

# **Non-parametric Statistical Methods**

APPLICATIONS IN MALDI IMAGING AND FINANCE

*A thesis submitted in partial fulfilment of the requirements  
for the degree of Dr. rer. nat.*

*submitted by*

**Jonathan von Schroeder**

Working Group Mathematical Statistics  
Institute for Statistics  
Faculty 3: Mathematics and Computer Science  
University of Bremen

First reviewer: [Prof. Dr. Thorsten Dickhaus](#)  
Second reviewer: [Assoc.-Prof. Dr. Taras Bodnar](#)

Date of defense: 28.01.2022

DOI: [10.26092/elib/1415](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0011-7)

*Science is a wonderful thing  
if one does not have to earn one's living at it.*

Albert Einstein



# CONTENTS

<b>Summary</b>	<b>ix</b>
<b>Zusammenfassung</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Efficient Calculation of the Joint Distribution of Order Statistics</b>	<b>7</b>
Declaration of Individual Contributions . . . . .	8
2.1 Introduction . . . . .	8
2.2 Order Statistics . . . . .	9
2.3 The Generalized Recursions . . . . .	11
2.3.1 Generalization of Bolshev's Recursion . . . . .	11
2.3.2 Generalization of Steck's Recursion. . . . .	11
2.3.3 Generalization of Noe's Recursion . . . . .	12
2.3.4 Computational Complexity and Numerical Properties . . . . .	12
2.4 Exact Evaluation of Bolshev's and Steck's Recursion . . . . .	14
2.5 Faithfully Rounded Evaluation of Noe's Recursion . . . . .	15
2.6 Applications . . . . .	16
2.6.1 Power of Goodness-Of-Fit Tests Against Contamination Alternatives . . . . .	16
2.6.2 Stepwise Multiple Hypothesis Testing. . . . .	16
2.6.3 Sample Size Calculation for Studies with Multiple Endpoints . . . . .	21
2.7 Discussion . . . . .	24
Acknowledgments . . . . .	24
<b>3 Reverse Stress Testing in Skew-Elliptical Models</b>	<b>25</b>
Declaration of Individual Contributions . . . . .	26
3.1 Introduction . . . . .	26
3.2 Problem Setting . . . . .	28
3.2.1 Reverse Stress Testing . . . . .	28
3.2.2 The Family of (Skew-)Elliptical Distributions . . . . .	29
3.3 Theory of Reverse Stress Testing in (Skew-)Elliptical Models. . . . .	30
3.3.1 Reverse Stress Testing in Elliptical Models. . . . .	30

3.3.2	Reverse Stress Testing in Skew-Elliptical Models . . . . .	31
3.3.3	Numerical Illustration . . . . .	31
3.4	Confidence Regions for the Most Likely Scenario . . . . .	32
3.4.1	Elliptical Models . . . . .	32
3.4.2	Skew-Normal and Skew-T Models . . . . .	35
3.4.3	Application to a Currency Portfolio . . . . .	36
3.5	Conclusion . . . . .	42
	Acknowledgments . . . . .	42
<b>4</b>	<b>Statistical Methods With Applications to MALDI MSI Data</b>	<b>43</b>
4.1	Introduction . . . . .	44
4.2	Feature Selection Using Univariate Tests for Homogeneity . . . . .	45
4.2.1	The Mann-Whitney U Test . . . . .	45
4.2.2	Chatterjee's Rank Correlation $\xi$ . . . . .	46
4.2.3	Bootstrapping Chatterjee's Rank Correlation $\xi$ . . . . .	47
4.2.4	Computation and Asymptotics of U-statistics for Feature Selection . . . . .	48
4.2.5	Feature Selection Using False Discovery Rate Control . . . . .	52
4.3	Application to MALDI MSI Data . . . . .	53
4.4	Statistical Modelling of MALDI MSI Data . . . . .	54
4.4.1	Likelihood of the Weights of a Finite Mixture . . . . .	55
4.4.2	The Expectation Maximization (EM) Algorithm . . . . .	56
4.4.3	Estimation of Mixture Weights Using the EM algorithm . . . . .	57
4.4.4	Proposed Mixture Model and Initial Estimation . . . . .	59
4.4.5	Improvement of the Initial Estimates Using EM Steps . . . . .	59
4.5	Conclusion . . . . .	61
<b>5</b>	<b>Conclusion</b>	<b>65</b>
	<b>Acknowledgements</b>	<b>67</b>
	<b>Bibliography</b>	<b>68</b>
<b>A</b>	<b>Appendix for Chapter 2</b>	<b>79</b>
A.1	When is the Result Faithfully Rounded? . . . . .	79
A.2	Multiple Hypothesis Testing . . . . .	80
A.3	Further Proofs and Algorithms . . . . .	82

<b>B</b>	<b>Appendix for Chapter 3</b>	<b>87</b>
B.1	Conditional Density . . . . .	87
B.2	Proofs . . . . .	87
B.2.1	Reverse Stress Testing in Elliptical Models . . . . .	88
B.2.2	Reverse Stress Testing in Skew-Elliptical Models . . . . .	89
B.2.3	Reverse Stress Testing in Skew-Normal Models . . . . .	92
B.2.4	Reverse Stress Testing in Skew-t Models . . . . .	94
B.2.5	Confidence Region for Elliptical Models . . . . .	98
B.3	Results on Method of Moment Estimation for the Skew-normal and Skew-t Distribution . . . . .	99
<b>C</b>	<b>Appendix for Chapter 4</b>	<b>101</b>
C.1	Proofs . . . . .	101
C.2	Calculation of U-statistics based on Chatterjee's Rank Correlation $\xi$ . . . . .	104
C.3	Generating Random Bits with Non-Uniform Probability . . . . .	105
C.3.1	Simple Implementation Using 64 Bits of Randomness . . . . .	106
C.3.2	More Efficient Utilisation of the Source Random Number Generator . . . . .	109
C.4	Efficient Calculation of $D_n$ and $\tau^*_{*n}$ . . . . .	111





# SUMMARY

During my time as a PhD student I was fortunate to have the opportunity to contribute to different topics in statistics. This is reflected in this thesis, which contains applications of (non-)parametric statistical methods (and the development of such techniques) with applications to three distinct topics:

- Chapter 2 contains a contribution to computational statistics. Its main topic is the (efficient and exact) calculation of the joint distribution of order statistics. Since ranks are fundamental to many statistical methods, these have many applications, some of which are detailed in Section 2.6.
- Chapter 3 contains a contribution to mathematical finance. It expands on the topic of my master thesis and contains results related to the topic of "reverse stress testing" which, roughly speaking, has the goal of performing a data-driven selection of likely scenarios for which a given portfolio exceeds a specified loss. Two notable contributions are the development of non-parametric confidence regions in elliptical models and a characterisation of the subspace which, in skew-elliptical models, contains the sought scenario.
- Chapter 4 contains contributions to mathematical statistics and some statistical insights into the analysis of biomedical images. It, among other things, contains results on statistical tests based on correlation coefficients when one of the random variables is a binary random variable. The derived results are utilised to elucidate some statistical properties of matrix-assisted laser desorption/ionization (MALDI) mass spectroscopy data.

In my work on all of these topics, my focus was on developing and applying statistical methods that are based only on the absolutely necessary assumptions. This is, of course, an aspirational goal. I am, however hopeful that I was able to make my own small contribution to the science of mathematical statistics.

Verden, 10.10.2021

J. von Schroeder



# ZUSAMMENFASSUNG

Während meiner Zeit als PhD-Student hatte ich die Möglichkeit, mich mit verschiedenen statistischen Themen zu beschäftigen. Dies spiegelt sich in dieser Thesis wieder, welche die Entwicklung und Anwendung (nicht-)parametrischer statistischer Methoden in drei verschiedenen Anwendungsgebieten darstellt:

- Kapitel 2 enthält numerische Methoden zur effizienten und exakten Berechnung der gemeinsamen Verteilung von Ordnungsstatistiken. Da Ränge fundamental für eine Vielzahl statistischer Verfahren sind, haben diese Methoden viele Anwendungen, von denen einige in Abschnitt 2.6 dargestellt werden.
- Kapitel 3 enthält einen Betrag im Bereich Finanzmathematik. Es handelt sich dabei um eine Weiterentwicklung eines Themas aus meiner Masterarbeit zum Thema 'Reverse Stress Testing'. Dabei geht es um die datenbasierte Auswahl von wahrscheinlichen Szenarien, unter denen ein vorgegebenes Portfolio hohe Verluste erleidet. Zwei relevante Kontributionen sind die Entwicklung von nicht-parametrischen Konfidenzregionen in elliptischen Modellen sowie die Charakterisierung des Unterraums, in dem sich das Szenario in Schief-Elliptischen Modellen befindet.
- Kapitel 4 enthält Beiträge zur mathematischen Statistik und einige statistische Erkenntnisse im Bereich der Analyse biomedizinischer Bilder. Dieses Kapitel enthält, unter anderem, Resultate über statistische Tests die auf Korrelationskoeffizienten basieren, wenn eine der Zufallsvariablen binär ist. Diese werden dazu verwendet, die statistischen Eigenschaften von Massenspektroskopie-Daten (die mit dem Verfahren 'Matrix-unterstützte Laser-Desorption/Ionisation (MALDI)' gemessen wurden) zu untersuchen.

Bei meiner Arbeit an all diesen Themen lag mein Fokus auf der Entwicklung und Anwendung statistischer Methoden, die nur auf den unbedingt notwendigen Annahmen beruhen. Ich bin hoffnungsvoll, dass ich meinen eigenen kleinen Beitrag zur Wissenschaft der mathematischen Statistik leisten konnte.

Verden, 10.10.2021

J. von Schroeder



# 1

## INTRODUCTION

*All models are approximations.  
Assumptions, whether implied or clearly stated, are never exactly true.  
All models are wrong, but some models are useful.*

George Box

*The properties of methods for statistical inference depend on the statistical model for the population. Sometimes prior knowledge can be utilized to justify restrictive assumptions that simplify inference. If, however, not many assumptions are justifiable, then the models' parameter space can be infinite-dimensional. In this setting, the usual, likelihood-based inference breaks down, and so called non-parametric methods are often a preferable alternative. This chapter starts with an instructive example and then describes the topology of measure spaces to explain, why likelihood based inference can break down and why the non-parametric methods used/presented in the following chapters are relevant.*

Given, for example, an observation  $y \in \mathbb{R}^n$  (of a random variable  $Y$ ) statistical inference can only be made if some model is, either explicitly or implicitly, assumed. The choice of model determines which procedures are valid, how their results can be interpreted and, finally, if a given procedure is even relevant for answering the question at hand. To illustrate this point consider, for example, the question which of two groups has, generally, the larger responses. Formally, denote by  $\mathbb{P}_Y$  the law of a random variable  $Y$  and assume that  $Y_1, \dots, Y_{n_0}$  are an independent and identically distributed (i.i.d.) sample from  $\mathbb{P}_{Y_1}$  and that  $Y_{n_0+1}, \dots, Y_n$  is an i.i.d. sample from  $\mathbb{P}_{Y_{n_0+1}}$ . Two popular procedures for statistical inference in this setting are the two-sample t-test (cf. [FP10, section 5.2] with  $Z_i = 0$  if  $i \leq n_0$  and  $Z_i = 1$  otherwise) and the Wilcoxon-Mann-Whitney U-test<sup>1</sup>. However, even the (seemingly simple) question 'Are the group means equal?' is quite hard, even if  $\mathbb{P}_{Y_1}$  and  $\mathbb{P}_{Y_{n_0+1}}$  are assumed to be Gaussian:

**Example 1** (Behrens-Fisher Problem). Denote by

$$\mathcal{P}_{\text{Behrens-Fisher}} := \{\mathbb{P}_{\mu_0, \mu_1, \sigma} = \mathcal{N}(\mu, \text{diag}(\sigma)) : (\mu_0, \mu_1, \sigma) \in \mathcal{F}_{\text{Behrens-Fisher}}\}$$

the set of Gaussian probability distributions (on  $\mathbb{R}^n$ ) with diagonal covariance matrices and equal within-group means and variances, and where the parameter space is given by

$$\mathcal{F}_{\text{Behrens-Fisher}} := \{(\mu_0, \mu_1, \sigma) : \sigma_0, \sigma_1 \in \mathbb{R}_{>0}^n \wedge \mu_0, \mu_1 \in \mathbb{R} \wedge \mu \in (\{\mu_0\}^{n_0} \times \{\mu_1\}^{n-n_0}), \sigma \in (\{\sigma_0\}^{n_0} \times \{\sigma_1\}^{n-n_0})\}.$$

Then for  $\mathbb{P}_Y \in \mathcal{P}_{\text{Behrens-Fisher}}$  the Behrens-Fisher problem is to test

$$H_0 : \mathbb{P}_Y \in \{\mathbb{P}_{\mu_0, \mu_1, \sigma} \in \mathcal{P}_{\text{Behrens-Fisher}} : \mu_0 = \mu_1\} \quad \text{vs} \quad H_1 : \mathbb{P}_Y \in \{\mathbb{P}_{\mu_0, \mu_1, \sigma} \in \mathcal{P}_{\text{Behrens-Fisher}} : \mu_0 \neq \mu_1\}$$

based on a sample  $(Y_1, \dots, Y_n) \sim \mathbb{P}$ . Neither the t-test nor the Wilcoxon-Mann-Whitney U-test are valid for this inference problem (cf. [FP10, Table 1, Perspective 14]).

If one restricts the model to  $\sigma_1 = \dots = \sigma_n$ , then the t-test is known to be the uniformly most powerful unbiased test (cf. [LR05, Problem 5.5]). Furthermore for  $\sigma_1 = \dots = \sigma_{n_0}$ ,  $\sigma_{n_0+1} = \dots = \sigma_n$  it has long been known that an exact test is possible (cf. [Kab66]).

Since the Wilcoxon-Mann-Whitney U-test is not based on a quantity related to the means, the model  $\mathcal{P}_{\text{Behrens-Fisher}}$  is not necessarily appropriate. In fact the test statistic is a plug-in estimate of the quantity

$$\phi(\mathbb{P}_{Y_1}, \mathbb{P}_{Y_{n_0+1}}) := \mathbb{P}(Y_1 > Y_{n_0+1}) + \frac{1}{2} \mathbb{P}(Y_1 = Y_{n_0+1})$$

if  $Y_1, \dots, Y_{n_0}$  and  $Y_{n_0+1}, \dots, Y_n$  are i.i.d. samples. The test is, if  $Y_1, \dots, Y_{n_0}$  and  $Y_{n_0+1}, \dots, Y_n$  are i.i.d. samples from continuous distributions, a valid test for the null-hypothesis  $H_0 : \mathbb{P}_{Y_1} = \mathbb{P}_{Y_{n_0+1}}$  against the alternative  $H_1 : \phi(\mathbb{P}_{Y_1}, \mathbb{P}_{Y_{n_0+1}}) \neq \frac{1}{2}$ ; cf. [FP10, Table 1, Perspective 3]. Unfortunately this model is hard to justify in practice since it does include distributions for which  $\phi(\mathbb{P}_{Y_1}, \mathbb{P}_{Y_{n_0+1}}) = \frac{1}{2}$  and  $\mathbb{P}_{Y_1} \neq \mathbb{P}_{Y_{n_0+1}}$ . A more helpful (even though sometimes too restrictive) view of the test is as one for stochastic ordering; cf. [FP10, Perspective 2]. For both of these models the t-test is not valid (cf. [FP10], see this source for more details on other relevant models).

As a side note, even though the one-sample t-test is not usually described as a non-parametric procedure, it is actually asymptotically maximin (in non-parametric models and for testing the mean) as long as the model is for i.i.d. observations (i.e. the probability measures are of the form  $\bigotimes_{i=1}^n \mathbb{P}$ ) and contains the family  $(\mathcal{N}(\theta, 1))_{\theta \in \mathbb{R}_{\geq 0}}$ ; cf. [LR05, Section 13.6.1]. Therefore the relevance of a given procedure should always be judged with a specific model in mind. While "there is not unanimity

<sup>1</sup>For the case when no ties are present see Section 4.2.1 with  $X_i = 0$  if  $i \leq n_0$  and  $X_i = 1$  otherwise. For the exact test in case of ties see [Wil45].

among statisticians in their use of the terms nonparametric and distribution-free" (cf. [Spr07]), I advocate for using the term "non-parametric" to describe statistical procedures that are successfully applied to infinite dimensional models (e.g. the space of all probability measures on  $\mathbb{R}$  with continuous distributions).

All of this demonstrates that the model plays a crucial role in determining the validity of a statistical procedure. Unfortunately the models for which desired properties of a given statistical procedure are mathematically guaranteed, are not necessarily models that agree with our preconceived knowledge of the data we wish to analyse. Thus model simplicity is a desirable criterion since it helps avoid nonintuitive or overly restrictive assumptions. Indeed, according to [Box76] "to devise simple but evocative models is the signature of the great scientist" and "overelaboration and overparameterization is often the mark of mediocrity."

But how can one avoid "overelaboration and overparameterization" when spaces of measures, that are not finitely supported, are notoriously difficult to handle mathematically? Over the years statisticians have developed many methods that are called "non-parametric" or "distribution-free". I consider these to be powerful tools to strive for this goal and have explored their applications in this thesis. Examples of this include the empirical likelihood method (cf. Section 3.4.1), the Mann-Whitney U test (cf. Section 4.2.1) and methods based on order statistics (a topic to which Chapter 2 is devoted).

The remainder of this chapter is devoted to describing the topology of measure spaces to elucidate some of their properties and the problems this causes in the design of inference methods.

Let  $(\Omega, \mathcal{A})$  be a measurable space, that is  $\Omega$  is a set and  $\mathcal{A} \subset \mathcal{P}(\Omega)$  is a  $\sigma$ -algebra; cf. [Bog07, Def. 1.2.2. and Def. 1.2.3.]. A countably additive (real valued) measure is a mapping  $\mu : \mathcal{A} \rightarrow \mathbb{R}$  such that

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

for any  $(A_i)_{i \in \mathbb{N}} \subseteq \mathcal{A}$  (cf. [Bog07, Def. 1.3.2.]). Every set  $A \in \mathcal{A}$  is called  $\mathcal{A}$ -measurable. In the following I will, for notational convenience, suppress the  $\mathcal{A}$  in my notation and, for example, just write measurable. I will however always assume that  $\Omega$  is equipped with a  $\sigma$ -algebra  $\mathcal{A}$ .

Furthermore the usual notions of absolute continuity and mutual singularity (of measures), written as  $\mu_1 \ll \mu_2$  and  $\mu_1 \perp \mu_2$  respectively, as well as atoms and atomicity will be adopted; cf. e.g. [Bog07, Def. 3.2.1.], [Bog07, p. 55].

Denote by  $\mathcal{M}(\Omega)$  the space of all measures on  $(\Omega, \mathcal{A})$ . Every measure  $\mu \in \mathcal{M}(\Omega)$  has a (unique) Jordan decomposition  $\mu = \mu^+ - \mu^-$  where  $\mu^+$  and  $\mu^-$  are positive measures, and a measure  $\tilde{\mu}$  is called positive iff  $\tilde{\mu}(A) \geq 0$  for all  $A \in \mathcal{A}$ . The total variation of  $\mu \in \mathcal{M}(\Omega)$  is, in terms of the Jordan decomposition, given by

$$\|\mu\|_{TV} := \mu^+(\Omega) + \mu^-(\Omega)$$

and  $(\mathcal{M}(\Omega), \|\cdot\|)$  is a Banach space; cf. [Bog07, p. 176] and [Bog07, Thm. 4.6.1]. The real vector space  $\mathcal{M}(\Omega)$  contains the convex cone (cf. [Zal02, p. 1])

$$\mathcal{M}_+(\Omega) := \{\mu \in \mathcal{M}(\Omega) : \mu \text{ is a positive measure}\}$$

of positive measures.  $\mathcal{M}_+(\Omega)$  is a cone because  $\mathbb{R}_+ \cdot \mathcal{M}_+(\Omega) \subseteq \mathcal{M}_+(\Omega)$ , but it is not a vector space since  $\mathbb{R} \cdot \mathcal{M}_+(\Omega) \not\subseteq \mathcal{M}_+(\Omega)$ .

The following lemma characterizes inner points of  $\mathcal{M}_+(\Omega)$ .

**Lemma 1.** A measure  $\mu \in \mathcal{M}_+(\Omega)$  is not in the topological interior  $\text{int}(\mathcal{M}_+(\Omega))$  of  $\mathcal{M}_+(\Omega)$  iff for every  $\varepsilon > 0$  there exists  $A_\varepsilon \in \mathcal{A}$  such that  $\mu(A_\varepsilon) \leq \varepsilon$  and  $A_\varepsilon \neq \emptyset$ .

*Proof.* Assume that such a measurable set exists for every  $\varepsilon$  and select for each some (arbitrary but fixed)  $\omega_\varepsilon \in A_\varepsilon$ . Then for every fixed  $\varepsilon > 0$  the measure  $\tilde{\mu} : \mathcal{A} \rightarrow \mathbb{R}$ , given by  $\tilde{\mu}(A) := \mu(A \setminus A_{\frac{\varepsilon}{2}}) - \frac{\varepsilon}{2} \cdot \chi(\omega_{\frac{\varepsilon}{2}} \in A)$ , is not an element of  $\mathcal{M}_+(\Omega)$  since  $\tilde{\mu}(A_{\frac{\varepsilon}{2}}) = -\frac{\varepsilon}{2} < 0$ . But

$$\|\mu - \tilde{\mu}\|_{TV} = \mu(A_{\frac{\varepsilon}{2}}) + \frac{\varepsilon}{2} \leq \varepsilon$$

and thus it is within the  $\varepsilon$ -neighbourhood of  $\mu$ . However, since  $\varepsilon > 0$  was arbitrary, this implies that  $\mu$  is not an interior point of  $\mathcal{M}_+(\Omega)$ .

Conversely assume that there exists  $\varepsilon > 0$  such that, for every non-empty set  $A \in \mathcal{A}$ , it holds that  $\mu(A) > \varepsilon$ . Then for every measure  $\tilde{\mu} \notin \mathcal{M}_+(\Omega)$  there exists a non-empty  $\tilde{A} \in \mathcal{A}$  such that  $\tilde{\mu}(\tilde{A}) < 0$ . But therefore it follows that

$$\|\mu - \tilde{\mu}\|_{TV} \geq \mu(\tilde{A}) + \tilde{\mu}(\tilde{A}) \geq \mu(\tilde{A}) > \varepsilon$$

and thus  $\tilde{\mu}$  is not an element of the  $\varepsilon$ -neighbourhood of  $\mu$ . But since  $\varepsilon > 0$  was arbitrary,  $\mu$  must be an inner point of  $\mathcal{M}_+(\Omega)$ .  $\square$

It remains to characterize exactly when the conditions in Lemma 1 can be fulfilled:

- If  $\mathcal{A}$  is finite (which is always the case when  $\Omega$  is finite), then  $\mu \in \text{int}(\mathcal{M}_+(\Omega))$  iff  $\mu(A) > 0$  for all measurable  $A \neq \emptyset$  and clearly such measures exist.
- If  $\mathcal{A}$  is not finite, then this condition is always fulfilled: It can be easily shown, by contradiction, that an infinite  $\sigma$ -algebra contains a pairwise disjoint sequence  $(A_n)_{n \in \mathbb{N}} \subseteq \mathcal{A}$  of non-empty sets. But since every  $\mu \in \mathcal{M}(\Omega)$  is finite, it holds that

$$\mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \sum_{i \in \mathbb{N}} \mu(A_i) \leq \mu(\Omega) < \infty$$

and a necessary condition for the convergence of  $\sum_{i \in \mathbb{N}} \mu(A_i)$  is that  $\lim_{i \rightarrow \infty} \mu(A_i) = 0$ . Thus for every  $\varepsilon > 0$  there exists an  $i \in \mathbb{N}$  such that  $\mu(A_i) \leq \varepsilon$ .

Thus, for  $1 \leq p < \infty$ ,  $\text{int}(\mathcal{M}_+(\Omega)) \neq \emptyset$  if either  $\Omega$  is finite or  $\mathcal{A}$  is finite and thus a very coarse  $\sigma$ -algebra. This is problematic since even the Gateaux derivative is only defined on open sets, but  $\mathcal{M}_+(\Omega)$  is not open in the topology of  $(\mathcal{M}(\Omega), \|\cdot\|_{TV})$  unless  $\text{int}(\mathcal{M}_+(\Omega)) \neq \emptyset$ . Furthermore, considering only the space  $\mathcal{M}_+(\Omega)$  (equipped with some other topology) is not an option since it is only a cone and not a vector space. But derivatives (of some kind) are fundamental to asymptotic results for likelihood-based inference since these results typically rely on a (local) first-order Taylor expansion; cf. [GN16, p. 551 ff.]. All of this is not to say that likelihood inference in this setting is impossible, but rather that it does not have the usual, desirable properties for many practically relevant cases.

For a measure  $\mu_0 \in \mathcal{M}_+(\Omega)$  denote by

$$\mathcal{M}(\Omega, \mu_0) := \left\{ A \mapsto \int_A \phi d\mu_0 : \phi \in L^1(\Omega, \mu_0) \right\}$$

the space of measures absolutely continuous wrt.  $\mu_0$  and where  $L^p(\Omega, \mu_0)$  is the space of functions such that  $f \in L^p(\Omega, \mu_0)$  iff  $|f|^p$  is  $\mu_0$ -integrable (or more precisely the factor-space wrt. the equivalence relation of functions being equal  $\mu_0$ -a.e., cf. [Bog07, p. 249 ff.]). That  $\mathcal{M}(\Omega, \mu_0) \subset \mathcal{M}(\Omega)$



is indeed the space of all  $\mu_0$ -dominated measures in  $M(\Omega)$  is an immediate consequence of the Radon-Nikodym theorem (cf. [Bog07, Thm. 3.2.2.]). For  $1 \leq p \leq \infty$  the spaces  $L^p(\Omega, \mu_0)$  are Banach spaces (cf. [Bog07, Thm. 4.1.3.]) when equipped with the usual norms and

$$i_{\text{can}} : M(\Omega, \mu) \rightarrow L^1(\Omega, \mu)$$

$$\mu_1 \mapsto \frac{d\mu_1}{d\mu},$$

which maps a measure  $\mu_1$  to its Radon-Nikodym density, is an isometry of Banach spaces; cf. [Ay+17, p. 140]. By the Hölder-inequality  $L^q(\Omega, \mu_0) \subseteq L^p(\Omega, \mu_0)$  if  $1 \leq p \leq q \leq \infty$ .

Denote by  $L_+^p(\Omega, \mu_0) := \{f \in L^p(\Omega, \mu_0) : A \mapsto \int_A f d\mu_0 \text{ is a positive measure}\}$  the convex cone of  $\mu_0$ -a.e. positive densities in  $L^p(\Omega, \mu_0)$ . To understand when the interior of  $L_+^p(\Omega, \mu_0)$  is empty, it is helpful to consider the main difference to  $M(\Omega)$ : For  $\mu \in L_+^p(\Omega, \mu_0)$  the measure  $\tilde{\mu}$  used in the proof need not be an element of  $L^p(\Omega, \mu_0)$ . For  $1 \leq p < \infty$ , this is exactly the case when  $\tilde{\mu}$  is not absolutely continuous wrt.  $\mu_0$  which happens only if  $A_\varepsilon$  is a  $\mu_0$  null set. Thus the question arises when it is impossible to pick  $A_\varepsilon$  with  $\mu_0(A_\varepsilon) > 0$ . It turns out that this is exactly the case when  $\mu_0$  is a purely atomic measure with only finitely many distinct atoms:

**Definition 1** (cf. [Bog07, Definition 1.12.7.]). *A  $\mathcal{A}$  is called an atom of  $\mu$  if  $\mu(A) > 0$  and for every  $B \in \mathcal{A}$  with  $B \subset A$  it holds that  $\mu(B) \in \{0, \mu(A)\}$ . Two atoms  $A_1, A_2$  are called distinct if  $\mu(A_1 \triangle A_2) > 0$  where  $A \triangle B := (A \cup B) \setminus (A \cap B)$  denotes the symmetric difference. If  $\{A_n\}$  is the (at most countable) set of distinct atoms, then  $\mu$  is called purely atomic if  $\mu(\Omega \setminus \bigcup_{n=1}^\infty A_n) = 0$ . If  $\mu$  has not atoms, then it is called atomless.*

The distinct atoms of a measure  $\mu$  can be chosen to be pairwise disjoint:

*Proof.* Let  $\{A_n\}$  be the (at most countable) set of distinct atoms of  $\mu$  and let  $\tilde{A}_i := A_i \setminus \left(\bigcup_{j=1}^{i-1} A_j\right)$ . The  $\tilde{A}_i$  are measurable since  $\tilde{A}_i = A_i \cap \left(\Omega \setminus \left(\bigcup_{j=1}^{i-1} A_j\right)\right)$  and obviously  $\bigcup_{n=1}^\infty \tilde{A}_n = \bigcup_{n=1}^\infty A_n$ . It remains to show that all of these are still atoms, that is  $\mu(\tilde{A}_i) > 0$  for all  $i \in \mathbb{N}$ , or equivalently that  $\mu(A_i \setminus \tilde{A}_i) = 0$ . To this end notice that

$$\mu(A_i \setminus \tilde{A}_i) = \mu\left(A_i \cap \left(\bigcup_{j=1}^{i-1} A_j\right)\right) \leq \sum_{j=1}^{i-1} \mu(A_i \cap A_j) = 0$$

which yields the desired result.  $\square$

Using this characterization of atoms it is easy to establish the following result:

**Lemma 2.** *Let  $\mu \in M(\Omega)$ . Then the following are equivalent:*

- *There exists  $\varepsilon > 0$  such that no  $A \in \mathcal{A}$  satisfies  $0 < \mu(A) \leq \varepsilon$*
- *$\mu$  is purely atomic with finitely many distinct atoms.*

*Proof.* By [Joh70, Thm. 2.1] there exist measures  $\mu_1, \mu_2 \in M(\Omega)$  such that  $\mu_1$  is purely atomic and  $\mu_2$  is atomless. If  $\mu_2$  is not the zero measure, then  $\mu$  is not purely atomic and by [Bog07, Cor. 1.12.10.] there exists, for every  $\varepsilon > 0$ ,  $A \in \mathcal{A}$  with  $\mu(A) = \min\{\mu(\Omega), \varepsilon\} > 0$ . If  $\mu_2$  is the zero measure it remains to show that the such an  $\varepsilon$  can only exist if  $\mu = \mu_1$  has only finitely many distinct atoms:

- If  $\mu$  has only finitely many distinct atoms  $A_1, \dots, A_k$  then, by definition,  $\varepsilon := \frac{1}{2} \min_i \mu(A_i)$  is such that all measurable sets either have a measure bigger than  $\varepsilon$  or are null sets.

- If  $\mu$  has infinitely many distinct (and pairwise disjoint) atoms  $(A_i)_{i \in \mathbb{N}}$  then, since  $\mu$  is finite, it follows that  $\sum_{i \in \mathbb{N}} \mu(A_i) = \mu(\Omega) < \infty$  and thus  $\lim_{i \rightarrow \infty} \mu(A_i) = 0$ . Thus for every  $\varepsilon > 0$  there exists an  $i \in \mathbb{N}$  such that  $\mu(A_i) \leq \varepsilon$ . But  $A_i$  is an atom and thus  $\mu(A_i) > 0$  which concludes the proof.

□

Thus, the interior of  $L_+^p(\Omega, \mu_0)$  is non-empty iff  $\mu_0$  takes only finitely many distinct values; cf. [Ay+17, p. 143, Footnote 4].

**Remark 1.** For  $L^\infty(\Omega, \mu_0)$  the situation is very different: It can contain open subsets of positive densities (cf. [GN16, Proposition 7.2.4]) enabling the definition of a differentiable likelihood and, for example, over a Sobolev ball the asymptotic normality of a non-parametric maximum-likelihood estimator can be established (cf. [GN16, Lemma 7.2.13]). This approach is, however, only mentioned for the sake of completeness as it cannot (in general) be extended to all of  $L^1(\Omega)$ .

For statistical purposes, usually the space of probability measures  $\mathcal{P}(\Omega) := \{\mu \in \mathcal{M}_+(\Omega) : \|\mu\|_{TV} = 1\}$  on  $(\Omega, \mathcal{A})$  is considered, but since  $\mathcal{P}(\Omega) \subseteq \mathcal{M}_+(\Omega)$  the previous discussion applies to  $\mathcal{P}(\Omega)$  as well.

# 2

## EFFICIENT CALCULATION OF THE JOINT DISTRIBUTION OF ORDER STATISTICS

**Jonathan von SCHROEDER, Thorsten DICKHAUS**

*Science is what we understand well enough to explain to a computer.  
Art is everything else we do.*

Donald E. Knuth

*The problem of computing the joint distribution of order statistics of stochastically independent random variables in one- and two-group models is considered. While recursive formulae for evaluating the joint cumulative distribution function of such order statistics exist, their numerical implementation remains a challenging task. This task is tackled by presenting novel generalizations of known recursions. They are utilized to obtain exact results (calculated in rational arithmetic) as well as faithfully rounded results. Finally, some applications in goodness-of-fit testing, step-wise multiple hypothesis testing, and sample size calculation for studies with multiple endpoints are discussed.*

## DECLARATION OF INDIVIDUAL CONTRIBUTIONS

This chapter is based on a research question posed by Prof. Dr. Thorsten Dickhaus. The proofs and algorithms for sections 2.3-2.5, all code needed to produce the graphics/tables as well as the implementation of the proposed methods are entirely my own work. The introduction as well as the applications discussed in section 2.6 are joint work with Prof. Dr. Thorsten Dickhaus to which both of us contributed equally.

Prof. Dr. Thorsten Dickhaus supported me in terms of presentation and writing in all parts of the manuscript. I am very thankful for all the fruitful discussions I had with him regarding all parts of the paper.

This chapter has been published in Computational Statistics & Data Analysis [vT20].

## 2.1. INTRODUCTION

THE joint distribution of order statistics  $X_{1:n}, \dots, X_{n:n}$  of stochastically independent random variables  $X_1, \dots, X_n$  plays a pivotal role in the theory of empirical processes and in nonparametric statistics; see, e. g., [SW09] and [Dic18]. For instance, the exact finite-sample null distributions of classical goodness-of-fit tests like the Kolmogorov-Smirnov and the Cramér-von Mises test as well as those of modern "higher criticism" goodness-of-fit tests rely on such joint distributions; cf. [GLF15], [GLF16], and [FG18] for recent developments and further references. In simultaneous statistical inference, the joint distribution of ordered  $p$ -values is needed to analyze the type I and type II error behaviour of stepwise rejective multiple test procedures; cf. Chapter 5 of [Dic14].

One important application area is sample size calculation for (clinical) studies with multiple end-points. Regulatory agencies have recognized that the type II error can be inflated if the planned sample size is not adjusted for multiplicity (see e.g. the guideline on Points to Consider on Multiplicity Issues in Clinical Trials by the CPMP (Committee for Proprietary Medicinal Products); [Med02]). This has led to the development of statistical methods to tackle this issue (see e.g. [SB07], [Ham+13], [VTS17]). Particularly challenging cases occur when the false discovery rate (FDR) is chosen as the type I error criterion and stepwise rejective multiple test procedures (operating on ordered test statistics or  $p$ -values, respectively) are utilized (see e.g. [Jun05], [FZ06], [Glu+08a] and [lzm18]). The mentioned results in the literature are either asymptotic (in the number of null hypotheses to be tested simultaneously) or limited to a small number of null hypotheses. Inexact power calculations can however lead to either too large samples (and therefore unnecessarily high costs) or underpowered (and therefore potentially futile) studies. Both can also be problematic from the ethical perspective. It is therefore desirable to have exact power calculations beyond the number of hypotheses feasible using the approach of [Glu+08a].

In the case that  $X_1, \dots, X_n$  are identically distributed (we refer to this case as a one-group model), classical recursive methods like Bolshev's recursion, Noe's recursion, and Steck's recursion allow for computing the joint cumulative distribution function (cdf) of  $X_{1:n}, \dots, X_{n:n}$  exactly; cf. Section 9.3 of [SW09]. A generalization of Steck's recursion to two-group models has been introduced by [Bla+14]. The other aforementioned recursions can be generalized in an analogous manner, as we will demonstrate in Section 2.3 of the present work.

While conceptually appealing, numerical properties of the aforementioned recursions are not well understood yet, and existing implementations into computer software often refer to rule-of-thumb-type upper bounds on  $n$  such that the respective implementation is trustworthy. For example, Art B. Owen reports in his implementation of the two-sided version of Noe's recursion in C (see <https://www.stat.washington.edu/jaw/RESEARCH/SOFTWARE/BERKJONES/BJ-RBJ-C-Code/>

noe.c) that the recursion works well for  $n \leq 1000$  but "For larger  $n$  (eg 1800 or more) [...] unexplained odd behavior." Similarly, in the R Package `mutoss` (cf. [Bla+10]) the following comment is made on the implementation of Bolshev's recursion: "Because of numerical issues  $n$  should not be greater than 100." Recently, [MN17] introduced a computational method for one-group models. However, they do not consider the numerical accuracy of their approach rigorously.

In this work, we contribute to the analysis of the numerical accuracy and the computational complexity of existing approaches for computing the joint distribution of  $X_{1:n}, \dots, X_{n:n}$  in a mathematically rigorous manner. Furthermore, we provide novel computational techniques for one- and two-group models which are guaranteed to provide accurate results for arbitrary sample size  $n$ . The rest of the material is structured as follows. In Section 3.2, we introduce the relevant quantities. The (generalized) recursions for one- and two-group models are provided in Section 2.3, together with a rigorous analysis of their computational complexities and their numerical properties. Our proposed exact computational methods rely on rational arithmetic (Section 2.4) and on faithful rounding (Section 2.5), respectively. Applications in power analysis, multiple hypothesis testing, and sample size planning are given in Section 2.6. We conclude with a discussion in Section 2.7. Lengthy proofs as well as pseudo code for the considered algorithms are deferred to the appendix.

## 2.2. ORDER STATISTICS

**T**HROUGHOUT the following sections, we let  $[n] := \{1, 2, \dots, n\}$  for a natural number  $n \in \mathbb{N}$ . Consider stochastically independent, real-valued random variables  $X_1, \dots, X_n$ , which are all driven by the same probability measure  $\mathbb{P}$ . Let  $I_1 := [n]$ , and recursively define

$$\begin{aligned} i_j &:= \min \{i \in I_j \mid \forall k \in I_j : X_i \leq X_k\} \\ I_j &:= I_{j-1} \setminus \{i_{j-1}\} \end{aligned}$$

for  $j \in [n]$ . Then we call  $X_{i_1}, \dots, X_{i_n}$  the order statistics of  $X_1, \dots, X_n$ , which we will denote by  $X_{1:n}, \dots, X_{n:n}$  in the remainder. The random variable  $X_{i:n}$  will be called the  $i$ -th order statistic of the random vector  $(X_1, \dots, X_n)^\top$ . Let  $F_i$  denote the marginal cdf of  $X_i$  for  $i \in [n]$ . This paper will present methods for the quick and numerically stable calculation of

$$\Psi_{n_1, n_2}^{G_1, G_2}(\mathbf{b}) := \mathbb{P}(X_{1:n} \leq b_1, \dots, X_{n:n} \leq b_n), \quad \mathbf{b} = (b_1, \dots, b_n)^\top \in \mathbb{R}^n,$$

assuming that  $\forall i \in [n] : F_i \in \{G_1, G_2\}$  where  $G_1, G_2$  are two continuous distribution functions on  $\mathbb{R}$  and with  $n_i = \left| \{j \in [n] : F_j = G_i\} \right|$  denoting the number of  $X_j$ 's distributed according to  $G_i$ ,  $i = 1, 2$ . Since it holds that

$$G_1(X_i) \sim \begin{cases} \text{Uni}[0, 1], & F_i = G_1, \\ G_2 \circ G_1^{-1}, & F_i = G_2, \end{cases}$$

it follows that  $\Psi_{n_1, n_2}^{G_1, G_2} = \Psi_{n_1, n_2}^{\text{Uni}[0, 1], F} \circ G_1$ , where  $F := G_2 \circ G_1^{-1}$  and  $\text{Uni}[0, 1]$  denotes the uniform distribution on the interval  $[0, 1]$ . Therefore, it is sufficient to consider the calculation of  $\Psi_{n_1, n_2}^{\text{Uni}[0, 1], F}(\mathbf{b})$  for an arbitrary continuous distribution function  $F : [0, 1] \rightarrow [0, 1]$  and argument  $\mathbf{b} \in [0, 1]^n$ . In the sequel, we suppress the dependence on  $F$  and  $\mathbf{b}$  notationally, and write  $\Psi(n_1, n_2) := \Psi_{n_1, n_2}^{\text{Uni}[0, 1], F}(\mathbf{b})$  for notational convenience.

As outlined in the introduction, for  $n_2 = 0$  there exist many well known recursions (see, e. g., Section 9.3 of [SW09]) for computing  $\Psi(n_1, n_2)$ . There are also newer approaches based on numerical integration (see [MNS16]) or based on the Poisson process (see [MN17]). Unfortunately, the former

cannot be easily generalized to the case  $0 < n_2 < n$ : For the one group case the density of the order statistics of i.i.d. uniform random variables is given by

2

$$f(x_1, \dots, x_n) = \begin{cases} n! & \text{if } 0 \leq x_1 \leq \dots \leq x_n \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Hence, it is piece-wise constant and comparatively simple to numerically integrate. Unfortunately for the multi-group case the density is not so simple anymore (cf. [BB89, Equation (1)]) since it involves the calculation of matrix-permanents. Thus one can probably do no better than the algorithm proposed by [Glu+08b] which has exponential  $O(n^n)$  complexity (cf. [Glu+08b, Theorem 4.2]), resulting in a very high computational effort for even moderate values of  $n$ . However, the method of [Glu+08b] can be used to compute  $k$ -variate marginal distributions for  $k \ll n$ , because in such cases the complexity of their approach reduces to  $O(n^k)$ . The approach based on the Poisson process is very fast due to the usage of the Fourier transform, but numerically unstable for small values of the  $b_i$ 's. This can for instance be demonstrated using the thresholds of the well-known linear step-up test (cf. [BH95]) for control of the FDR; see Figure 2.1.

Since we are mostly concerned with the full joint distribution, we extend the approach suggested by [Bla+14] and provide generalizations of Bolshev's and Noe's recursions. We compare them to the generalization of Steck's recursion proposed by [Bla+14] and demonstrate that Bolshev's and Steck's recursion are suitable for exact computations in rational arithmetic, whereas Noe's recursion is numerically stable when computed in standard (IEEE 754) floating point arithmetic.

All our numerical calculations were performed on a Windows 10 machine with an AMD Ryzen™ 5 2600 CPU (6 cores) with 16 gigabytes of RAM. Parallelization was implemented using Intel® Threading Building Blocks (TBB).

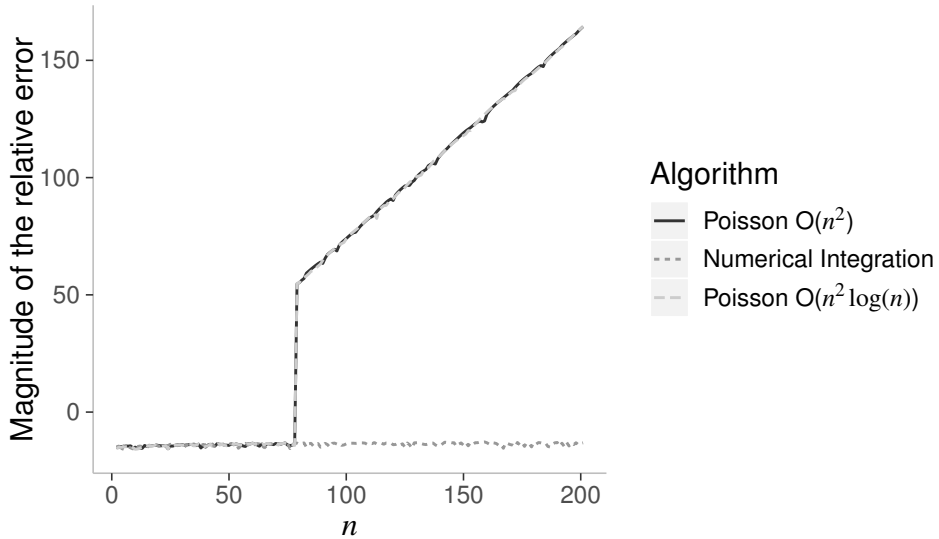


Figure 2.1: Relative error (on the  $\log_{10}$  scale) of the methods presented by [MNS16] and [MN17], respectively, when calculating  $\Psi(n, 0)$  for the thresholds  $b_i \equiv b_i^{(n)} := 0.05 \times i/n$ . A value of  $-16$  implies at least 15 accurate non-zero digits in base 10. For  $n \leq 77$  the relative error of the three methods is visually barely distinguishable. For  $n \geq 77$  the dotted line below zero corresponds to the "Numerical Integration".

## 2.3. THE GENERALIZED RECURSIONS

Let  $n := n_1 + n_2$ ,  $n_1, n_2 \in \mathbb{N}$ . Furthermore, let  $X_1, \dots, X_{n_1} \sim \text{Uni}[0, 1]$  and  $X_{n_1+1}, \dots, X_{n_1+n_2} \sim F$  be jointly stochastically independent, where  $F$  is continuous and supported on  $[0, 1]$ . Let

$$\Psi(n_1, n_2) := \mathbb{P}(X_{1:n} \leq b_1, \dots, X_{n:n} \leq b_n), \quad (2.1)$$

where  $(b_i)_{i \in [n]}$  is an increasing sequence with values in  $[0, 1]$ . The following subsections provide formulas for efficiently calculating  $\Psi(n_1, n_2)$ , and we discuss their computational and numerical properties.

### 2.3.1. GENERALIZATION OF BOLSHEV'S RECURSION

**Lemma 3** (Generalization of Bolshev's Recursion). *The function  $\Psi$  from (2.1) satisfies the recursion*

$$\Psi(m_1, m_2) = 1 - \sum_{\substack{0 \leq k_1 \leq m_1 \\ 0 \leq k_2 \leq m_2 \\ k_1 + k_2 < m_1 + m_2}} M_{k_1, k_2}^{(m_1, m_2)} \cdot \Psi(k_1, k_2),$$

where

$$M_{k_1, k_2}^{(m_1, m_2)} := \binom{m_1}{k_1} \binom{m_2}{k_2} (1 - b_{k_1 + k_2 + 1})^{m_1 - k_1} \cdot (1 - F(b_{k_1 + k_2 + 1}))^{m_2 - k_2}. \quad (2.2)$$

Moreover, we have the following recursive relationships for  $M$ .

$$M_{k_1, k_2}^{(m_1 + 1, m_2)} = \begin{cases} 1 & k_2 = m_2 \wedge k_1 = m_1 + 1 \\ M_{m_1, k_2 + 1}^{(m_1, m_2)} \cdot \frac{k_2 + 1}{m_2 - k_2} \cdot (1 - F(b_{m_1 + (k_2 + 1) + 1})) & k_2 < m_2 \wedge k_1 = m_1 + 1 \\ M_{k_1, k_2}^{(m_1, m_2)} \cdot \frac{(m_1 + 1)}{m_1 + 1 - k_1} \cdot (1 - b_{k_1 + k_2 + 1}) & \text{otherwise,} \end{cases}$$

$$M_{k_1, k_2}^{(m_1, m_2 + 1)} = \begin{cases} 1 & k_1 = m_1 \wedge k_2 = m_2 + 1 \\ M_{k_1 + 1, m_2}^{(m_1, m_2)} \cdot \frac{k_1 + 1}{m_1 - k_1} \cdot (1 - b_{k_1 + m_2 + 2}) & k_1 < m_1 \wedge k_2 = m_2 + 1 \\ M_{k_1, k_2}^{(m_1, m_2)} \cdot \frac{(m_2 + 1)}{m_2 + 1 - k_2} \cdot (1 - F(b_{k_1 + k_2 + 1})) & \text{otherwise.} \end{cases}$$

For  $n_1 = 0$  or  $n_2 = 0$  this is simply the well-known Bolshev recursion.

### 2.3.2. GENERALIZATION OF STECK'S RECURSION

**Lemma 4** (Generalization of Steck's Recursion). *Let  $b_0 := 0$ . Then  $\Psi$  from (2.1) satisfies the recursion*

$$\Psi(m_1, m_2) = (b_{m_1 + m_2})^{m_1} F(b_{m_1 + m_2})^{m_2} - \sum_{\substack{0 \leq k_1 \leq m_1 \\ 0 \leq k_2 \leq m_2 \\ k_1 + k_2 \leq m_1 + m_2 - 2}} M_{k_1, k_2}^{(m_1, m_2)} \cdot \Psi(k_1, k_2), \quad (2.3)$$

where

$$M_{k_1, k_2}^{(m_1, m_2)} := \binom{m_1}{k_1} \binom{m_2}{k_2} (b_{m_1 + m_2} - b_{k_1 + k_2 + 1})^{m_1 - k_1} (F(b_{m_1 + m_2}) - F(b_{k_1 + k_2 + 1}))^{m_2 - k_2}.$$

Letting

$$a(k, j) := \binom{k}{j} \text{ and } a(k, j) = \begin{cases} 1 & k = j \\ \frac{j+1}{k-j} \times a(k, j+1) & j < k, \end{cases} \quad (2.4)$$

we can write

$$M_{0, j}^{(m_1, m_2)} = a(m_2, j) \cdot (b_{m_1 + m_2} - b_{j+1})^{m_1} (F(b_{m_1 + m_2}) - F(b_{j+1}))^{m_2 - j}, \quad (2.5)$$

$$M_{j,m_2}^{(m_1,m_2)} = a(m_1, j) \cdot (b_{m_1+m_2} - b_{j+m_2+1})^{m_1-j}. \quad (2.6)$$

Furthermore, we have the following recursion for  $M$ .

$$M_{k_1+1,k_2-1}^{(m_1,m_2)} = M_{k_1,k_2}^{(m_1,m_2)} \times \frac{F(b_{m_1+m_2}) - F(b_{k_1+k_2+1})}{b_{m_1+m_2} - b_{k_1+k_2+1}} \times \frac{m_1 - k_1}{k_1 + 1} \times \frac{m_2 - k_2 + 1}{k_2}$$

for  $0 \leq m_1 \leq n_1$  and  $0 \leq m_2 \leq n_2$ .

*Proof.* See [Bla+14, Proposition 1] □

### 2.3.3. GENERALIZATION OF NOE'S RECURSION

**Lemma 5** (Generalization of Noe's Recursion). *Let  $b_0 := 0$ ,  $Q_{0,0}(0) := 1$ ,  $Q_{i_1,i_2}(1) := b_1^{i_1} \cdot F(b_1)^{i_2}$  and for  $m > 1$*

$$Q_{i_1,i_2}(m) := \sum_{\substack{0 \leq k_1 \leq i_1 \\ 0 \leq k_2 \leq i_2 \\ m-1 \leq k_1+k_2}} M_{k_1,k_2}^{i_1,i_2}(m) \cdot Q_{i_1-k_1,i_2-k_2}(m-1),$$

$$M_{k_1,k_2}^{i_1,i_2}(m) := \binom{i_1}{k_1} \binom{i_2}{k_2} \times (b_m - b_{m-1})^{i_1-k_1} \times (F(b_m) - F(b_{m-1}))^{i_2-k_2}$$

for  $0 \leq i_1 \leq n_1, 0 \leq i_2 \leq n_2, m \leq i_1 + i_2 \leq n$ .

Then the function  $\Psi$  from (2.1) satisfies

$$\Psi(i_1, i_2) = Q_{i_1,i_2}(i_1 + i_2)$$

for  $i_1 \leq n_1$  and  $i_2 \leq n_2$ .

Letting

$$a^{(m),1}(j) := (b_m - b_{m-1})^j \quad \text{and} \quad a^{(m),2}(j) := (F(b_m) - F(b_{m-1}))^j,$$

we can write

$$M_{k_1,k_2}^{i_1,i_2}(m) = \binom{i_1}{k_1} \binom{i_2}{k_2} \times a^{(m),1}(i_1 - k_1) \times a^{(m),2}(i_2 - k_2). \quad (2.7)$$

### 2.3.4. COMPUTATIONAL COMPLEXITY AND NUMERICAL PROPERTIES

The computational complexity (defined to be the number of elementary arithmetic operations on floating point numbers) of each of the aforementioned recursions is given by the following lemma.

**Lemma 6.** *The proposed recursions can be implemented using*

**Bolshev**  $O(n_1^2 n_2^2)$

**Steck**  $O(n_1^2 n_2^2 \log_2(n_1 n_2))$

**Noe**  $O(n_1^2 n_2^2 (n_1 + n_2))$

elementary arithmetic operations (addition, subtraction, multiplication, division). Furthermore  $O(n_1 n_2)$  memory is needed (assuming fixed-precision storage of all results).



The results of Lemma 6 suggest that Noe's recursion might not be the best choice. However, for small values of the  $b_i$ 's, Bolshev's recursion and Steck's recursion are inherently numerically unstable. Consider for example  $n_1 = 11, n_2 = 0$  and

$$b_i := \begin{cases} 2^{-10} & \text{if } i \leq 10, \\ 2^{-1} & \text{if } i = 11. \end{cases}$$

Then both recursions, when implemented in double precision floating point arithmetic, result in negative values and huge relative errors (cf. Table 2.1), which can be explained by inaccurate rounding and/or catastrophic cancellation, respectively.

Table 2.1: Calculation of the probability that  $n = 11$  uniform order statistics are (component-wise) bounded from above by  $\mathbf{b} \in [0, 1]^{11}$ , where  $b_i = 2^{-10}$  if  $i \leq 10$  and  $b_{11} = 2^{-1}$ . The rows give the intermediate steps of the recursions. The first column reports the first few non-zero digits of the exact probabilities (computed in rational arithmetic), the second column reports the steps of Steck's recursion (calculated in double precision floating point arithmetic) and the fourth column reports the steps of Bolshev's recursion (also calculated in double precision floating point arithmetic). In the third and the fifth column the relative error of the intermediate value is reported.

Exact Probability	Steck	Rel. Err. (Steck)	Bolshev	Rel. Err. (Bolshev)
9.76562E-04	9.76562E-04	0.00000E+00	9.76562E-04	0.00000E+00
9.53674E-07	9.53674E-07	0.00000E+00	9.53674E-07	0.00000E+00
9.31323E-10	9.31323E-10	0.00000E+00	9.31323E-10	0.00000E+00
9.09495E-13	9.09495E-13	0.00000E+00	9.09495E-13	0.00000E+00
8.88178E-16	8.88178E-16	0.00000E+00	8.88178E-16	0.00000E+00
8.67362E-19	8.67362E-19	0.00000E+00	1.73472E-18	1.00000E+00
8.47033E-22	8.47033E-22	0.00000E+00	1.10114E-20	1.20000E+01
8.27181E-25	8.27181E-25	0.00000E+00	6.85071E-21	8.28100E+03
8.07794E-28	8.07794E-28	0.00000E+00	-2.70517E-20	3.34884E+07
7.88861E-31	7.88861E-31	0.00000E+00	-1.12683E-16	1.42842E+14
4.33103E-30	-1.75898E-20	4.06134E+09	2.83880E-16	6.55456E+13

For the example depicted in Table 2.1 Noe's recursion implemented in double precision arithmetic does not result in numerical errors, and we have therefore omitted this column. In general Noe's recursion, if implemented in a reasonable manner, never results in negative values. Furthermore by [JR18, Equation (3)] the relative error is bounded (if the coefficients are computed with a bounded relative error) since all summands are non-negative.

**Remark 2.** Noe's recursion can be easily parallelized since the  $Q_{i_1, i_2}(m)$ 's appearing in Lemma 5 can be, for any fixed  $m$ , computed in parallel.

By examining Algorithm 1 we can get a sense of the constants involved in the computational complexity of Bolshev's recursion for the case where  $n_2 = 0$ , and hence  $n_1 = n$ : Even though Bolshev's recursion involves binomial coefficients our proposed Algorithm 1 (cf. the appendix) for the one-group case evaluates it using only

$$\# \text{Operations} = n + (n + 1) + \sum_{k=2}^n \left[ 2 + \sum_{j=1}^{k-1} 6 \right] = 3n^2 + n - 1$$

elementary arithmetic operations (addition, subtraction, multiplication, division).

For the general case  $n_1, n_2 > 0$  our proposed Algorithm 2 (cf. the appendix) implements the two-group case in  $O(n_1^2 n_2^2)$  elementary arithmetic operations. Consequently, for equal sample sizes  $n_1 = n_2 = \ell$  the number of operations is of  $O(\ell^4)$ . Notice that this is a marked improvement over the exponential complexity  $\ell^\ell$  reported by [Glu+08b, Theorem 4.2]. Figure 2.2 illustrates the observed execution time for calculating  $f(\ell) := \Psi(\ell, \ell)$  and  $b_i \equiv b_i^{(n)} := 0.05 \times i/n$ , where  $n = n_1 + n_2 = 2\ell$ .

For not necessarily equal sample sizes  $n_1 \geq 1$  and  $n_2 \geq 1$ , our implementation of Algorithm 2 (Bolshev's generalized recursion) needs

$$f(n_1, n_2) := 1.5 \cdot n_1^2 \cdot n_2^2 + 4.5 \cdot (n_1^2 \cdot n_2 + n_1 \cdot n_2^2) + 3 \cdot (n_1 + n_1^2 + n_2 + n_2^2) + 7.5 \cdot n_1 \cdot n_2 + 2$$

arithmetic operations.

Our implementation of the generalization of Steck's recursion needs

$$f(n_1, n_2) := 0.5 \cdot (n_1^3 \cdot (n_2 + 1) + n_2^3 \cdot (n_1 + 1)) + 3 \cdot n_1^2 \cdot n_2^2 + 9.5 \cdot (n_1^2 \cdot n_2 + n_2^2 \cdot n_1) \\ + 6.5 \cdot (n_1^2 + n_2^2) - 5 \cdot (n_1 \cdot n_2 + n_1 + n_2) + 2$$

arithmetic operations. To be able to give an exact formula we did not utilize exponentiation by squaring (cf. [Knu98, p. 462, Algorithm A]) and therefore this is a useful upper bound for the optimal complexity.

Lastly our implementation of the generalization of Noe's recursion needs

$$f(n_1, n_2) := (1 + 1.25 \cdot n_2 + 0.25 \cdot n_2^2) \cdot n_1^3 + (1.5 + 4.25 \cdot n_2 + 3 \cdot n_2^2 + 0.25 \cdot n_2^3) \cdot n_1^2 \\ + (1.5 + 1.5 \cdot n_2 + 4.25 \cdot n_2^2 + 1.25 \cdot n_2^3) \cdot n_1 + n_2^3 + 1.5 \cdot n_2^2 + 0.5 \cdot n_2 \quad (2.8)$$

arithmetic operations.

## 2.4. EXACT EVALUATION OF BOLSHEV'S AND STECK'S RECURSION

**E**VALUATION of the recursions in floating point arithmetic can, as demonstrated in the previous section, lead to the accumulation of numerical errors. This can be improved by utilizing a higher precision floating point type, but it is not clear what precision of intermediate results is necessary to achieve the desired precision and the answer to this question will depend on the  $b_i$ 's. It is therefore desirable to have a method for obtaining exact results (either as a benchmark for numerical approximations or when the exact result is needed directly to, for example, investigate exact equalities).

If only elementary arithmetic operations are utilized it is possible to exactly evaluate expressions using rational arithmetic (our code uses the [GNU Multiple Precision Arithmetic Library](#)). This is the case for the (generalized) recursions described in the previous section. Since it has clearly the highest computational complexity we do not recommend exact calculations using Noe's recursion. Even though one might expect Bolshev's recursion to result in shorter run times than Steck's recursion (due to the theoretical computational complexity) it turns out that Steck's recursion can be faster in a practical implementation (see Figure 2.2). Since we observed an even larger relative advantage (in computation time) when utilizing floating point arithmetic, we suspect that this has to do with the fact that our implementation of the generalized Steck recursion does not need to keep track of full matrices of coefficients (as is done in our implementation of the generalized Bolshev's recursion) and is therefore likely to be more cache (and register) friendly.

Unfortunately the cdf  $F$  of many interesting distributions is not available in a closed form and thus the thresholds  $(F(b_i))_{i \in [n]}$  might either not be exactly calculable or simply not exactly representable as rational numbers. Lemma 7 analyzes the error propagation when  $F$  and / or  $\mathbf{b}$  are inexact.

**Lemma 7.** Let

$$x_i := \begin{cases} b_1 & \text{if } i = 1 \\ b_i - b_{i-1} & \text{if } 1 < i \leq n \\ F(b_1) & \text{if } i = n + 1 \\ F(b_i) - F(b_{i-1}) & \text{if } n + 1 < i \leq 2n \end{cases} \quad (2.9)$$

and denote by  $(\tilde{x}_i)_{i \in [n]}$  approximations thereof, which are obtained by replacing  $(b_i)_{i \in [n]}$  and  $F$  by approximations  $(\tilde{b}_i)_{i \in [n]}$  and  $\tilde{F}$ . If for  $\varepsilon \in (0, 1)$  it holds that  $\forall i \in [2n] : x_i \in (1 - \varepsilon, 1 + \varepsilon)$ , then it follows that for all  $(i_1, i_2) \in [n_1] \times [n_2]$

$$\tilde{\Psi}(i_1, i_2) \in \Psi(i_1, i_2) \cdot \left( (1 - \varepsilon)^{i_1+i_2}, (1 + \varepsilon)^{i_1+i_2} \right),$$

where  $\tilde{\Psi}$  denotes the approximation of  $\Psi$  obtained by using  $\tilde{F}$  and  $\tilde{b}_i$  instead of  $F$  and  $b_i$ ,  $i \in [n]$ .

We conclude this section by giving a short description of rational arithmetic and the practical aspects of evaluating the recursions using rational arithmetic. In rational arithmetic (rational) numbers are stored as a tuple of numerator and denominator. Numerator and denominator are usually represented as infinite precision integers. By this term we denote integer data types that use as much storage as necessary to represent all digits without any loss of precision. These are for example implemented in the [GNU Multiple Precision Arithmetic Library](#). After arithmetic operations the numerator and denominator are normalized (where the normalized rational is defined as the unique representation such that numerator and denominator have no (non-trivial) common factors) by utilizing the extended Euclidean algorithm. The relevant algorithms can be, for example, found in [\[Knu98, Chapter 4\]](#). Due to the (potentially dynamically adjusted) storage of the infinite precision integers as well as the use of the Euclidean algorithm the basic arithmetic operations like addition and multiplication cannot (in general) be considered to be constant time anymore. Instead the computational complexity can depend on the actual numbers the operation is performed on.

## 2.5. FAITHFULLY ROUNDED EVALUATION OF NOE'S RECURSION

UNLESS the exact value of  $\Psi(n_1, n_2)$  has less than 16 (base 10) non-zero digits an exact evaluation in double precision floating point arithmetic is not possible. The best one can hope for is a faithfully rounded result. That is, the result is either exact (if the exact value is a double precision floating point number) or it is one of the two closest floating point numbers. This is the best one can achieve because the double precision floating point numbers are in essence a discrete space that contains only some real numbers.

[\[LR17\]](#) describe algorithms that can perform multiple arithmetic operations such that the result is faithfully rounded. We have implemented these algorithms as a portable single-header C++11 library.<sup>1</sup> Utilizing this library we implemented the generalization of Noe's recursion presented in Section 2.3.3 obtaining faithfully rounded results if no underflow occurs. In our experience this is usually the case if the values of  $\Psi$  are not too close to the smallest (in absolute value) normal double, which equals  $2^{-1022} \approx 2.225 \cdot 10^{-308}$  on most computer architectures. In case of an underflow the results are smaller than the true values of  $\Psi$ , but never less than zero. Figure 2.2 compares the runtime of our implementation of our generalization of Noe's recursion to that of the algorithm from the previous section. It becomes apparent that Noe's recursion with faithful rounding is much faster than Bolshev's and Steck's recursion implemented in rational arithmetic. It furthermore incurs only a constant overhead (approximately five times when compared to evaluation in double precision floating point arithmetic) whereas the use of rational arithmetic can increase the computational complexity of an algorithm. For practical applications, we therefore recommend Noe's recursion with faithful rounding, at least if a fixed numerical precision is sufficient.

The reason why we can apply the technique of [\[LR17\]](#) to the evaluation of Noe's recursion, but not to that of the other recursions is, that the guarantee of faithful rounding only holds if the implemented algorithm satisfies the NIC principle (*No Inaccurate Cancellation*, cf. [\[LR17, Definition 2.2\]](#)):

<sup>1</sup>Available at <https://github.com/jvschroeder/PairArithmetic/>.

An algorithm satisfies the NIC principle if there are no sums where at least one summand is not an input to the algorithm and the summands have opposite signs.

2

This is satisfied by Noe's recursion, but not by the other recursions we presented. Furthermore a close examination of our implementation yields that, by [LR17, Theorem 4.2], the result will be faithfully rounded (assuming that no over- or underflow occurs) if  $n_1, n_2 \leq 8184$ . This can be improved significantly by changing a detail of the implementation, but at a big computational cost. For the details see A.1.

## 2.6. APPLICATIONS

### 2.6.1. POWER OF GOODNESS-OF-FIT TESTS AGAINST CONTAMINATION ALTERNATIVES

WE begin our examples by considering one of the most basic statistical test problems, namely, testing for goodness-of-fit: Given an independent sample  $Y_1, \dots, Y_n$ , where  $Y_i$  is distributed according to  $F_i$ , it is given by the pair of hypotheses

$$H_0 : F_i = G_0 \text{ for all } i \in [n] \quad \text{vs} \quad H_1 : F_i \neq G_0 \text{ for at least one } i \in [n],$$

where  $G_0$  is a known and continuous distribution function. The null hypothesis  $H_0$  is a homogeneity hypothesis in the sense of [Dic18, p. 51]. It is interesting to consider the power of a test of  $H_0$  against a specific, one-point, alternative. One such model is the contamination alternative  $\{F_1, \dots, F_n\} = \{G_0, G_1\}$  where the random variables belong to one of two groups. Denote by  $k = \{i \in [n] : F_i = G_0\}$  the number of random variables that belong to the first group. Then under  $H_0$  it holds that  $k = n$  and under the alternative  $k < n$ . If the goodness-of-fit test is based on the order statistics  $Y_{1:n}, \dots, Y_{n:n}$ , we can use the previously discussed results to calculate the power against a specific alternative given by a continuous cdf  $G_1$  and an integer  $k \in [n]$ . Consider, for example, the goodness-of-fit test with equal local levels proposed by [GLF16]. Then Figure 2.3 depicts the power of the test (for  $n = 100$  and at level  $\alpha = 0.025$ ) for  $G_0 = \mathcal{N}(0, 1), G_1 = \mathcal{N}(-1.1, 1)$  as a function of  $k$ . Figure 2.4 considers  $G_1 = \mathcal{N}(\mu_1, 1)$  and depicts the power of the test (for  $k = 65$  fixed) as a function of  $\mu_1$ .

### 2.6.2. STEPWISE MULTIPLE HYPOTHESIS TESTING

FOLLOWING [Bla+14] we consider  $m \geq 2$  null hypotheses  $H_1, \dots, H_m$  which are simultaneously under consideration under one and the same statistical model. We assume that associated  $p$ -values  $p_1, \dots, p_m$  are available on which the multiple test operates. Furthermore, we assume that  $p_1, \dots, p_m$  (regarded as random variables) are jointly distributed according to one of the following models

**FM**( $m, m_0, F$ ) The  $p_i$ 's are stochastically independent with marginal distributions

$$p_i \sim \begin{cases} \text{Uni}[0, 1] & \text{if } 1 \leq i \leq m_0, \\ F & \text{if } m_0 + 1 \leq i \leq m, \end{cases}$$

where  $m_0$  denotes the number of true null hypotheses among  $H_1, \dots, H_m$  and  $F$  is a given continuous cdf on  $[0, 1]$ .

**RM**( $m, \pi_0, F$ ) Let  $M_0$  denote a binomially distributed random variable,  $M_0 \sim \mathcal{B}(m, \pi_0)$ . Conditionally on  $M_0 = m_0$ , the  $p_i$ 's are jointly distributed according to  $\text{FM}(m, m_0, F)$ .

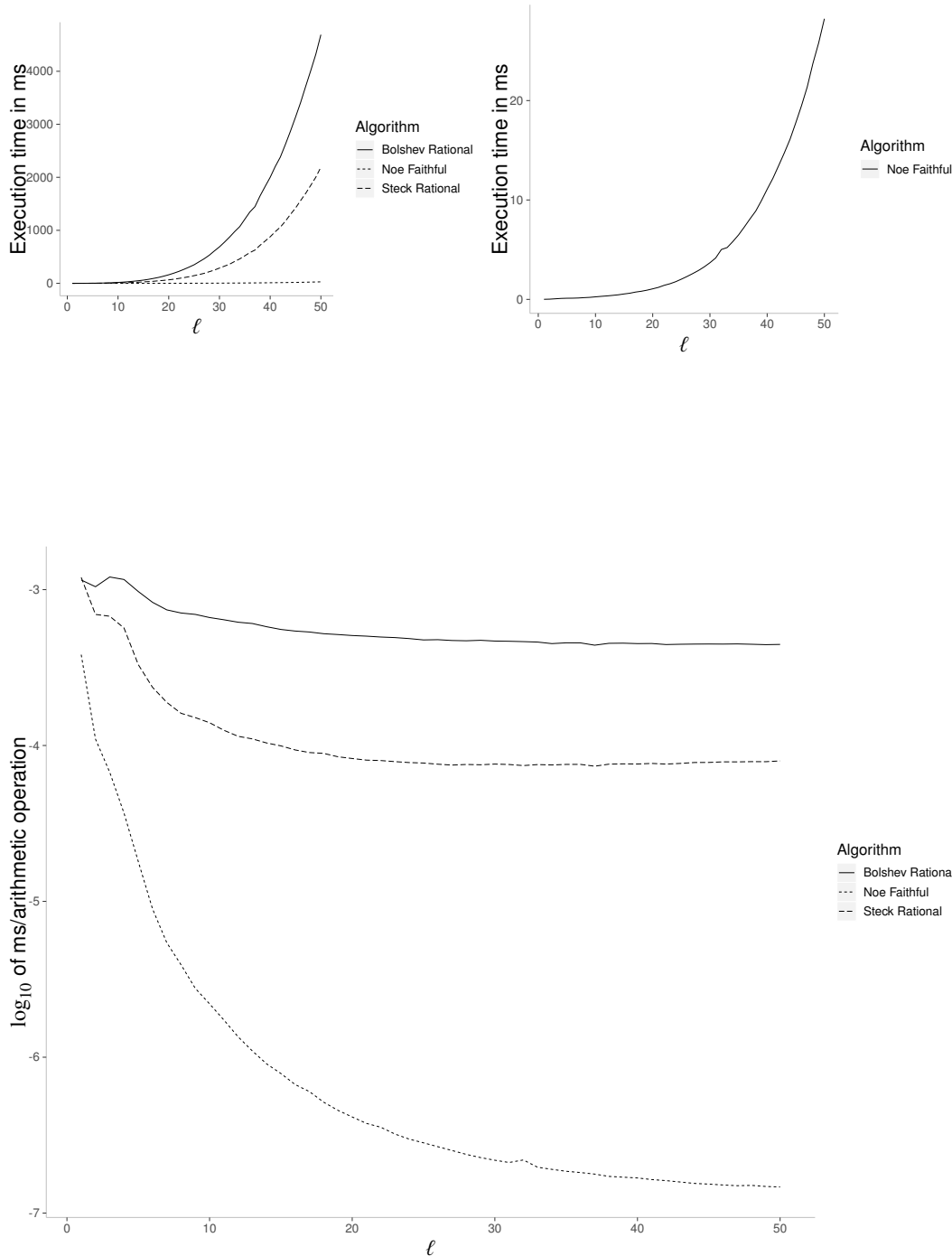


Figure 2.2: Comparison of the runtimes of Bolshev's and Steck's recursion (where  $n_1 = n_2 = \ell$ ) implemented in rational arithmetic with Noe's recursion implemented in faithfully rounded floating point arithmetic. To determine the time per arithmetic operation we simply divided by the theoretical number of arithmetic operations (cf. Section 2.3.4).

As discussed by [RV11], the joint distribution of the number of rejections  $R$  and the number of false rejections  $V$  under these models can be calculated exactly for step-up tests if one can calculate  $\Psi$  (for a summary of the relevant details see A.2).

The random variables  $V$  and  $R$  play an important role when analyzing the type I and type II error behavior of such multiple tests. One important observation is that the previously discussed recursions for calculating  $\Psi(m, m)$  also calculate all  $\Psi(i_1, i_2)$  (for  $0 \leq i_1, i_2 \leq m$ ) as intermediate results.

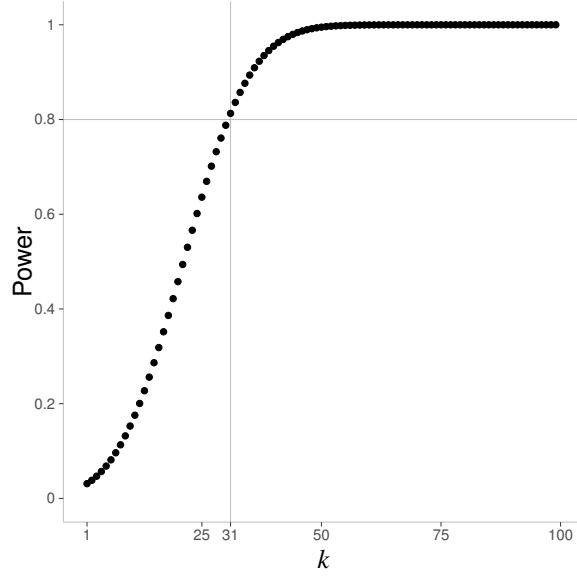


Figure 2.3: Power of the goodness-of-fit test with equal local levels (at significance level  $\alpha = 0.025$ ) proposed by [GLF16] against the contamination alternative that only  $k$  random variables are distributed according to  $G_0 = \mathcal{N}(0, 1)$  and the remaining  $n - k$  random variables are distributed according to  $G_1 = \mathcal{N}(-1.1, 1)$  where  $n = 100$ .

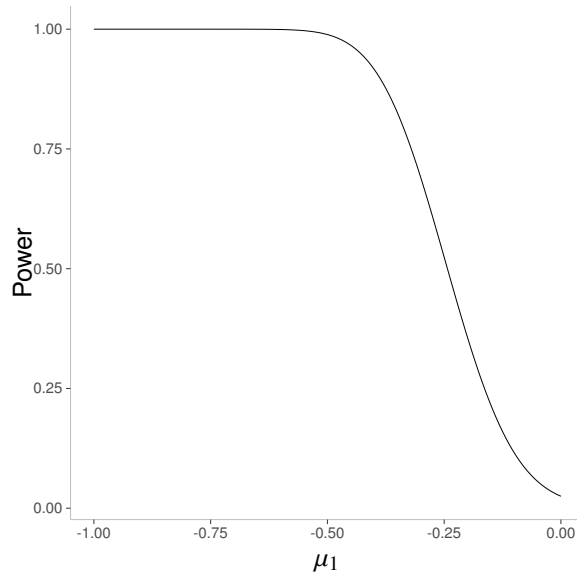


Figure 2.4: Power of the goodness-of-fit test with equal local levels (at significance level  $\alpha = 0.025$ ) proposed by [GLF16] against the contamination alternative that only  $k$  random variables are distributed according to  $G_0 = \mathcal{N}(0, 1)$  and the remaining  $n - k$  random variables are distributed according to  $G_1 = \mathcal{N}(\mu_1, 1)$  where  $n = 100$  and  $k = 65$ .

In order to provide some numerical illustrations, we first consider the average power (cf. (A.1)) under  $FM(m, m_0, F)$ , where

$$F(t) := 1 + \Phi\left(\Phi^{-1}\left(\frac{t}{2}\right) - \sqrt{N}\right) - \Phi\left(\Phi^{-1}\left(1 - \frac{t}{2}\right) - \sqrt{N}\right) \quad (2.10)$$

for  $t \in [0, 1]$  and  $N = 5$ . This is the setting considered in [Glu+08a, Table 2] where the average power for  $m \leq 5$  was calculated for the Benjamini-Hochberg procedure (controlling the FDR at  $\alpha = 0.05$ ) for  $m$  independent two-sided one sample z-tests. In our notation, the Benjamini-Hochberg (linear step-up test) procedure equals  $SU_t$  with  $t_i = i \times \alpha / m$  for  $i \in [m]$ . Table 2.2 illustrates the results obtained

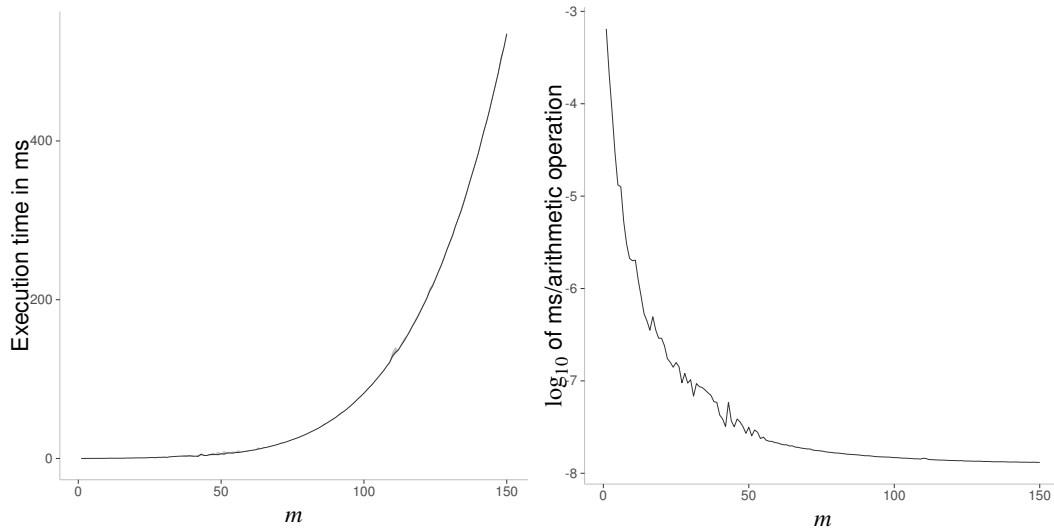


Figure 2.5: Time needed to calculate the average power of the Benjamini-Hochberg procedure for  $m$  hypotheses. To determine the time per arithmetic operation we simply divided by the theoretical number of operations given by (2.8). No adjustment was made for the small additional computational effort required for deriving the average power from the distribution of the order statistics.

for  $m, m_0 \leq 50$ . Due to space constraints only the first six columns (corresponding to  $m_0 \leq 5$ ) and some rows are presented. The calculation of the full table (not presented here) took less than a second for an  $m$  one magnitude larger (50 instead of 5) than the one considered by [Glu+08a]. Figure 2.5 illustrates the time needed to calculate one row of such a table corresponding to some  $m \in \mathbb{N}$  when utilizing our proposed algorithms.

Table 2.2: Average power of the Benjamini-Hochberg procedure (controlling the FDR at  $\alpha = 0.05$ ) for  $m$  independent two-sided one sample z-tests (sample size  $N = 5$ , common variance  $\sigma^2 = 1$ ,  $\mu_0 = 0, \mu_1 = 1$ , cf. (2.10)) when exactly  $m_0$  null hypotheses are true. The bold values are exactly those in [Glu+08a, Table 2].

$m$	$m_0 = 0$	$m_0 = 1$	$m_0 = 2$	$m_0 = 3$	$m_0 = 4$	$m_0 = 5$
2	<b>0.56539</b>	<b>0.50342</b>				
3	<b>0.54576</b>	<b>0.49842</b>	<b>0.44439</b>			
4	<b>0.53446</b>	<b>0.49583</b>	<b>0.45256</b>	<b>0.40451</b>		
5	<b>0.52712</b>	<b>0.49440</b>	<b>0.45819</b>	<b>0.41837</b>	<b>0.37494</b>	
6	0.52201	0.49357	0.46241	0.42840	0.39148	0.35175
7	0.51827	0.49310	0.46574	0.43606	0.40399	0.36955
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
45	0.49975	0.49513	0.49041	0.48559	0.48068	0.47565
46	0.49969	0.49516	0.49055	0.48584	0.48103	0.47612
47	0.49963	0.49520	0.49068	0.48607	0.48137	0.47657
48	0.49957	0.49523	0.49081	0.48629	0.48169	0.47700
49	0.49951	0.49526	0.49093	0.48651	0.48201	0.47741
50	0.49946	0.49529	0.49104	0.48672	0.48230	0.47781

As a second example, we consider the computation of the  $\lambda$ -power  $\text{Pow}_\lambda(\text{SU}_t)$  from (A.2). Again, we choose  $t$  as in the Benjamini-Hochberg case. The simulations in [lzm18, Table 3] consider a two-group model for the test statistics such that they are stochastically independent and follow a central t-distribution with  $\nu$  degrees of freedom under the null hypothesis. Under the alternative, they follow a non-central t-distribution with  $\nu$  degrees of freedom and non-centrality parameter  $\mu$ . We denote by  $F_{\nu, \mu}$  the (non-central) t-distribution with  $\nu$  degrees of the freedom and non-centrality parameter  $\mu$ . Thus (in our notation), the  $p$ -values are uniform under the null hypothesis

Table 2.3: Faithfully rounded calculation of the  $\lambda_{90}$ -power of the Benjamini Hochberg procedure (controlling the FDR at  $\alpha = 0.15$ ) when applied to  $m = 200$  test-statistics (and where we reject for large absolute values of the test-statistics which corresponds to two-sided tests) which follow a central t-distribution  $F_{v,0}$  under the null hypothesis and a non-central t-distribution  $F_{v,\mu}$  under the alternative, where  $v = 2n - 2$  and  $\mu = \sqrt{n/2}\theta$ . The results are obtained under the RM-model where  $\pi_0 = \mathbb{E}(M_m)/m$ . For comparison the fourth column contains the Monte Carlo estimates (sample size 1000) given in [lzm18, Table 3]. The following column gives the faithfully rounded power calculated using our method and the last column states the absolute difference between the previous two columns divided by the standard deviation of the Monte Carlo estimation (which was estimated using 300 replicates). As expected most Monte Carlo approximations are within one or two standard deviations of the faithfully rounded result. For the twelfth row no value is given since the estimated standard deviation was zero. This is not unexpected since the faithfully rounded result is approximately equal to  $4.58 \cdot 10^{-7}$  which is two orders of magnitude smaller than  $(300 \cdot 1000)^{-1}$ .

Eff Sz. $\theta$	$\mathbb{E}(M_m)$	n	est. $\lambda_{90}$ -pwr	$\lambda_{90}$ -pwr	Diff in std
0.60000	5	70	0.24900	0.26691	1.23987
0.60000	5	80	0.39600	0.39977	0.24081
0.60000	5	90	0.53800	0.53479	0.18527
0.60000	5	100	0.65700	0.65626	0.05057
0.60000	20	50	0.02800	0.02538	0.48379
0.60000	20	60	0.15700	0.13890	1.69096
0.60000	20	70	0.37800	0.36864	0.61103
0.60000	20	80	0.59900	0.62231	1.50514
0.60000	60	40	0.00200	0.00143	0.43439
0.60000	60	50	0.09900	0.08584	1.49443
0.60000	60	60	0.49200	0.49307	0.06522
0.60000	100	30	0.00000	0.00000	Inf
0.60000	100	40	0.00600	0.00658	0.22041
0.60000	100	50	0.27000	0.30726	2.80406
0.80000	5	40	0.25200	0.26037	0.59552
0.80000	5	50	0.50200	0.49870	0.19588
0.80000	5	60	0.70400	0.70951	0.39928
0.80000	20	30	0.03600	0.03969	0.60858
0.80000	20	40	0.35700	0.36732	0.64187
0.80000	60	20	0.00000	0.00004	0.18826
0.80000	60	30	0.14700	0.15775	0.88238
0.80000	100	20	0.00300	0.00013	8.59402
0.80000	100	30	0.50600	0.48563	1.42549
1.00000	5	30	0.39200	0.39421	0.14658
1.00000	20	20	0.04500	0.04534	0.05420
1.00000	20	30	0.61400	0.63660	1.40588
1.00000	60	20	0.22500	0.19941	2.00156
1.00000	100	20	0.58600	0.57569	0.64138

and distributed according to

$$F(t) := \bar{F}_{v,\mu}\left(\bar{F}_{v,0}^{-1}\left(\frac{t}{2}\right)\right) - F_{v,\mu}\left(-\bar{F}_{v,0}^{-1}\left(\frac{t}{2}\right)\right), t \in [0, 1], \quad (2.11)$$

under the alternative, where  $\bar{F}_{v,\mu} = 1 - F_{v,\mu}$  denotes the upper-tail cdf. [lzm18] provides asymptotic approximations of the  $\lambda$ -power. Our results can be used to calculate  $\text{Pow}_\lambda(\text{SU}_t)$  with high numerical precision. Table 2.3 gives the faithfully rounded values for the 0.9-power for the parameters considered in [lzm18, Table 3].

We conclude by giving an example for the exact distribution of the FDP which shows why the FDR is not always an appropriate summary statistic. Consider again the multiple two-sided z-test described in [Glu+08a], meaning that  $F$  is given by (2.10), for  $N = m = 50$  and  $m_0 = 5$ . It is clear that in Figure 2.6 the distribution of the FDP is neither symmetric about its mean (the FDR, which is depicted as dotted vertical line) nor concentrated below the FDR. A similar argumentation has been used by, among others, [Bla+14] and [DR15] in order to motivate the computation of the full distribution of the FDP and to control its quantiles. The latter task is inherently computationally demanding. Figure 2.7 demonstrates that there is a loss in accuracy when utilizing simple floating point arithmetic for Noe's recursion.



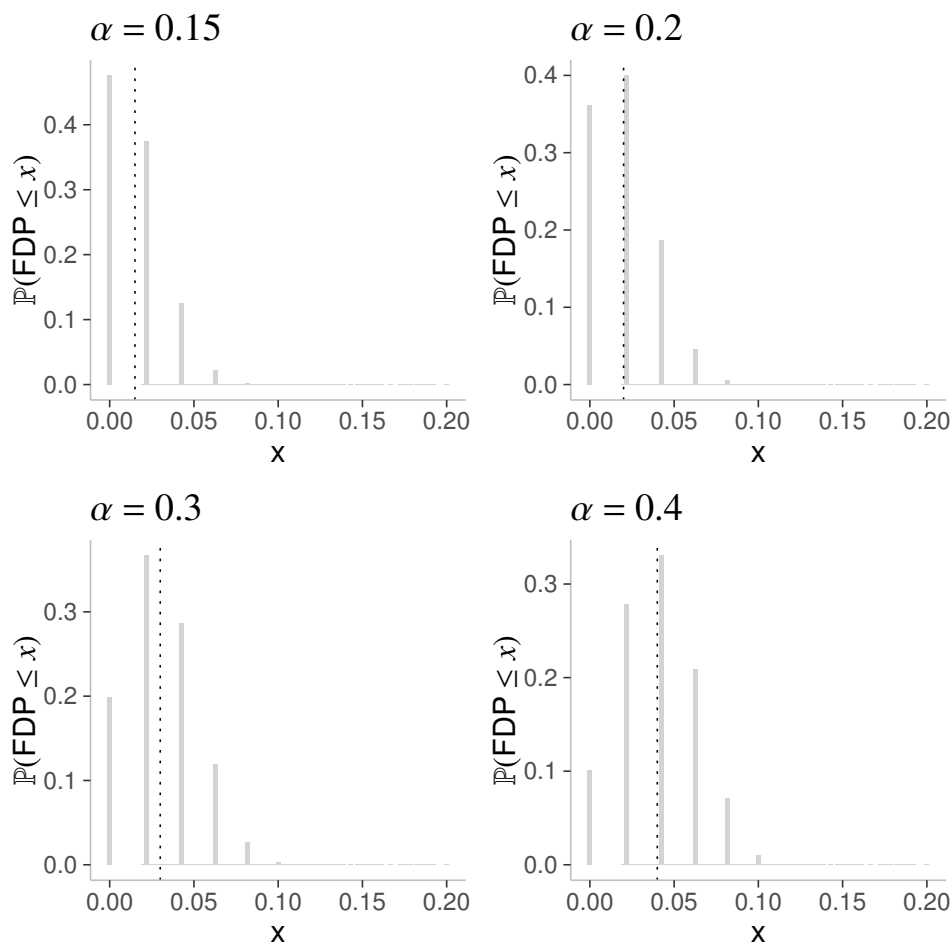


Figure 2.6: Distribution of the FDP for the Benjamini-Hochberg procedure (controlling the FDR at  $\alpha$ ) for  $m = 50$  independent two-sided one sample z-tests (sample size  $N = 50$ , common variance  $\sigma^2 = 1$ ,  $\mu_0 = 0, \mu_1 = 1$ , cf. (2.10)). The dotted vertical line is the actual FDR of the test.

### 2.6.3. SAMPLE SIZE CALCULATION FOR STUDIES WITH MULTIPLE ENDPOINTS

GENETIC association studies often (cf. [Dic14, Section 9.2]) consist of two stages: A screening and a validation stage. Both operate on independent samples. In the following we demonstrate how one can, based on the screening study, perform a sample size calculation for the validation study.

As an example we consider the study of [Not+01] which collected  $n = 18$  paired colon/adenocarcinoma samples.<sup>2</sup> They analyzed only the 4000 genes and ESTs with an average expression intensity of at least 10. Of these,  $m = 126$  genes and ESTs had an at least four-fold higher/lower average expression in cancerous tissue than in normal tissue. If one considers this to be a screening study and wishes to follow up with a validation stage, testing only these  $m = 126$  genes and ESTs, a sample size calculation is desirable:

What sample size would be necessary for a follow-up study to (likely) find most differentially expressed subscripts among the  $m = 126$  hypotheses when applying a **statistical** criterion?

<sup>2</sup>The data-set is available at <http://genomics-pubs.princeton.edu/oncology/>.

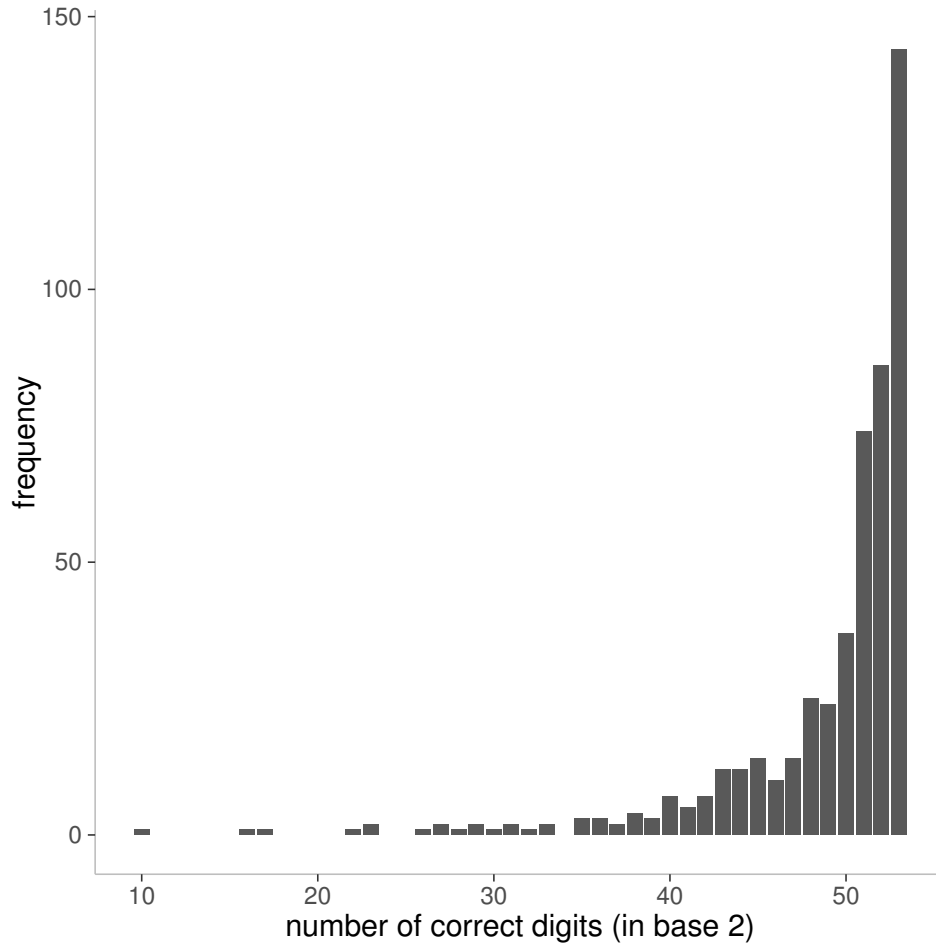


Figure 2.7: Number of correct digits (in base 2) when the FDP distributions in Figure 2.6 are calculated with Steck's recursion (using double precision floating point numbers). The correct result (as a double precision floating point number) was obtained using the faithfully rounded approach based on Noe's recursion. The presence of 53 correct digits in base 2 implies that there is no difference to the faithfully rounded result (due to the use of double precision floating point numbers). It is clear that in some cases only a few digits are correct.

As the statistical (type I error) criterion, we choose here control of the FDR at level  $\alpha$ , a standard criterion for large  $m$ . The  $p$ -values are obtained from a paired t-test. If we assume that they are stochastically independent and that the number of true hypotheses  $m_0$  is fixed (but unknown), then they are distributed according to  $FM(m, m_0, F)$  where  $F = F_{\nu, \mu}$  is given by (2.11), where  $\nu = n - 1 = 17$  and where the effect size  $\mu$  is unknown. It follows that the following steps are necessary for sample size planning:

- Find plausible values for  $\mu$ . This can, for example, be done by considering the realizations of the test statistics (in the study of [Not+01] the quartiles are 2.81 and 6.71).
- Find a plausible upper bound on the number of true hypotheses  $m_0$ . Remark 3 describes one possible approach.
- Choose an FDR-controlling procedure and an appropriate notion of power, e.g., the Benjamini-Hochberg procedure and the average power (cf. (A.1)).

The following remark demonstrates that, in this setting, we can recast finding a suitable upper bound for  $m_0$  as a numerical optimization problem:

**Remark 3.** Suppose we have a realization  $\tilde{p}$  of  $p$ -values from  $FM(m, m_0, F)$ , where  $m$  and  $F$  are known, but  $m_0$  is not. For an arbitrary  $\widetilde{m}_0 \in [m]$  let

$$f(\gamma, \widetilde{m}_0) := P_{FM(m, \widetilde{m}_0, F)}(\widetilde{m}_0 \leq m - \gamma R(SU_t, \mathbf{p}))$$

which is the probability that, under  $FM(m, \widetilde{m}_0, F)$ , the random quantity  $m - \gamma R(SU_t, \mathbf{p})$  is an upper bound for  $\widetilde{m}_0$ . Then we can (since we can calculate the distribution of  $R$  as discussed in A.2) obtain an  $1 - \kappa$  upper bound on the unknown  $m_0$ :

$$\hat{m}_0(\kappa) := m - R(SU_t, \tilde{p}) \times \min_{\widetilde{m}_0 \in [m]} \max \{ \gamma \in [0, 1] : f(\gamma, \widetilde{m}_0) \geq 1 - \kappa \}$$

Figure 2.8 depicts the sample size needed to achieve a power of at least 0.8 for an assumed minimal effect size of  $\mu \geq 2.81$  and an FDR level of  $\alpha = 0.05$ .

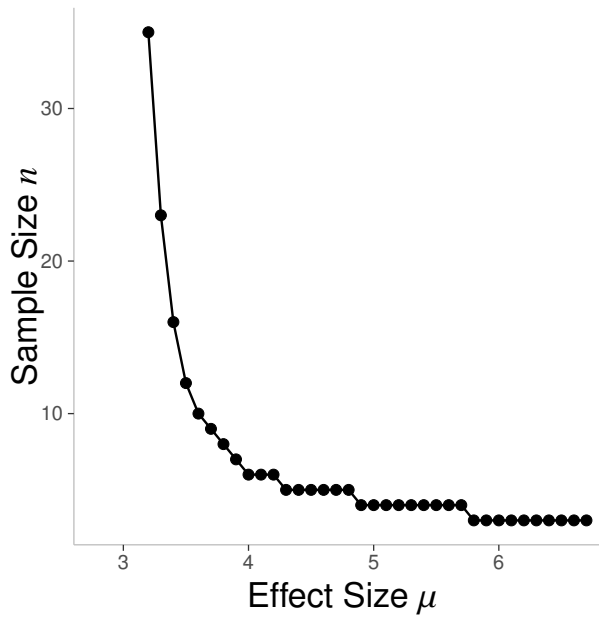


Figure 2.8: The (with probability  $1 - \kappa = 0.95$ ) needed sample size to achieve an average power of at least 0.8 when utilizing the Benjamini-Hochberg step-up procedure to control the FDR at level  $\alpha = 0.05$  for  $m = 126$  as a function of the effect size  $\mu$ . The sample sizes have been calculated from the  $p$ -value realization  $\tilde{p}$  provided by the [Not+01] data set and under the assumptions that the  $p$ -values are stochastically independent, the number of true null hypotheses is fixed but unknown and that the effect size  $\mu$  is at least 2.81 (the lower quartile of the observed test statistics). These assumptions were made for the sake of this computational example even though gene expressions are, in general, not independent.

## 2.7. DISCUSSION

WE have presented computationally efficient and numerically stable methods for calculating the joint distribution of order statistics. Such joint distributions have a multitude of important applications that require their repeated evaluation (to numerically solve optimization problems). Apart from the applications that we have presented in Section 2.6, they include, among others, the calibration goodness-of-fit tests with equal local levels (see Section 1.4 of [GLF16]) and the adjustment of the asymptotically optimal rejection curve as proposed by [FGD12], see Equation (19) in their paper, and [FDR09, Equation (6.1)] with the goal of obtaining valid critical values for a step-up-down procedure (guaranteeing strict FDR control). The latter applications have not been considered explicitly in the present work, because they merely refer to the one-group case. For this case, the methods of [MNS16] are already sufficiently accurate and fast. For most other applications Noe's recursion using standard (double precision) floating point arithmetic is likely most suitable. It is available in the preliminary version of our package (see below) as `OrdStat::pordstat2` when passing the argument `quick=TRUE`.

Future extensions to our methods could include a normalization (in Noe's recursion) to avoid underflows and the exploration of potential efficiency gains in the exact computation of Bolshev's recursion by a trade-off between the memory consumption and the frequency of normalizations of the intermediate rational numbers.

A preliminary version of our planned package (which utilizes RCPP, cf. [Edd13]) for the R language ([R C20]) is available at <https://github.com/jvschroeder/OrdStat/> and can be installed using the [devtools package](#):

```
install.packages("devtools")
devtools::install_github("jvschroeder/OrdStat")
```

The code used to generate the graphics and numerical examples is available at <https://github.com/jvschroeder/OrdStatExamples/>. The graphics were created using `ggplot2` (cf. [Wic16]) and `tikzDevice` (cf. [SB18]).

## ACKNOWLEDGMENTS

Jonathan von Schroeder was supported by the Deutsche Forschungsgemeinschaft (DFG) within the framework of RTG 2224 " $\pi^3$ : Parameter Identification - Analysis, Algorithms, Applications".

The authors wish to thank two anonymous reviewers the Associate Editor of Computational Statistics & Data Analysis, and the responsible Co-Editor for their useful comments.

# 3

## REVERSE STRESS TESTING IN SKEW-ELLIPTICAL MODELS

**Jonathan VON SCHROEDER, Taras BODNAR and Thorsten DICKHAUS**

*Show me the incentive and I will show you the outcome.*

Charlie Munger

*Stylized facts about financial data comprise skewed (log-)returns. Therefore, we extend previous results on reverse stress testing under elliptical models to the broader class of skew-elliptical models. In particular, under the assumption of a linear Profit and Loss function, we are concerned with finding the most likely scenarios given that the loss exceeds a given threshold. In the elliptical case, an explicit formula for the solution is provided. In the skew-elliptical case, we characterize the solution in terms of an easy-to-implement numerical optimization problem. As specific examples, we investigate the classes of skew-normal and skew-t models in detail. Since the solutions depend on population parameters, which are often unknown in practice, we also tackle the statistical task of estimating these parameters and provide confidence regions for the most likely scenarios. Finally, we present an application to real currency exchange data.*

---

A previous version of this chapter was published as a research report; cf. [https://kurser.math.su.se/pluginfile.php/19448/mod\\_folder/content/0/2019/2019\\_4\\_report.pdf?forcedownload=1](https://kurser.math.su.se/pluginfile.php/19448/mod_folder/content/0/2019/2019_4_report.pdf?forcedownload=1)

## DECLARATION OF INDIVIDUAL CONTRIBUTIONS

The introduction (sections 3.1 and 3.2) and the conclusion (section 3.5) are joint work with Prof. Dr. Thorsten Dickhaus and Prof. Taras Bodnar.

Section 3.3 is joint work with Prof. Taras Bodnar, to which we contributed equally. The contributions to the corresponding proofs in appendix B.2 are as follows: We contributed equally to the proofs in sections B.2.1 and B.2.2 contributed equally. The proofs in sections B.2.3 and B.2.4 are entirely my own work.

Prof. Dr. Thorsten Dickhaus contributed to section 3.4.1 by suggesting the method and to section 3.4.3 by suggesting the methods to test for skewness. All other parts of section 3.4 are entirely my own work. The corresponding proofs in section B.2.5 and section B.3 are entirely my own work.

All code needed to produce the graphics/tables is entirely my own work.

Furthermore, I am very thankful to Prof. Dr. Thorsten Dickhaus for supporting me in terms of presentation and writing in all parts of the manuscript.

## 3.1. INTRODUCTION

RECENTLY the Basel Committee on Banking Supervision has recognized the importance of including extreme scenarios as part of an overall stress testing programme:

A stress testing programme should also determine what scenarios could challenge the viability of the bank (for example by reverse stress testing) and thereby uncover hidden risks and interactions among risks. [Ban09, p. 18 ff.]

Thus one of the methods relevant to stress testing is the scenario analysis. The utilized ‘[s]cenarios usually involve some kind of coherent, logical narrative or ‘story’ as to why certain events and circumstances can occur and in which combination and order’ [Gov12, p. 9]. Stress scenarios should reflect an ‘organization’s unique vulnerabilities to factors that affect its exposures, activities, and risks’ [Gov12, p. 9]. This however means that an analyst has to select plausible scenarios based on knowledge about the organization which will invariably introduce a bias. Furthermore plausibility is a constraint that, based on historical data, can exclude certain ‘break the bank’ scenarios that can occur during extreme events like the financial crisis of 2008. One of the main goals of reverse stress testing is to overcome this limitation by ‘assum[ing] a known adverse outcome ... and then deduc[ing] the types of events that could lead to such an outcome’ [Gov12, p. 9]. To that end the question of what happens to a given collection of assets (the portfolio under investigation), if a market factor changes in a certain way, is reversed by asking instead what could cause a certain portfolio event (e.g. a loss exceeding a certain threshold, cf. [Kop+15]). Reverse stress testing is usually based on the density of the Profit and Loss (P&L) distribution, which has to be estimated for real-world portfolios. Addressing the arising confounding issue requires ‘that the distribution that serves as the foundation for reverse stress tests is consistent with stylized facts of actual tail behavior’ [Kop+15]. These stylized facts, as described by [Con01], include heavy tails which are also present in high frequency data (cf. [BDE03]) and which imply that normality can certainly not be assumed. Furthermore many empirical loss distributions are highly skewed and thus even the family of elliptical distributions is often not rich enough; cf. [MFE05, p. 44].

Non-normality of (log-)returns can also be deduced from standard financial models. For example, one may consider the popular stochastic volatility model by [Hes93]. This model may be regarded as a generalization of the Black-Scholes model, where the latter implies log-normal prices. As, for instance, shown by [Zha+16], this generalization leads to skewness. The stationary density of the

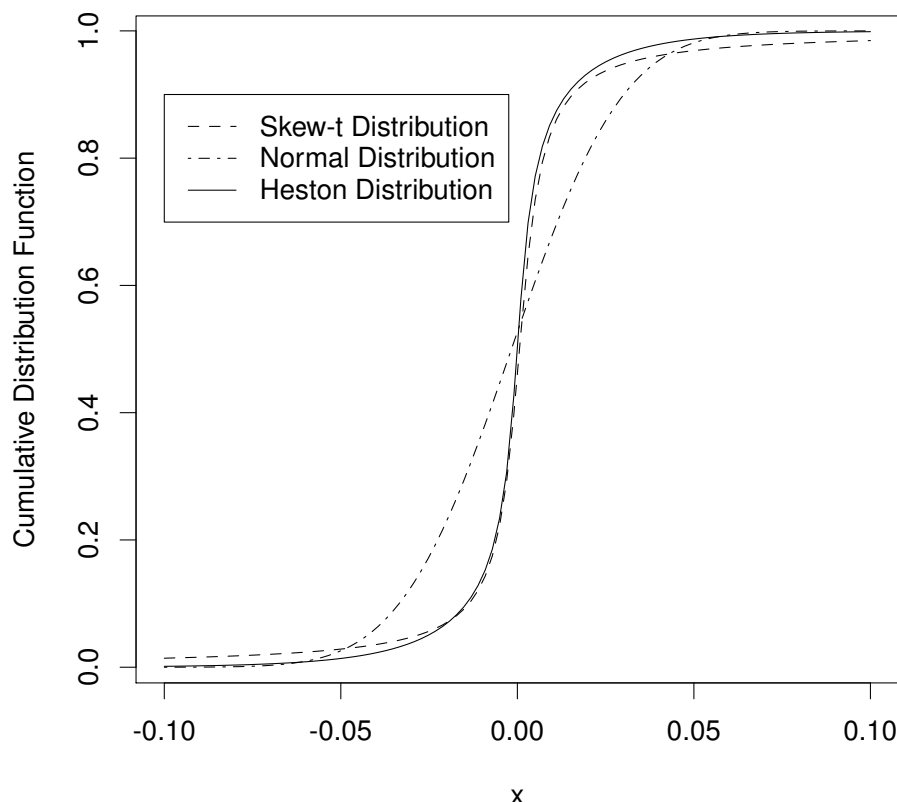


Figure 3.1: Comparison between a chosen stationary Heston (solid line) and approximations to it. The approximations were obtained by drawing a sample of size 200 and performing a maximum-likelihood fit for normal (dot-dash line) and skew-t distribution (dashed line).

returns under the Heston model has been derived by [DY02]. Its complicated mathematical form makes it intractable for statistical inference purposes. However, this distribution can be closely approximated by a skew-t distribution; see Figure 3.1.

In conclusion, the normality assumption of [Kop+15] as well as the more general assumption of ellipticity of [GKK15] are not always appropriate in the reverse stress testing context. Furthermore the approach of [Tra+19] to utilize the Mahalanobis distance to judge the likelihood of a scenario is only appropriate, when the distribution is rotationally symmetric, i.e. does not present skewness. In this work, we therefore explore some aspects of the representation of the solution given in Proposition 1 of [GKK15] and give conditions for global optimality. We then extend our results to the more general class of skew-elliptical distributions and skew-t distributions. This choice is motivated by its practical relevance and its mathematical tractability. Heavy tails are often observed in financial data. Although the skew-normal distributions can capture some heavy-tailed behaviour, more sophisticated models are needed, especially to properly model financial markets during turbulent periods. Skew-elliptical models are a natural extension of the skew-normal distribution which appear to be very useful for this goal. This family of distributions includes also mean-variance mixtures of normal distributions as special cases (see, e.g., [Azz13], [BMP19]). Both skew-normal and skew-elliptical distributions have already been used in financial literature. For instance, [CGT03], [BG15], [SNT17] use the skew-normal distribution as a model of the asset returns, while the applications of skew-t distribution in portfolio theory, risk management, and actuarial sciences are discussed by [Adc10], [Adc14], and [AEL15]. Furthermore several results for the general class of skew-elliptical distributions from the perspective of financial applications are discussed by [Shu17] and [AA20].

The rest of the material is structured as follows. In Section 3.2, we introduce basic notation and assumptions. Section 3.3 contains our main results. The specific cases of skew-normal and skew-t models are investigated in Section 3.3.2. Section 3.4 contains results regarding confidence regions for the most likely scenarios. We conclude with a discussion in Section 3.5. Mathematical details leading to our main results and proofs are deferred to the appendix.

## 3

## 3.2. PROBLEM SETTING

### 3.2.1. REVERSE STRESS TESTING

To formalize the notion of identifying the most likely scenario or scenarios ‘that result in losses exceeding a given magnitude for a particular portfolio or firm’ (cf. [GKK15]) we need a formal definition of losses. To this end consider an  $\mathbb{R}^d$ -valued random variable  $X = (X_1, \dots, X_d)^T$ , where  $X_i$  is the change in value of the  $i$ -th portfolio asset over the holding period. Following [MFE05, p. 4] we define the profit and loss (P&L) function  $v$  of the portfolio under consideration by

$$v(X) = c^T X = \sum_{i=1}^d c_i X_i \quad (3.1)$$

for a fixed (non-stochastic) vector  $c \in \mathbb{R}^d$  of portfolio weights. This definition includes that of [GKK15] by an obvious choice of the vector  $c$ . Specifically singling out the case of a linear portfolio has the advantage that it is not necessary to assume that  $(X, v(X))$  is jointly in the same distribution class as  $X$  which is for example not the case for the class of skew-normal distributions which is not closed under convolution (cf. [Azz13, p. 27]). There are classes of skew-elliptical distributions which are closed under convolution, notably the closed skew-normal distribution introduced by [GDG04] and the closed-skew t distribution introduced by [Ive10]. We are, however, not aware of any general development of skew-elliptical distributions closed under convolution. In the present work, we only have to model the joint distribution of the portfolio factors due to the linearity assumption in (3.1). If a nonlinear relationship between portfolio factors and loss has to be assumed, our proposed methods can still be applied if we assume that the joint distribution of portfolio factors and loss is in the class of (skew-)elliptical distributions. [GKK15] assumed the latter for the elliptical case.

The assumption of linear portfolios is in line with modern portfolio theory as suggested by [Mar52], where the portfolio return is equal to a linear combination of asset returns (see also [Bod09], [Bra10], [Elt+14], and [Bod+19]). Moreover, for analytical and / or numerical tractability a restriction to the linear or quadratic case is often necessary; see e.g. [Stu97].

For a given loss threshold  $\ell$  the goal of a reverse stress test is to, following [GKK15], find the most likely loss scenarios  $x^*(\ell)$  given that the loss exceeds  $\ell$ . If  $f$  is the density of  $X$  and large positive values of  $v(X)$  represent large losses this can be mathematically formalized as<sup>1</sup>

$$x^*(\ell) \in \arg \max_{x \in \mathbb{R}^d} f(x | c^T X \geq \ell),$$

i. e.,  $x^*(\ell)$  maximizes the likelihood conditional on having observed a loss of at least  $\ell$ . This is equivalent to (cf. Appendix B.1)

$$\begin{aligned} & \underset{x}{\text{maximize}} && f(x) \\ & \text{subject to} && c^T x \geq \ell. \end{aligned} \quad (3.2)$$

<sup>1</sup>For details regarding the conditional density  $f(\cdot | c^T X \geq \ell)$  see Appendix B.1.



The first result about (3.2) when  $f$  is an elliptical density was presented by [GKK15]. It states that there is a connection to the conditional expectation (which can be estimated using the empirical likelihood method):

**Theorem 1** ([GKK15, Proposition 1]). *Assuming appropriate tail behaviour of the distribution of  $X$  there exists a sequence  $\{\kappa_\ell\}_\ell$  (which depends on the tail behaviour) such that  $\kappa_\ell \rightarrow \kappa \in \mathbb{R}$  as  $\ell \rightarrow \infty$  and*

$$x^*(\ell) = \kappa_\ell \cdot \mathbb{E}[X | c^T X \geq \ell]. \quad (3.3)$$

3

### 3.2.2. THE FAMILY OF (SKEW-)ELLIPTICAL DISTRIBUTIONS

This section aims to give a short introduction to the family of skew-elliptical distributions as well as relevant special properties. The exhibition closely follows [Azz13, Chapter 6].

**Definition 2** (Elliptical Distribution). *Let  $\mu \in \mathbb{R}^d$  be a given vector and  $\Sigma \in \mathbb{R}^{d \times d}$  a positive definite symmetric matrix. Any density  $f$  on  $\mathbb{R}^d$  fulfilling*

$$f(x; \mu, \Sigma) \propto g((x - \mu)^T \Sigma^{-1} (x - \mu))$$

*for some function  $g : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ , is called the density of an elliptical distribution on  $\mathbb{R}^d$ . The function  $g$  is then called the density generator of  $f$ ,  $\mu$  is called the location parameter, and  $\Sigma$  is called the dispersion matrix.*

Simultaneous statistical inference theory under elliptical models has been developed by [BD17]. It is interesting to remark that  $\Sigma$  is (in general) not the covariance matrix. However, we have the following corollary to [GVB13, Theorem 2.11 (b)]:

**Corollary 1.** *If  $X$  is elliptically distributed according to the preceding definition and possesses second moments, then the covariance matrix of  $X$  is a real multiple of  $\Sigma$ , i.e. there exists a real constant  $\kappa > 0$  such that  $\text{Cov}(X) = \kappa \Sigma$ .*

An approach for estimating the covariance matrix in elliptical models when the number of samples is small is discussed in [Zho+14].

**Remark 4.** *The generators for common elliptical probability distributions are (see e.g. [LV03, p. 62 ff.]) as follows.*

- *Normal distribution:*  $g(u) := \exp\left(-\frac{u}{2}\right)$
- *Student-t distribution:*  $g(u) := \left(1 + \frac{u}{c_p}\right)^{-p}$  for  $c_p \in \mathbb{R}$ ,  $p > \frac{1}{2}$
- *Logistic Distribution:*  $g(u) := \frac{e^{-u}}{(1+e^{-u})^2}$
- *Exponential distribution:*  $g(u) := \exp(-ru^s)$  for  $r, s > 0$

Since it is of great importance for the results of this paper we note that the following holds:

**Lemma 8.** *For the normal, Student-t, logistic, and exponential distributions the generator is strictly decreasing on  $\mathbb{R}_+$ .*

The family of elliptical distributions can be extended to allow for skewness by multiplying elliptical densities with a scaling function; cf. [Azz13, Equation (6.11)].

**Definition 3** (Skew-elliptical distribution). Let  $\tilde{f}$  be the density of an elliptical distribution with density generator  $\tilde{g}$  (cf. Definition 2) and let  $F$  be the cumulative distribution function of a univariate elliptical distribution. Then

$$f(x; \mu, \Sigma, \lambda) := 2 \cdot \tilde{f}(x; \mu, \Sigma) \cdot F(\lambda^T(x - \mu)) \quad (3.4)$$

is the density of a skew-elliptical distribution. The parameter  $\mu$  is called the location parameter,  $\Sigma$  is called the dispersion matrix and  $\lambda$  is called the skewness parameter.

3

### 3.3. THEORY OF REVERSE STRESS TESTING IN (SKEW-)ELLIPTICAL MODELS

#### 3.3.1. REVERSE STRESS TESTING IN ELLIPTICAL MODELS

SINCE the conditional expectation in (3.3) depends on the tail behaviour (and so does  $\kappa_\ell$ ), the solution given by (3.3) could depend on the tail behaviour of the distribution of  $X$ . Furthermore only observations for which the loss exceeds  $\ell$  can be utilized when estimating the conditional expectation using the empirical likelihood method. In this section we derive an explicit equation for  $x^*(\ell)$  (in terms of the population parameters) which proves that  $x^*(\ell)$  does not depend on the tail behaviour of the distribution of  $X$ . In particular, we are therefore able to present empirical likelihood estimates and confidence regions that utilize all observations (instead of just those corresponding to losses larger than  $\ell$ , cf. [GKK15, Section 3]) in Section 3.4.1.

The following theorem asserts that the reverse stress testing problem has, under elliptical models, a global optimum and gives an explicit expression in terms of the parameters of the density:

**Theorem 2.** If  $f$  is an elliptical density with a decreasing generator  $g$  (cf. Definition 2), then the unique solution of (3.2) is given by

$$x^*(\ell) := \begin{cases} \mu & \text{if } \ell \leq c^T \mu, \\ \mu + (\ell - c^T \mu) \frac{\Sigma c}{c^T \Sigma c} & \text{otherwise.} \end{cases} \quad (3.5)$$

By Lemma 8 the generators of the normal, Student-t, logistic, and exponential distribution are strictly decreasing. In practice the generator, and therefore its properties, might be unknown. In this situation we suggest to empirically check that the distribution is unimodal (see e.g. [BP09]), since, by a result of [BB14, Proposition 2], the generator of an elliptical distribution is decreasing if and only if the elliptical distribution is unimodal.

Equation (3.5) shows that, for fixed  $\ell$ , the solution does not depend on (the tail behaviour of the generator)  $g$ , but only on the parameters  $\mu$  and  $\Sigma$  of the (elliptical) distribution of  $X$ . As a result, the derived expression of the most likely loss scenario is applicable to the whole family of elliptical distributions without the need to specify the density generator  $g$  explicitly. The location parameter  $\mu$  and the dispersion matrix  $\Sigma$  are unknown quantities in practically relevant applications which have to be estimated by using historical data (see Section 4 for further discussion). In order to simplify the estimation procedure, the following remark is insightful.

**Remark 5.** The solution

$$x^*(\ell, \Sigma) := \mu + \left( \frac{\ell - c^T \mu}{c^T \Sigma c} \right) \Sigma c \quad (3.6)$$

does not depend on the scaling of  $\Sigma$ , i.e. for  $\alpha \in \mathbb{R}_{>0}$  it holds that  $x^*(\ell, \alpha \Sigma) = x^*(\ell, \Sigma)$ . Therefore, assuming existence of second moments of  $X$  and by Corollary 1, the dispersion matrix  $\Sigma$  can be replaced by the covariance matrix of  $X$ . The latter matrix can be estimated from data by standard

techniques, e.g. the method of moments. The location parameter  $\mu$  is usually estimated by the arithmetic mean of the data. Once estimators  $\hat{\mu}$  and  $\hat{\Sigma}$  are available, we propose for practical purposes to evaluate the right-hand side of (3.5) with  $\mu$  replaced by  $\hat{\mu}$  and  $\Sigma$  replaced by  $\hat{\Sigma}$ .

### 3.3.2. REVERSE STRESS TESTING IN SKEW-ELLIPTICAL MODELS

To derive the solution of (3.2) in the case of a skew-elliptical model we, in the following, assume that  $\lambda \neq \mathbf{0}$ , since otherwise the results of Theorem 2 can be applied. Furthermore let  $\lambda = (\lambda_1, \lambda_2^T)^T$  with  $\lambda_1 \neq 0$  which can be assumed without loss of generality, since otherwise the components of the vector  $x - \mu$  can be rearranged.

The following result, which is a direct consequence of Corollary 4, characterizes the most likely scenario in skew-elliptical models.

**Theorem 3.** *Let  $f$  be the density of a skew-elliptical distribution with strictly decreasing generator  $\tilde{g}$  (cf. Definition 3). Then, there exist real numbers  $k_1$  and  $k_2$  such that*

$$x^*(\ell) = \mu + k_1 \cdot \Sigma \lambda + k_2 \cdot \Sigma c \quad (3.7)$$

is a solution of (3.2).

This result settles the question of the existence of a solution, but does not guarantee uniqueness. If  $f$  is a skew-normal density, i.e.  $\tilde{f}$  is a normal density and  $F$  is the cumulative distribution function of the univariate normal distribution in Definition 3, then we obtain the following stronger result by making use of Corollary 7.

**Theorem 4.** *Under the conditions of Theorem 3 and if  $f$  is a skew-normal density, the solution of (3.2) is unique. The numbers  $k_1, k_2$  in (3.7) can in this case be found by solving a one-dimensional optimization problem (cf. (B.9)).*

It is possible to define a skew-t distribution by choosing  $\tilde{f}$  in Definition 3 as a multivariate Student's t-density. We, however, derive a result similar to that of Theorem 4 for the more popular formulation of a skew-t distribution given in (B.10). Namely, Theorem 5 is a simplified restatement of Theorem 10.

**Theorem 5.** *If  $f$  is a skew-t density (in the sense of (B.10)), the solution of (3.2) is unique. Furthermore the solution is of the form (3.7) and the numbers  $k_1$  and  $k_2$  can be found by solving a one-dimensional optimization problem (cf. (B.13)).*

### 3.3.3. NUMERICAL ILLUSTRATION

It is now possible to illustrate how the selected scenario in the skew-normal model and skew-t model differs from the one obtained when misspecifying the skewness as zero. As an example consider the points

$$\lambda_\varphi = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$$

for  $\varphi \in (0, 2\pi) \setminus \{\pi/2\}$  which lie on the unit circle in  $\mathbb{R}^2$ . Denote by  $x_\varphi$  the scenario for  $\lambda = \lambda_\varphi$  and  $c = \lambda_{\pi/2}$ . Furthermore denote by  $y_\varphi$  the scenario obtained by setting  $\lambda = \mathbf{0}$  (i.e. when  $\lambda$  is misspecified as zero, that is the scenario obtained by using Equation (3.6) and the population mean and variance according to Lemma 18). Thus we consider here only the effect of model-misspecification, and we quantify this effect by plotting  $\varphi$  against the  $L_1$ -norm of the difference

vector  $x_\varphi - y_\varphi$ . The confidence regions resulting from estimation uncertainty are discussed in Section 3.4.

Figure 3.2 considers four different cases for the covariance structure:

(1) **Negative correlation and equal variances**  $\Sigma = \begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix}$

(2) **Positive correlation and equal variances**  $\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$

(3) **Uncorrelated and equal variances**  $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

(4) **Uncorrelated and unequal variances**  $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$

As can be seen in Figure 3.2 the covariance structure does not significantly alter the qualitative behaviour, meaning that in both columns all four graphs look qualitatively similar: If  $\lambda_\varphi$  points approximately in the direction of  $c$ , then the difference is comparatively small. The further observed local minima depend on the inner product induced by  $\Sigma$  as well as the other parameters of the distribution. They can therefore not be characterized simply in terms of  $\lambda_\varphi$  and  $c$ . The location of comparatively large differences depends on  $\Sigma$ , since the latter can cause a rotation of the distribution. From the practical point of view, the results indicate that skewness only matters when its direction differs markedly from the direction of the portfolio vector. We have observed similar phenomena (not shown here) also in higher dimensions  $d \geq 3$ .

The impact of the skewness is relatively small under the skew-normal model: Up to 5-8% of the scenario are allocated incorrectly when the most likely scenario is calculated according to the elliptical model. Under the skew-t model the misallocation can reach up to 20-30%. In both cases the misallocation can therefore be of the same order of magnitude as the prescribed loss level  $\ell = 1$ .

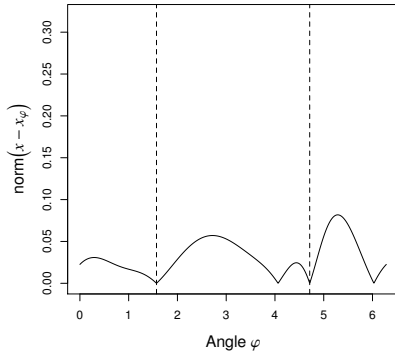
Additionally, Figure 3.3 compares the run times and the accuracy of our proposed methods with those of the all-purpose constrained optimization routine `constrOptim` in R (cf. [R C20]). The figure demonstrates that the advantages (runtime and accuracy) of our proposed numerical method increase with growing dimension  $d$ .

### 3.4. CONFIDENCE REGIONS FOR THE MOST LIKELY SCENARIO

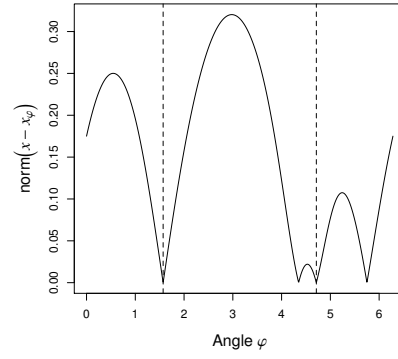
Point estimation of the most likely stress scenario is, due to the results of the previous section, straightforward once the population parameters have been estimated. For the elliptical case these can be estimated using e.g. the method of moments and for the skew-normal and skew-t models maximum likelihood estimation (cf. [Azz19]) can be used. This section builds on this to describe approaches to obtain confidence regions for the most likely stress scenario.

#### 3.4.1. ELLIPTICAL MODELS

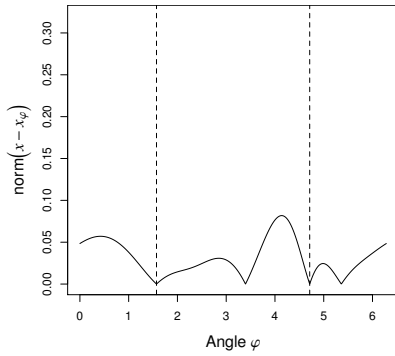
THE goal of this section is to apply the empirical likelihood (EL) method (cf. [Owe01]), which provides likelihood-based inference methods without assuming a parametric model for the data, to construct a confidence region for the scenario obtained under the setting of Theorem 2. It is provided by the following corollary to Corollary 8:



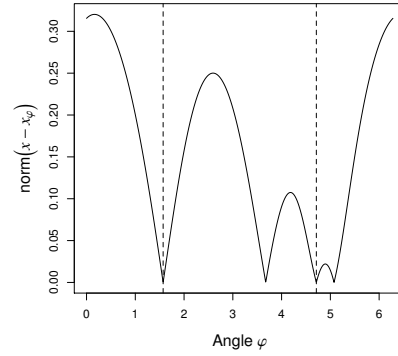
(a) Negative correlation and equal variances (skew normal)



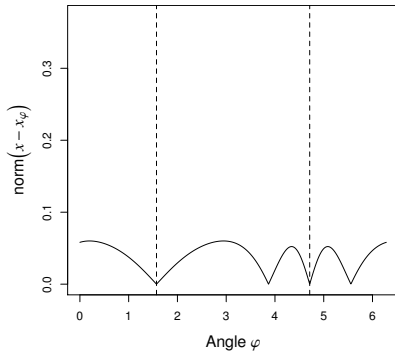
(b) Negative correlation and equal variances (skew-t)



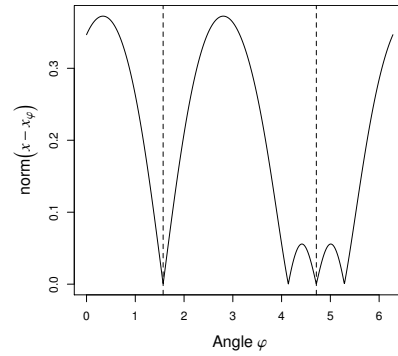
(c) Positive correlation and equal variances (skew normal)



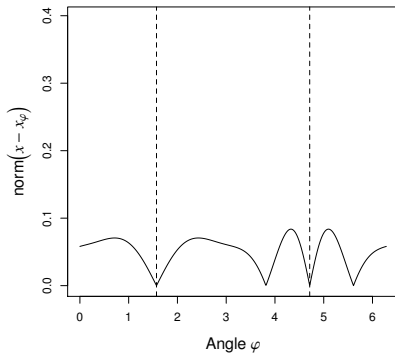
(d) Positive correlation and equal variances (skew-t)



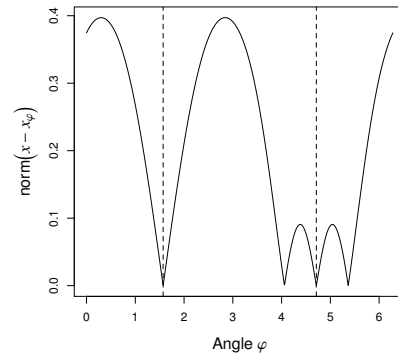
(e) Equal variances and uncorrelated (skew normal)



(f) Equal variances and uncorrelated (skew-t)



(g) Unequal variances and uncorrelated (skew normal)



(h) Unequal variances and uncorrelated (skew-t)

Figure 3.2: Relative distance between the selected scenario in the skewed and non-skewed model in  $\mathbb{R}^2$  as a function of  $\varphi$ , where  $\text{norm}(v) := \|v\|_1$  denotes the  $L_1$ -norm of a vector  $v$ . The number  $\varphi \in [0, 2\pi]$  represents the angle between the skewness vector  $\lambda_\varphi$  and the vector  $c$  of the portfolio weights. The dashed vertical lines mark the 'direction' (as well as its opposite) in which  $c$  points. For further details see Section 3.3.3.

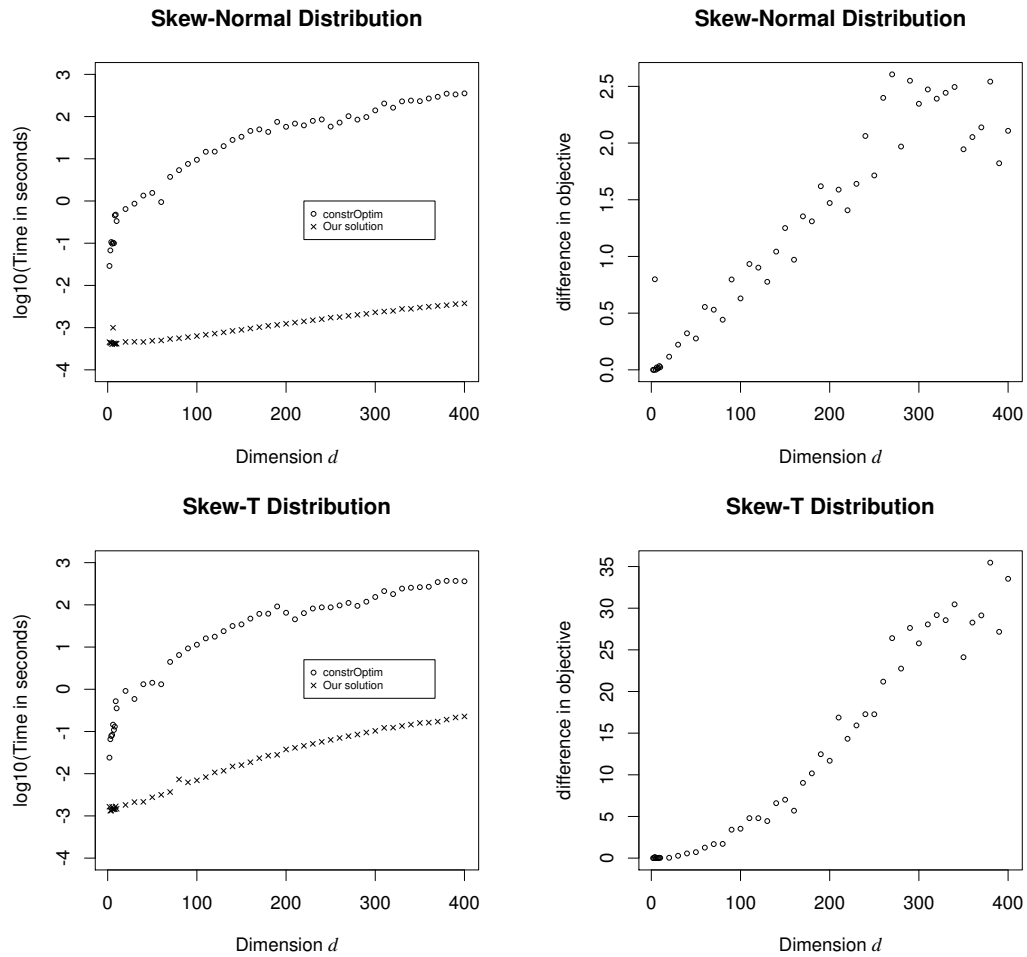


Figure 3.3: Comparison of the run time and achieved objective value when the dimensionality  $d$  of the optimization problem increases. We compared our solution to a naive implementation using R's `constrOptim` method. In the right sub-figures, we calculated the difference of the value of the objective function obtained with our method minus the corresponding value obtained by `constrOptim`. The positive values displayed in the right column are in favour of our method.

**Corollary 2.** Let  $Y_1, \dots, Y_n$  be stochastically independent and identically distributed according to an elliptical density with parameters  $\mu$  and  $\Sigma$ . Furthermore let

$$X_1 := \begin{pmatrix} Y_1 \\ (c^T Y_1) \cdot Y_1 \end{pmatrix}, \dots, X_n := \begin{pmatrix} Y_n \\ (c^T Y_n) \cdot Y_n \end{pmatrix}$$

and

$$h(m) = h(\mu, \mu^{(1)}) = \mu + \frac{\ell - c^T \mu}{c^T (\mu^{(1)} - c^T \mu \cdot \mu)} \times (\mu^{(1)} - c^T \mu \cdot \mu),$$

where  $\mu^{(1)}$  is the expectation of  $(c^T Y_1) \cdot Y_1$ . Then the Jacobian of  $h$  (which is given by (B.14)) has rank  $q = d - 1$  and an asymptotic  $(1 - \alpha)$  confidence region for  $x^*(\ell)$  is given by

$$\left\{ h(\hat{m}) + \sum_{i=1}^n w_i \left( \frac{\partial h}{\partial m}(\hat{m}) X_i \right) \middle| \prod_{i=1}^n n w_i \geq \exp\left(-\frac{c_\alpha}{2}\right), w_i \geq 0, \sum_{i=1}^n w_i = 1 \right\},$$

where  $\hat{m} = n^{-1} \sum_{i=1}^n X_i$  and  $c_\alpha$  is the  $(1 - \alpha)$  quantile of the  $\chi^2$ -distribution with  $q$  degrees of freedom.

In comparison to the confidence region proposed in [GKK15, Theorem 1] the confidence region from Corollary 2 does not depend on a nuisance parameter describing the tail behaviour and can utilize all observations (instead of just those for which the loss exceeds  $\ell$ ). It is, however, based on the EL confidence region for twice as many parameters. Furthermore the only available marginal confidence intervals from Corollary 2 are projections onto the coordinate axes, which have asymptotic joint coverage probability of  $1 - \alpha$ . In contrast to this it is possible to directly derive marginal  $1 - \alpha$  confidence regions from [GKK15, Theorem 1] that are (in general) distinct from the ones obtained by projections onto the coordinate axes. Both approaches are applied to an example in Section 3.4.3.

### 3.4.2. SKEW-NORMAL AND SKEW-T MODELS

FOR the skew-normal and skew-t models the estimating equations for the maximum likelihood estimator are quite complex and the solution to the reverse stress testing problem is only available implicitly. Therefore the empirical likelihood approach is intractable. Fortunately the (parametric) bootstrap approach (cf. [Hor19, Section 2.2]) is a suitable alternative: Maximum likelihood estimation for the (multivariate) skew-normal and the skew-t distribution is readily available in the R language (cf. [Azz19]). This can be utilized to estimate the relevant parameters from a data set. The point estimate of the most likely scenario for which the loss exceeds a (fixed) threshold can then be obtained by solving the corresponding one-dimensional optimization problem. The parametric bootstrap approach can be utilized to obtain pseudo samples of  $x^*(\ell)$  from which marginal confidence regions can be obtained using e.g. the method described by [HY13]. This approach is demonstrated in Section 3.4.3.

The discussed approach, which involves the estimation of all population parameters, might not be desirable (or even not feasible) for high-dimensional data. We therefore dedicate the remainder of this section to considering the method of moment estimation of the population parameters, relevant for the task at hand, in skew-normal and skew-t models. As detailed in Appendix B.3 the method of moments estimators of all relevant parameters are not guaranteed to exist. We, however, have the following useful result (which follows directly from the results in Appendix B.3):

**Theorem 6.** Let  $X$  be an  $\mathbb{R}^d$ -valued random variable that is distributed according to an elliptical, a skew-normal or a skew-t distribution with location parameter  $\xi$ , dispersion matrix  $\Omega$  and skewness parameter  $\alpha$ . Then it holds that

$$x^*(\ell) \in \{\xi + a_1 \cdot \Omega c + a_2 \cdot \Omega \alpha \mid a_1, a_2 \in \mathbb{R}\}. \quad (3.8)$$

If  $X$  is skew-normal or skew- $t$  distributed but  $x^*(\ell)$  is incorrectly calculated according to (3.6) (with  $\mu = \mathbb{E}[X]$  and  $\Sigma = \text{Cov}(X)$ , see Lemma 18), then (3.8) still holds. Furthermore if  $\hat{\xi}$  is a consistent estimator of the location parameter (whose components can be estimated marginally), then (B.15) and (B.16) yield consistent plugin estimators of the affine two-dimensional subspace in (3.8).

If a confidence region is obtained using the assumption of an elliptical distribution and the method described in Section 3.4.1, then this result can be utilized to identify a one-dimensional subset of the confidence region (from which multiple scenarios can be easily selected) by its intersection with

$$\{x \in \mathbb{R}^d \mid c^T x = \ell\} \cap \left\{ \hat{\xi} + a_1 \cdot \widehat{\Omega}c + a_2 \cdot \widehat{\Omega}\alpha \mid a_1, a_2 \in \mathbb{R} \right\}.$$

In practice this seems to work well as demonstrated in the application in the following section.

### 3.4.3. APPLICATION TO A CURRENCY PORTFOLIO

In this section we consider the application of the developed theory to "[...] a basket of currencies, half held in British pounds (GBP), the rest divided evenly among the Australian dollar (AUD), the euro (EUR), the Japanese yen (JPY) and the Swiss franc (CHF)" (cf. [GKK15, Section 5.2]). The portfolio is valued in US dollars and the monthly returns are analyzed. These are defined as  $R_i := (V_i - V_{i-1})/V_{i-1}$ , where  $V_i$  is the average value (in US dollars) of one unit of the currency under consideration in month  $i$ ,  $1 \leq i \leq n$ .

Since currencies are traded around the world there is no canonical data set. For our analysis we choose the monthly data that can be freely obtained from the Federal Reserve System<sup>2</sup> for the period from January 2000 to December 2019 which, in the period from January 2000 to December 2011, is very similar to the data set analyzed by [GKK15].

To apply the method of [GKK15] we performed a maximum likelihood fit of the data to a multivariate  $t$ -distribution (using `QRM: fit.mst`, cf. [PM20]) which yielded  $\hat{\nu} = 4.765659$  as an estimate for the degrees of freedom. Choosing  $\ell = 2.8776\%$  (which is the upper 5% quantile of the observed losses) yields the estimate (according to [GKK15, Proposition 1])

$$(AUD, EUR, JPY, CHF, GBP) \quad (3.9)$$

$$= (-3.7197\%, -3.2303\%, 0.2639\%, -2.9468\%, -4.0723\%) \quad (3.10)$$

of the most likely scenario exceeding this loss. The plug-in estimate of (3.5) (using the empirical mean and covariance) is, on the other hand, given by

$$(AUD, EUR, JPY, CHF, GBP) \quad (3.11)$$

$$= (-3.0412\%, -3.0973\%, -1.3745\%, -2.7961\%, -3.1780\%) \quad (3.12)$$

which differs from the former estimate in two ways:

- The most likely loss in Japanese yen is negative and not positive. The marginal 99% confidence regions are  $[-1.5478\%, 2.7895\%]$  and  $[-2.5564\%, 0.8401\%]$ , respectively. Hence, we do not have confidence in the sign of the component of  $x^*$  referring to JPY.
- The most likely losses are smaller. This can be easily explained by noticing that the loss of the scenario (3.10) is approximately 3.2403% which is larger than  $\ell$  whereas the scenario

<sup>2</sup><https://www.federalreserve.gov/datadownload/Download.aspx?rel=H10&series=8bbd346e2878c9b2137ff7alc9e95be9&lastobs=&from=01/01/2000&to=12/31/2019&filetype=csv&label=include&layout=seriescolumn>



(3.12) has, by design, a loss exactly equal to  $\ell$ . The latter estimate is based on the assumption that  $\ell$  is bigger than the expected loss, which seems reasonable since  $\ell$  has been chosen as a quantile far from the median.

Figures 3.4 and 3.5 each contain two of the ten bivariate marginals and the confidence regions obtained using both methods mentioned before. We rasterized the boundaries of the confidence regions using the method described in [CH13, Section 4] using 200 points.

The confidence regions obtained by [GKK15] "for the most likely loss reflect the skewness in the joint distribution of the EUR/USD and CHF/USD returns". It is therefore natural to consider modelling this skewness using a skew-t distribution. Before proceeding we test the null hypothesis  $H_0 : \lambda = 0$  versus the alternative  $H_1 : \lambda \neq 0$ . We test this (global) hypothesis in terms of the individual hypotheses

$$H_0^I : \lambda_I = 0 \text{ versus } H_1^I : \lambda_I \neq 0 \quad (3.13)$$

for every subset  $I$  of  $\{1, \dots, 5\}$ . In (3.13),  $\lambda_I$  denotes the skewness parameter of the corresponding marginal distribution. We test each  $H_0^I$  using the corresponding likelihood-ratio statistic which, under  $H_0^I$ , asymptotically follows a chi-squared distribution with  $|I|$  degrees of freedom (cf. [Wil38]). After applying the Bonferroni correction (cf. [Dic14, p. 30]) we obtain the minimal adjusted p-value 0.002741076 which provides clear evidence for the presence of skewness. Furthermore, Table 3.1 provides (unadjusted) bootstrap-based 95% confidence intervals for the components of the skewness parameter vector. Only one of them covers zero. Hence, we work with the full model including the skewness parameter.

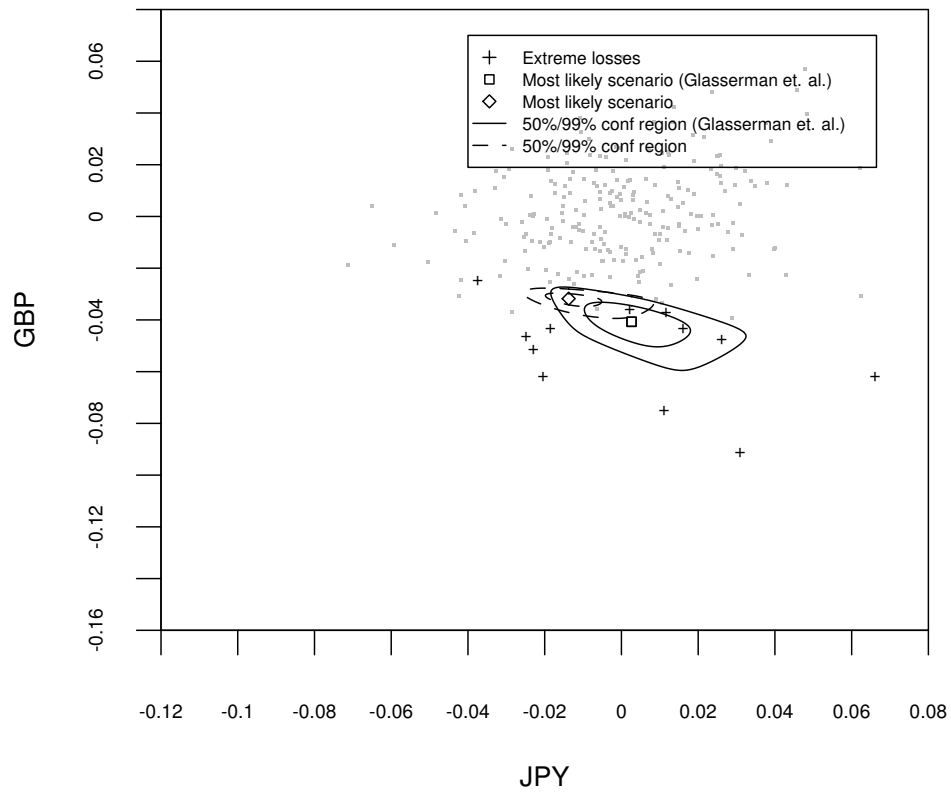
A maximum likelihood fit using the R function `sn::mst.mple` (i.e. without penalty, cf. [Azz19]) yields a slightly larger value of  $\hat{\nu} = 4.940044$  for the degrees of freedom. For selected bivariate marginals, Figures 3.6 and 3.7 depict confidence regions for the most likely scenario originating from elliptical and skew-t models, respectively. The latter were estimated from  $10^4$  (parametric) bootstrap samples. It is interesting to note that all  $10^4$  realizations of  $x^*(\ell)$  obtained using the parametric bootstrap had only strictly negative components which is more in line with (3.12) than with (3.10).

Finally, we also utilized the method described at the end of Section 3.4.2 which is a practical way to select interesting scenarios while avoiding the estimation of all parameters of the skew-elliptical distribution. From the resulting one-dimensional subset of the confidence region, we selected 21 scenarios for illustrative purposes. Figure 3.8 depicts the  $L_1$ -distances of these 21 scenarios to the most likely scenario estimated under the full skew-t model. Among these 21 scenarios, there is the scenario estimated under the elliptical model (indicated by a cross in Figure 3.8). The figure demonstrates that some of the scenarios selected with the method proposed at the end of Section 3.4.2 yield an improvement.

Table 3.1: Unadjusted bootstrap-based 95% confidence intervals, based on 1000 bootstrap samples, for the components of the skewness parameter vector.

Parameter	95% CI
$\lambda_1$	(-1.9064, -0.0010)
$\lambda_2$	(-1.7539, -0.0027)
$\lambda_3$	(-0.0177, 1.2450)
$\lambda_4$	(0.0022, 1.8977)
$\lambda_5$	(0.0026, 3.2029)

## Return based on monthly value in Dollar



## Return based on monthly value in Dollar

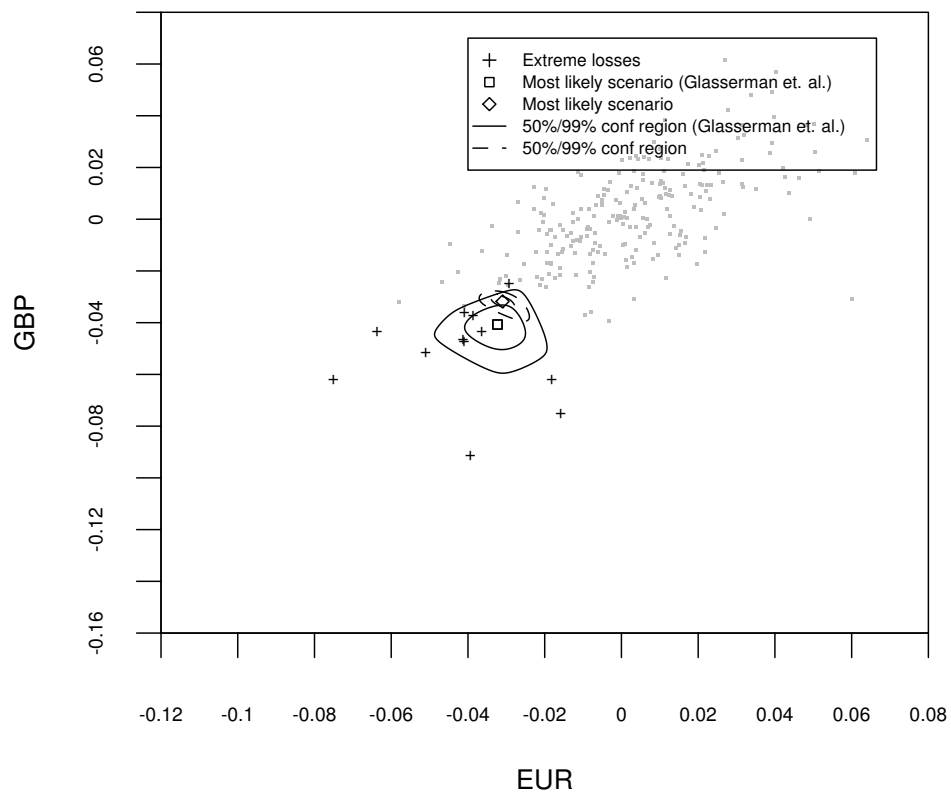


Figure 3.4: Comparison of our proposed methodology with the one by [GKK15] (bivariate marginals JPY / GBP as well as EUR / GBP): Illustration of the point estimates of the most likely scenario (the square and the diamond), the confidence regions (the solid and the dashed lines), the losses (small grey points) and the 5% most extreme losses (crosses).

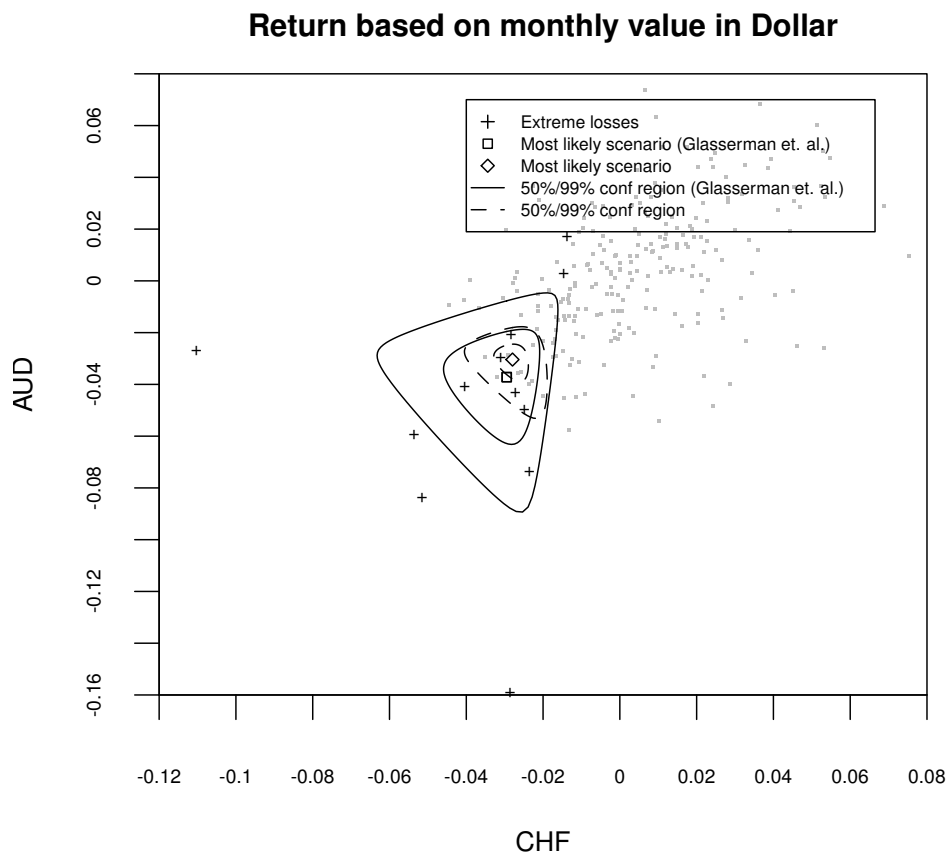
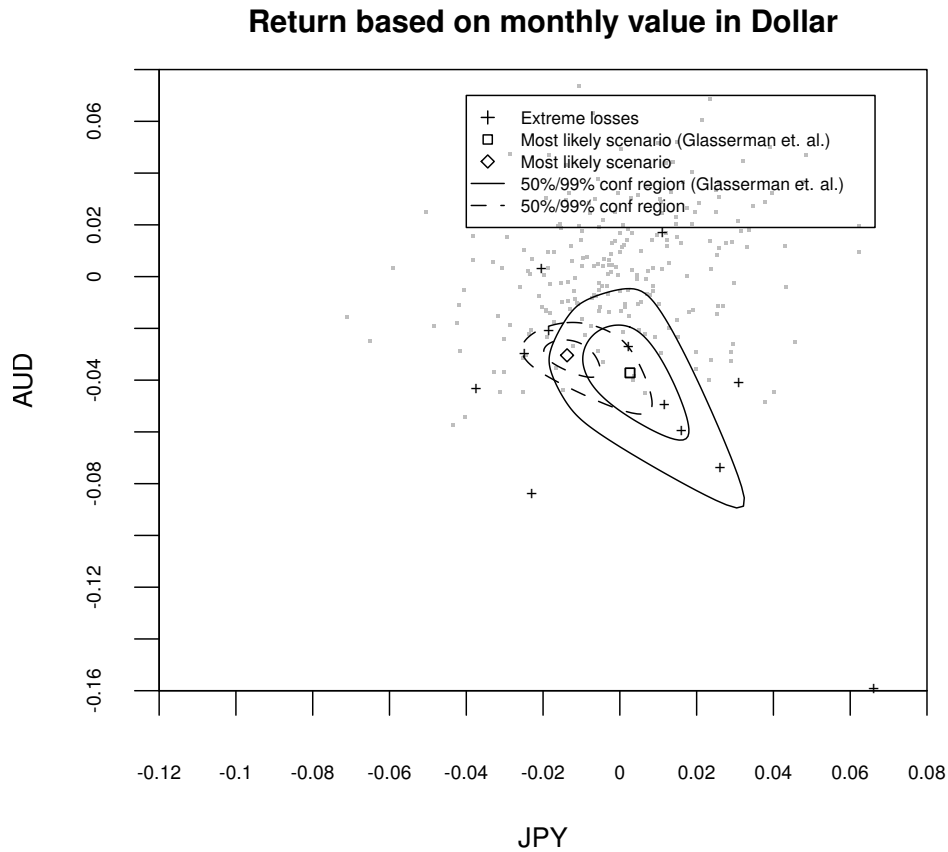


Figure 3.5: Comparison of our proposed methodology with the one by [GKK15] (bivariate marginals JPY / AUD as well as CHF / AUD): Illustration of the point estimates of the most likely scenario (the square and the diamond), the confidence regions (the solid and the dashed lines), the losses (small grey points) and the 5% most extreme losses (crosses).

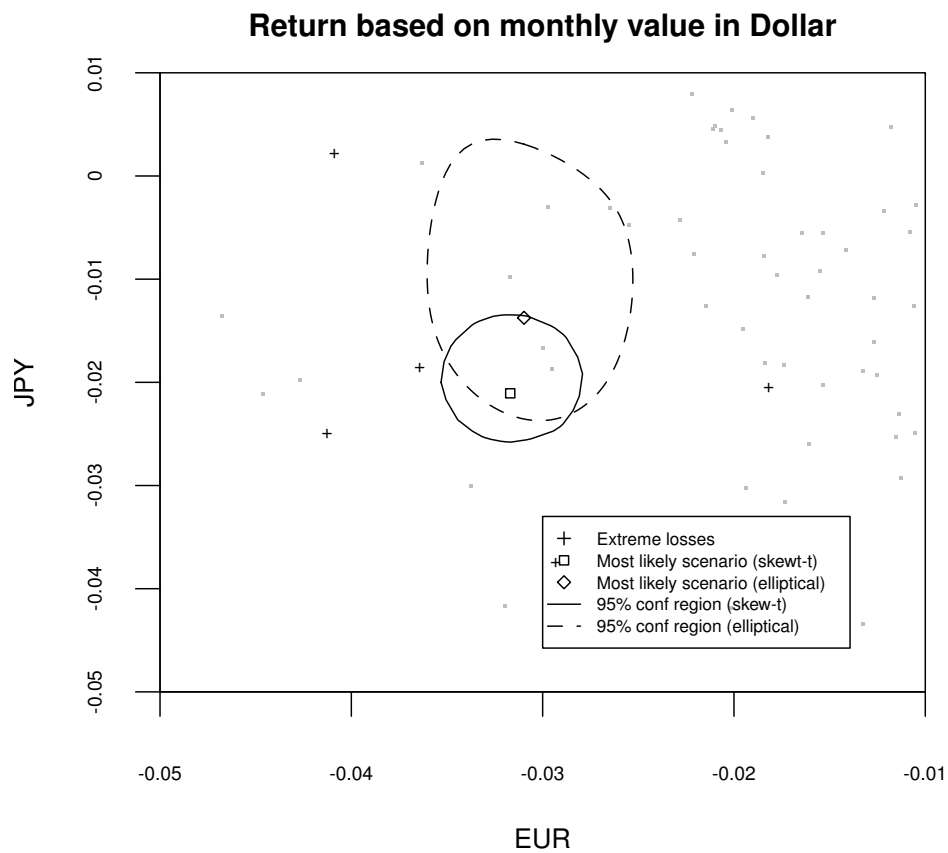
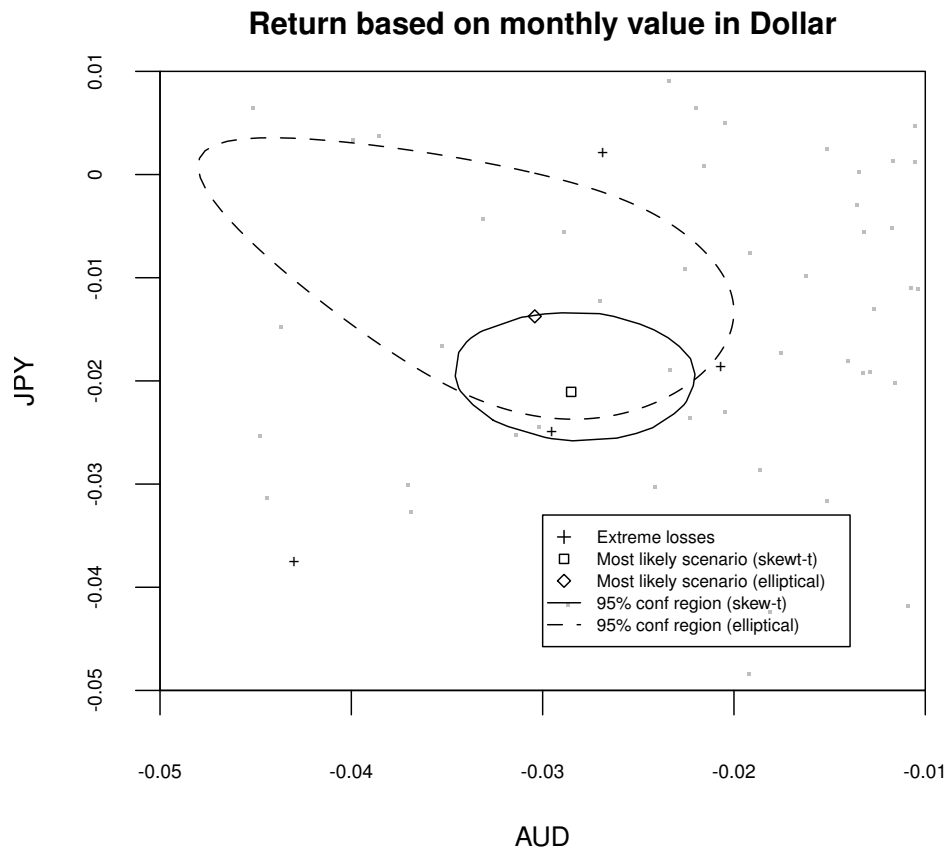


Figure 3.6: Comparison between models with and without skewness (bivariate marginals AUD / JPY as well as EUR / JPY): Illustration of the point estimates of the most likely scenario (the square and the diamond), the confidence regions (the solid and the dashed lines) as well the losses (small grey points) and the 5% most extreme losses (crosses).

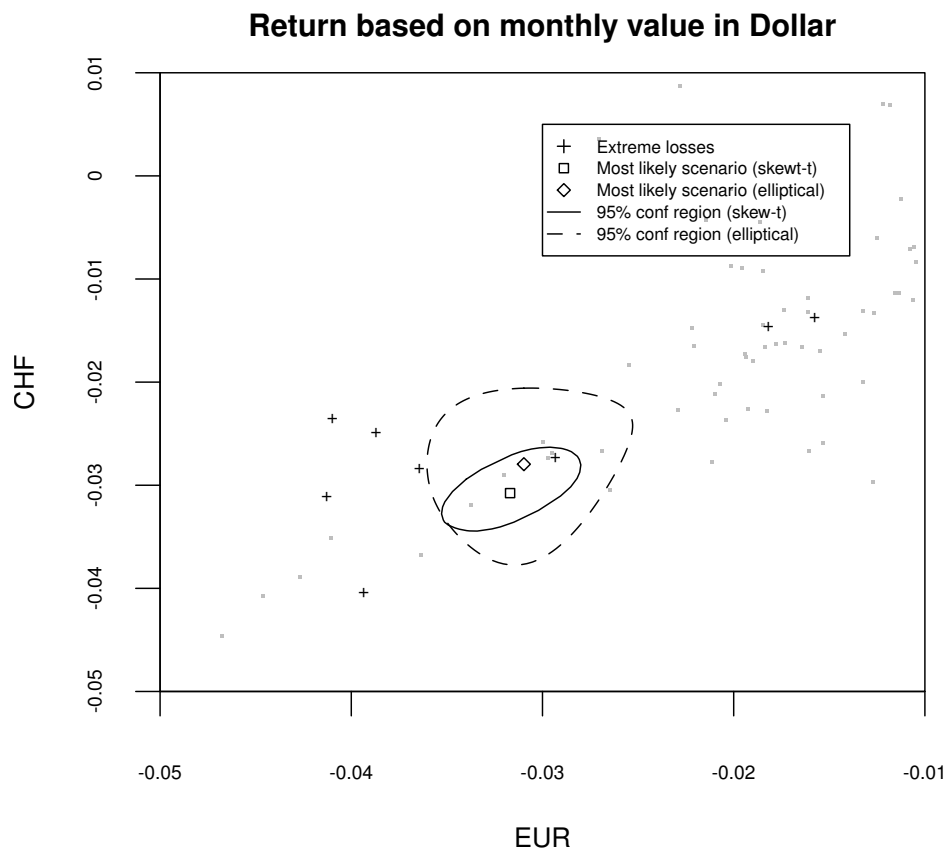
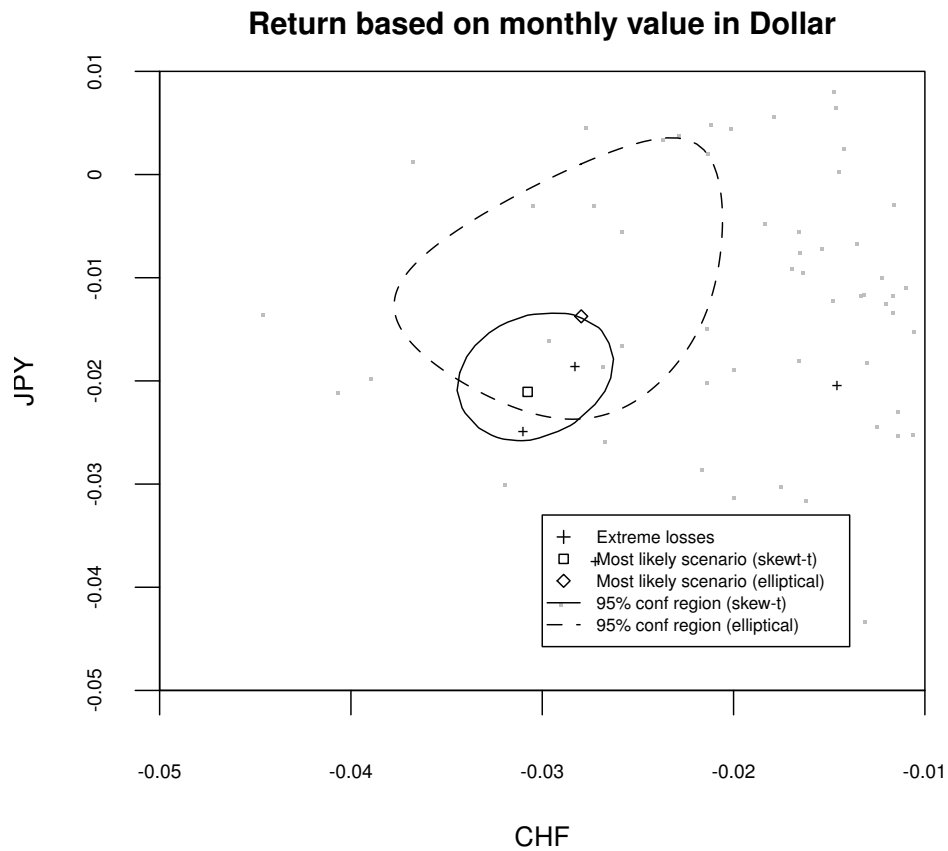


Figure 3.7: Comparison between models with and without skewness (bivariate marginals CHF / JPY as well as EUR / CHF): Illustration of the point estimates of the most likely scenario (the square and the diamond), the confidence regions (the solid and the dashed lines) as well the losses (small grey points) and the 5% most extreme losses (crosses).

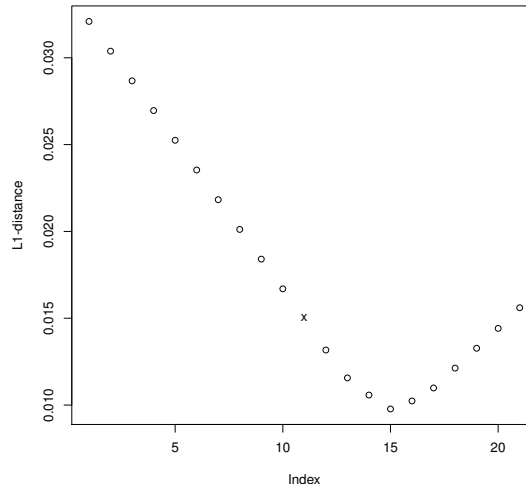


Figure 3.8: Illustration of the  $L_1$ -distance between the scenario estimated from a skew-t model and the scenarios selected from the subspace estimated using the method of moments. The scenario estimated under the elliptical model is indicated by a cross.

### 3.5. CONCLUSION

**R**EVERSE stress testing is a highly relevant task in the context of bank regulation. Therefore, it is essential that reliable and numerically stable methods are available. With the present work, we have contributed (i) an explicit solution for the most likely scenario  $x^*(\ell)$  given that the loss exceeds  $\ell$  under the scope of elliptical models, and (ii) a characterization of  $x^*(\ell)$  in terms of a numerically stable and easy-to-implement optimization problem under the broader scope of skew-elliptical and skew-t models by showing that the solution is an element of a two-dimensional affine subspace. We have furthermore tackled the statistical task of estimating the relevant population parameters and quantifying the uncertainty which propagates from the latter estimation to the obtained value of  $x^*(\ell)$  by providing methods to construct confidence regions and select scenarios from these confidence regions.

Potential extensions of our work would be to consider more general (non-linear) P & L functions as well as different distributional models, for instance by utilizing copula theory. It might also be interesting to explore, whether an approach similar to ours could be utilized to extend the work of [FK15] on the selection of multiple, well-distributed, stress scenarios in elliptical models to skew-elliptical models.

Computer programs, with which all results of the present paper can be reproduced, are available from the first author upon request.

### ACKNOWLEDGMENTS

We thank two anonymous referees for their constructive comments. We are grateful to Peter Ruckdeschel for fruitful discussions. Jonathan von Schroeder was supported by the Deutsche Forschungsgemeinschaft (DFG) within the framework of GRK 2224 /  $\pi^3$ : Parameter Identification - Analysis, Algorithms, Applications.

# 4

## STATISTICAL METHODS WITH APPLICATIONS TO MALDI MSI DATA

負けるが勝ち

Japanese Proverb

*Perfection is the enemy of progress.*

Winston Churchill

*This chapter<sup>1</sup> aims to give some insights into the statistical properties of matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI MSI) data. Specifically I propose a novel (multiplicative) trend normalisation for MALDI MSI data and demonstrate its utility by applying it successfully to a real data set. Its effect on the stability of feature selection is investigated using an approach inspired by the work of [MB10], by comparing the change in stability of different correlation coefficients. For the setting relevant to this chapter (binary classification) the finite sample distribution of the correlation coefficient recently proposed by [Cha20] is derived and the computational and theoretical properties of other correlation coefficients are discussed in detail. Finally a marginal model for the (high-dimensional) MALDI MSI observations is proposed.*

---

<sup>1</sup>Parts of this chapter were published as a preprint; cf. [Sch20].

## 4.1. INTRODUCTION

Matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI MSI) is a method suitable for high throughput imaging that shows potential for tissue typing, especially tumor subtyping (cf. [Bos+17]). It is an "an established tool for the investigation of formalin-fixed paraffin-embedded (FFPE) tissue samples", but "the applicability of this method [...] is often hampered by inevitable technical variation and limited reproducibility."; cf. [Bos+21]. Thus it is of great interest to gain insights into different types of (technical) variation and how to either avoid or correct them. Towards this end many different methods have been proposed. These include:

- Total ion count (TIC) normalization, which rescales spectra such that their L1 norm (sum of intensities) equals 1. Alternatives to this include the root mean square normalisation and the z-score normalisation; cf. e.g. [Wol+05].
- Intensity profile normalization (IPN), which "transform[s] spectral intensities such that their statistical distributions become more similar"; cf. [Bos+21].
- Variance stabilization, which is based on the idea that, for certain classes of distributions, it is possible to find a deterministic transformation such that the transformed random variable has approximately constant variance; cf. [Yu09]. An example of this is the square-root transformation for the Poisson distribution (cf. [Bar36]). The square-root and the logarithm transformation have been successfully applied to MALDI data (cf. [Wol+05], [Dei+11]).

When reanalysing (the annotated regions of) the data set presented in [Kri+16] I noticed a clear difference in the trend of the variance between the 8 different tissue micro arrays (TMAs) that were imaged to obtain the data set; see Figure 4.1. Notably the trend differences do not differ much by class, which indicates that they are not due to biological variability. To the best of my knowledge this type of variation has not yet been documented in literature. Furthermore it does not look like IPN or the variance stabilization can efficiently eliminate the observed variation. Therefore I propose to simply estimate and normalise this observed trend.

This proposed normalisation is very simple: Borrowing the idea from time series analysis<sup>2</sup>, that the trend of a time series can be estimated by smoothing the time series I propose the following simple implementation (given in R and assuming that the spectra are the rows of the matrix data):

```
1 normalize_ <- function(data,run) {
2   x <- apply(data,2,mean)
3   adj <- 1/runmed(x,run)
4   t(apply(data,1,function(x){
5     v <- x*adj
6   })))
7 }
8 normalize <- function(data,groups,run=301) {
9   for(g in unique(groups)) {
10    idx <- groups == g
11    data[idx,] <- normalize_(data[idx,],run)
12  }
13  data
14 }
```

The argument `groups` should contain categorical values indicating either the TMA or (which I have used in the following) the individual from whose tissue sample the spectrum was obtained.

<sup>2</sup>This is sensible because a spectrum obtained using time of flight mass spectroscopy is a time series since each  $m/z$  value corresponds to an arrival time of ions of a specific mass to charge ratio.



In my opinion the latter is preferable since it works just as well and does not assume that many tissue cores are measured at the same time as part of a TMA. This proposed approach seems to have favourable effects on classification performance and the stability of feature selection, as demonstrated in the following sections.

## 4.2. FEATURE SELECTION USING UNIVARIATE TESTS FOR HOMOGENEITY

Feature selection is an important task in statistics and machine learning, as it allows for dimensionality reduction and selection of features that merit further (potentially very expensive/laborious) investigation. This section describes the connection between univariate tests for homogeneity and feature selection for binary classification. Specifically it considers the AUC (whose stability when utilized for feature selection for the analysis of MALDI MSI data was recently discussed in [Wil+19]) and a recently by [Cha20] proposed correlation coefficient. Since the Chatterjee coefficient was found to have lower power than alternatives (cf. [SDH20]) I compare it with a resampled version (which can be seen to be an incomplete U-statistic using the coefficient as kernel function). Furthermore Section 4.2.4 gives some insights into the computational and theoretical properties of Hoeffding's D (cf. [Hoe48]), Blum-Kiefer-Rosenblatt's R (cf. [BKR61]), Bergsma-Dassio-Yanagimoto's  $\tau^*$  (cf. [BD14]) as well as U-statistics based on Chatterjee's  $\xi$ .

Denote by  $[n] := \{1, \dots, n\}$  the first  $n \in \mathbb{N}$  natural numbers, let  $[n]_0 := [n] \cup \{0\}$  and let, in the following,  $(X_1, Y_1), \dots, (X_n, Y_n)$  be a sample of bivariate, jointly independent random variables where (for all  $i \in [n]$ )  $X_i \in \{0, 1\}$  and  $Y_i$  is real valued with absolutely continuous distribution function. One way to test the suitability of  $Y_i$  for predicting  $X_i$  is to test the homogeneity hypothesis (c.f. [Dic18, p. 51])

$$H_0 : \mathbb{P}_{Y_1} = \mathbb{P}_{Y_2} = \dots = \mathbb{P}_{Y_n}. \quad (4.1)$$

To this end let  $N_0 := |\{i \in [n] : X_i = 0\}|$ ,  $N_1 := n - N_0 = |\{i \in [n] : X_i = 1\}|$  and let  $\pi = \pi(Y_1, \dots, Y_n) : [n] \rightarrow [n]$  be the (almost surely unique) random permutation such that  $Y_{\pi(1)} \leq \dots \leq Y_{\pi(n)}$ . It holds that  $(X_{\pi(1)}, \dots, X_{\pi(n)})|N_0 = n_0$  is, under  $H_0$ , distributed uniformly on

$$B_{n_0, n_1} = \left\{ x \in \{0, 1\}^n \mid \sum_{i \in [n]} x_i = n_1 \right\} \quad (4.2)$$

which is the set of all binary sequences of length  $n$  with exactly  $n_0$  zeros and  $n_1 := n - n_0$  ones, since, under  $H_0$ ,  $\pi$  sampled uniformly at random from the set of all permutations of  $[n]$ . Thus this will be the setting for which the following sections will discuss the distribution of the Mann-Whitney U statistic and that of Chatterjee's  $\xi_n$ .

### 4.2.1. THE MANN-WHITNEY U TEST

The Mann-Whitney U test is a nonparametric test for the homogeneity hypothesis (4.1). It is based on the test statistic

$$U_{n_0, n_1} = U_{n_0, n_1}((X_1, Y_1), \dots, (X_n, Y_n)) := -\frac{n_0(n_0 + 1)}{2} + \sum_{i \in I_0} \pi^{-1}(i) \quad (4.3)$$

where  $I_0 := \{i \in [n] : X_i = 0\}$  and  $\pi^{-1}$  denotes the inverse permutation of  $\pi$  (i.e.  $\pi^{-1}(i)$  is the rank of  $Y_i$ ).

Exact and approximate  $p$ -values under  $H_0$  can be e.g. obtained in R using the `wilcox.test` command. The expectation and variance of  $U_{n_0, n_1}$  are given by  $\mathbb{E}[U_{n_0, n_1}] = \frac{n_0 n_1}{2}$  and  $\text{Var}(U_{n_0, n_1}) = \frac{n_0 n_1 (n+1)}{12}$  (c.f. [LD75, p. 14]). Furthermore  $U_{n_0, n_1}$  takes values in  $[0, n_0 n_1]$  (c.f. [LD75, p. 9]).

In many applications  $U_{n_0, n_1}$  is replaced by  $\text{AUC}_{n_0, n_1} := \frac{U_{n_0, n_1}}{n_0 n_1}$  and then called the area under the ROC curve. It takes values in  $[0, 1]$  and has variance  $\frac{n_0 n_1 (n+1)}{12 n_0^2 n_1^2}$ .

#### 4.2.2. CHATTERJEE'S RANK CORRELATION $\xi$

Recently a new coefficient of correlation was proposed in [Cha20]. This section discusses its properties and gives the finite sample distribution of this coefficient for the setting relevant to this paper. The coefficient is defined as

$$\xi_n(Y, X) := 1 - \frac{n \sum_{i=1}^{n-1} |r_{i+1} - r_i|}{2 \sum_{i=1}^n l_i \times (n - l_i)} \quad (4.4)$$

where  $r_i := \left| \left\{ j \in [n] : X_{\pi(j)} \leq X_{\pi(i)} \right\} \right| = \pi^{-1}(i)$ ,  $l_i := \left| \left\{ j \in [n] : X_{\pi(j)} \geq X_{\pi(i)} \right\} \right|$ . It is obviously symmetric, simple to compute and has the property that  $\xi(Y, X) := \lim_{n \rightarrow \infty} \xi_n(Y, X) = 0$  iff  $Y \perp X$  and  $\xi(Y, X) = 1$  iff there exists a measurable  $f$  such that  $X = f(Y)$  almost surely (assuming  $X$  is not a.s. constant, cf. [Cha20, Theorem 1.1.]).

Since  $X$  is a binary random variable it is immediate that

$$\begin{aligned} r_i &= N_0 + \chi(X_{\pi(i)} = 1) \times N_1 \\ l_i &= \chi(X_{\pi(i)} = 0) \times N_0 + N_1 \end{aligned}$$

and furthermore it follows that

$$\sum_{i=1}^N l_i \times (n - l_i) = \sum_{i=1}^{N_1} N_0 N_1 = N_0 N_1^2$$

and

$$\begin{aligned} \sum_{i=1}^{n-1} |r_{i+1} - r_i| &= \sum_{i=1}^{n-1} N_1 \chi(r_{i+1} \neq r_i) \\ &= N_1 \times \left| \left\{ i \in [n-1] : X_{\pi(i+1)} \neq X_{\pi(i)} \right\} \right|. \end{aligned}$$

Thus it follows that (4.4) can be written as

$$\begin{aligned} \xi_n(Y, X) &:= 1 - \frac{n N_1 \times \left| \left\{ i \in [n-1] : X_{\pi(i+1)} \neq X_{\pi(i)} \right\} \right|}{2 N_0 N_1^2} \\ &= 1 - \frac{n}{2 N_0 N_1} \times \left| \left\{ i \in [n-1] : X_{\pi(i+1)} \neq X_{\pi(i)} \right\} \right|. \end{aligned}$$

Since this is only well-defined if  $N_0, N_1 > 0$  one needs to choose a value for the (uninformative) cases  $N_0 = 0$  or  $N_1 = 0$ . Since in this case there is neither evidence against nor for independence of  $X$  and  $Y$  it is reasonable to choose  $\xi_n(Y, X) = 0$  if  $N_0 = 0 \vee N_1 = 0$ . This choice does not affect the properties of the limit  $\xi$  unless  $X$  is almost surely constant.

To derive the finite sample distribution of  $\xi_n$  when  $X$  and  $Y$  are independent and conditional on  $N_0 = n_0, N_1 = n_1$ , denote by

$$\tau_n(x) = |\{i \in [n-1] : x_i \neq x_{i+1}\}|$$

the number of 'jumps' in the binary sequence  $x \in B_{n_0, n_1}$ . If  $X$  is sampled uniformly at random from  $B_{n_0, n_1}$ , then the distribution of  $\tau_n(X)$  is given by the following corollary to the results of [Tha01]:

**Corollary 3.** If  $X$  is sampled uniformly at random from  $B_{n_0, n_1}$ , then the distribution of  $\tau(X)$  is given by

$$\mathbb{P}(\tau(X) = x) = \begin{cases} (x+1)^2 \times G_{n_0, n_1}(x+1) & \text{if } x \text{ is odd} \\ (nx - x^2) \times G_{n_0, n_1}(x) & \text{if } x \text{ is even} \end{cases} \quad (4.5)$$

where  $n := n_0 + n_1$ ,  $x \in [2 \times (n_0 \wedge n_1) - \chi(n_0 = n_1)]$  and where

$$G_{n_0, n_1}(x) := \left[ 2n_0 n_1 \times \binom{n}{n_0} \right]^{-1} \times \binom{n_0}{\frac{x}{2}} \binom{n_1}{\frac{x}{2}}.$$

The [proof](#) is deferred to Appendix C.1.

The following lemma summarizes some properties of this distribution:

**Lemma 9.** The expectation of the probability distribution defined by (4.5) is given by

$$\mathbb{E}[\tau(X)] = \frac{2n_0 n_1}{n}$$

For  $n_0 = n_1 = m$  it holds that  $\mathbb{P}(\tau(X) = m + a) = \mathbb{P}(\tau(X) = m - a)$  for all  $a \in \mathbb{N}$ . Furthermore, if  $n_0 = n_1 = m$ , the variance of the probability distribution defined by (4.5) is given by

$$\text{Var}(\tau(X)) = m \frac{2m^2 - 1}{2m - 1} - m^2 = \frac{m(m-1)}{2m-1}.$$

The [proof](#) is deferred to Appendix C.1.

For  $n_0 \in [n-1]$  it holds that if  $X$  and  $Y$  are independent  $1 - \frac{n}{2n_0(n-n_0)} \tau(Z_{n_0}) \equiv \xi_n(Y, X) | N_0 = n_0$  since  $\pi$  is then just a permutation sampled uniformly at random from the set of all permutations of  $[n]$  and where  $Z_{n_0}$  is sampled uniformly at random from  $B_{n_0, n-n_0}$ . It follows immediately that  $\mathbb{E}[\xi_n(Y, X) | N_0 = n_0] = 0$ .

#### 4.2.3. BOOTSTRAPING CHATTERJEE'S RANK CORRELATION $\xi$

A popular technique to reduce the variance of an estimator of a parameter is the (subsampling) bootstrap. It is simple to demonstrate that the ordinary bootstrap (i.e. resampling with replacement instead of subsampling) is inappropriate for  $\xi_n$ : If  $(X_1^*, Y_1^*), \dots, (X_n^*, Y_n^*)$  is a bootstrap sample (sampled with replacement), then

$$\begin{aligned} \mathbb{E}_{\mathbb{P}^*} \left[ \left| \{i \in [n-1] : X_{\pi(i+1)}^* \neq X_{\pi(i)}^*\} \right| \right] &= n-1 - \mathbb{E} \left[ \left| \{i \in [n-1] : X_{\pi(i+1)}^* = X_{\pi(i)}^*\} \right| \right] \\ &\leq (n-1) \times \left( 1 - \left( 1 - n^{-1} \right)^n \right) \end{aligned}$$

since the expected fraction of non-unique points in the bootstrap sample is  $\left( 1 - n^{-1} \right)^n$ . Consequently

$$\mathbb{E}_{\mathbb{P}^*} [\xi_n(X^*, Y^*)] \geq 1 - \frac{n}{2n_1 n_2} \times (n-1) \times \left( 1 - \left( 1 - n^{-1} \right)^n \right)$$

and thus for  $O(n^{-1} n_1) = O(n^{-1} n_2) = O(1)$  it follows that  $\lim_{n \rightarrow \infty} \mathbb{E}_{\mathbb{P}^*} [\xi_n(X^*, Y^*)] \geq e^{-1}$ . But it follows from [Cha20, Theorem 1.1] that  $\lim_{n \rightarrow \infty} \mathbb{E}[\xi_n(X, Y)] = 0$  if  $X$  and  $Y$  are independent and therefore the bootstrap fails to be asymptotically unbiased in this case. Sampling without replacement does not cause artificial ties and therefore does not suffer from this issue.

If all subsamples of a fixed size  $m$  are evaluated and the resampled statistic is symmetric, then the average over all such subsamples is an U-statistic, which is a type of unbiased and asymptotically normal test statistic that was introduced by [Hoe48]. Since both of the previously discussed test statistics are symmetric functions of the sample one can use them as the kernel of a U statistic (c.f. [BC18, Definition 1.1]) and define<sup>3</sup>

$$T_{(m,n)}^{\text{AUC}} := \binom{n_0 + n_1}{m}^{-1} \sum_{1 \leq i_1 < \dots < i_m \leq n} \text{AUC}_{\tilde{n}_0, \tilde{n}_1}((X_{i_1}, Y_{i_1}), \dots, (X_{i_m}, Y_{i_m}))$$

$$T_{(m,n)}^{\xi} := \binom{n_0 + n_1}{m}^{-1} \sum_{1 \leq i_1 < \dots < i_m \leq n} \xi_{\tilde{n}_0, \tilde{n}_1}((X_{i_1}, Y_{i_1}), \dots, (X_{i_m}, Y_{i_m}))$$

where  $\tilde{n}_1 = \tilde{n}_1(i_1, \dots, i_m) := \sum_{j=1}^m X_{i_j}$  and  $\tilde{n}_0 = \tilde{n}_0(i_1, \dots, i_m) := m - \tilde{n}_1$ . In practice  $\binom{n_0 + n_1}{m}^{-1}$  will be so large for many applications, that an exact evaluation of  $T_{(m,n)}^{\text{AUC}}$  and  $T_{(m,n)}^{\xi}$  is computationally hard for large  $m$ . It is however possible to replace these test statistics by Monte Carlo approximations. To this end denote by  $I^{(1)}, \dots, I^{(\ell)}$  a sample of size  $\ell$  drawn uniformly at random from  $\{(i_1, \dots, i_m) \in [n]^m : 1 \leq i_1 < \dots < i_m \leq n\}$ . Then

$$\tilde{T}_{(m,n)}^{\text{AUC}, \ell} := \ell^{-1} \sum_{j=1}^{\ell} \text{AUC}_{\tilde{n}_0, \tilde{n}_1}((X_{I_1}^{(j)}, Y_{I_1}^{(j)}), \dots, (X_{I_m}^{(j)}, Y_{I_m}^{(j)})) \quad (4.6)$$

$$\tilde{T}_{(m,n)}^{\xi, \ell} := \ell^{-1} \sum_{j=1}^{\ell} \xi_{\tilde{n}_0, \tilde{n}_1}((X_{I_1}^{(j)}, Y_{I_1}^{(j)}), \dots, (X_{I_m}^{(j)}, Y_{I_m}^{(j)})) \quad (4.7)$$

are randomized approximations of the previously defined quantities that converge almost surely as  $\ell \rightarrow \infty$  due to the strong law of large numbers.  $P$ -values for these test statistics can be approximated, according to the methodology described in [PS10] (which is implemented in the R-package `statmod`).

**Remark 6.** In [Sch20] I stated that the calculation of  $T_{(m,n)}^{\xi}$  was infeasible for many practical applications. Since then I have, however, discovered that it can be, once the ranks  $Y$  are known, calculated in  $O(m^3 \cdot n)$  using  $O(m^3)$  storage. An implementation in the Julia programming language (cf. [Bez+17]) can be found in Appendix C.2.

#### 4.2.4. COMPUTATION AND ASYMPTOTICS OF U-STATISTICS FOR FEATURE SELECTION

Let  $X \in \{0, 1\}^n$  and  $Y \in \mathbb{R}^n$  be observations. Denote by  $I$  the indicator function which equals 1 iff its argument is true and 0 otherwise. Furthermore let  $V_i$ , for any vector  $V$ , be the subvector given by  $V_i := (V_{i_1}, V_{i_2}, \dots, V_{i_k})$ . Then an estimator of Hoeffding's D is given by (cf. [SDH20, Eq. (6)])

$$D_n(X, Y) := \frac{1}{n(n-1) \dots (n-4)} \sum_{\substack{i=(i_1, \dots, i_5) \subset [n]^5 \\ i_1 \neq \dots \neq i_5}} \frac{1}{4} f_D(X_i) \cdot f_D(Y_i) \quad (4.8)$$

where  $f_D(X) := [I(X_1 \leq X_5) - I(X_2 \leq X_5) \leq X][I(X_3 \leq X_5) - I(X_4 \leq X_5)]$ .

Similarly Blum-Kiefer-Rosenblatt's  $R$  can be estimated by (cf. [SDH20, Eq. (7)])

$$R_n(X, Y) := \frac{1}{n(n-1) \dots (n-5)} \sum_{\substack{i=(i_1, \dots, i_6) \subset [n]^6 \\ i_1 \neq \dots \neq i_6}} \frac{1}{4} f_R(X_{(i_1, i_2, i_3, i_4, i_5)}) \cdot f_R(Y_{(i_1, i_2, i_3, i_4, i_5)}) \quad (4.9)$$

<sup>3</sup>This is, of course, not particularly useful for the AUC as it already is an U-statistic. Thus this done here merely to check that lack of improvement in the asymptotic properties does, indeed, imply that the finite sample properties do not improve either. The empirical results in the application section confirm this.

where  $f_R = f_D$ .

Lastly an estimator of Bergsma-Dassios-Yanagimoto's  $\tau^*$  is given by (cf. [SDH20, Eq. (8)])

$$\tau_n^*(X, Y) := \frac{1}{n(n-1) \cdots (n-3)} \sum_{\substack{i=(i_1, \dots, i_4) \subset [n]^4 \\ i_1 \neq \dots \neq i_4}} f_{\tau^*}(X_i) \cdot f_{\tau^*}(Y_i) \quad (4.10)$$

where

$$f_{\tau^*}(X) := I(\max\{X_1, X_3\} < \min\{X_2, X_4\}) + I(\max\{X_2, X_4\} < \min\{X_1, X_3\}) \\ - I(\max\{X_1, X_4\} < \min\{X_2, X_3\}) - I(\max\{X_2, X_3\} < \min\{X_1, X_4\}).$$

4

If the ranks of the elements of  $Y$  are unique, then the following lemma can be verified e.g. by complete enumeration of the values (for  $\tau_n^*$  see also [WDL15]):

**Lemma 10.** *Let*

$$k_D(X) := \begin{cases} 16 & \text{if } X \in \{(0, 0, 0, 1, 1), (1, 1, 0, 0, 0)\} \\ -8 & \text{if } X \in \{(1, 0, 0, 1, 0), (0, 1, 0, 1, 0), (1, 0, 0, 0, 1), (0, 1, 0, 0, 1)\} \\ 0 & \text{otherwise} \end{cases}$$

$$k_R(X) := \sum_{\substack{i=(i_1, \dots, i_5) \in [6]^5 \\ i_1 < i_2 < \dots < i_5}} k_D(x_i)$$

$$k_{\tau^*}(X) := \begin{cases} 16 & \text{if } X \in \{(0, 0, 1, 1), (1, 1, 0, 0)\} \\ -8 & \text{if } X \in \{(1, 0, 0, 1), (0, 1, 1, 0), (1, 0, 1, 0), (0, 1, 0, 1)\} \\ 0 & \text{otherwise} \end{cases}$$

*If  $X \in \{0, 1\}^n$  and the ranks  $\pi$  of the elements of  $Y$  are unique, then it holds that*

$$D_n(X, Y) = \frac{1}{n(n-1) \cdots (n-4)} \sum_{1 \leq i_1 < \dots < i_5 \leq n} \frac{1}{4} k_D(X_{(\pi(i_1), \dots, \pi(i_5))})$$

$$R_n(X, Y) = \frac{1}{n(n-1) \cdots (n-5)} \sum_{1 \leq i_1 < \dots < i_6 \leq n} \frac{1}{4} k_R(X_{(\pi(i_1), \dots, \pi(i_6))})$$

$$\tau_n^*(X, Y) = \frac{1}{n(n-1) \cdots (n-3)} \sum_{1 \leq i_1 < \dots < i_4 \leq n} k_{\tau^*}(X_{(\pi(i_1), \dots, \pi(i_4))})$$

*and furthermore  $D_n(X, Y) = R_n(X, Y)$  if  $n \geq 6$ .*

Thus, like  $\xi_n$ , these coefficients can be calculated by sorting  $X$  according to the ranks of  $Y$  and then simply counting the number of times the patterns given in the definition of the kernels occur as subsequences of the sorted  $X$ .

The U-statistics in Lemma 10 posses first order degeneracy (cf. [Lee90, p. 78]) and therefore their asymptotic distributions are not normal. They can instead be found using [Lee90, Corollary 1 on p. 83] as demonstrated for Hoeffding's  $D$  in the following lemma:

**Lemma 11.** *Let  $X_1, \dots, X_n$  be i.i.d and  $Y_1, \dots, Y_n$  be i.i.d., and assume further that the  $X_i$  and  $Y_i$  are independent.*

Then, for  $p := \mathbb{P}(X_1 = 1)$ , it holds that

$$nD_n(X, Y) \rightarrow \frac{2p(1-p)^2}{3} \sum_{v=1}^{\infty} \lambda_v (Z_v^2 - 1)$$

in distribution, where the  $Z_v$  are i.i.d. standard normal and the  $\lambda_v$  are the eigenvalues of

$$\varphi \mapsto \int_0^1 \tilde{k}(y_1, y_2) \varphi(y_2) dy_2$$

where  $\tilde{k}(y_1, y_2) = 3(y_1^2 + y_2^2) - 6\max\{y_1, y_2\} + 2$ .

## 4

The [proof](#) is deferred to [Appendix C.1](#).

The main difficulty in utilising [Lemma 11](#) is finding the eigenvalues of  $\tilde{k}$  since it is a non-smooth kernel. It is however piecewise smooth and therefore by [[KR12](#), Theorem 3.1]) the classical Nyström approximation can be used to obtain numerical approximations that converge quickly. It is very simply to implement this method in e.g. Julia to approximate the first  $n$  eigenvalues:

```

1 function k(x,y)
2     3*(x^2+y^2)-6*max(x,y)+2
3 end
4
5 function nyström_midpoint_mat(k,n)
6     i = 1:(n+1);
7     h = 1/n;
8     s = (i .- 0.5)*h;
9     m = zeros(n,n);
10    w = ones(n) * h;
11    for i=1:n
12        for j=1:n
13            m[i,j] = w[j] * k(s[i],s[j])
14        end
15    end
16    m
17 end
18
19 import LinearAlgebra
20 n = 1000;
21 m = LinearAlgebra.Symmetric(nyström_midpoint_mat(k,n));
22 eigvals = LinearAlgebra.eigvals(m);

```

These can be used with any of the approximations discussed in [[BA15](#)] or the method of [[LPB00](#)] which are all implemented in the R package [[Bod16](#)]. They are all based on the first few cumulants (cf. [[BA15](#), Eq. (4)]) which are, for  $\sum_{v=1}^{\infty} \lambda_v Z_v^2$ , given by<sup>4</sup>

$$\kappa_r = 2^{r-1}(r-1)! \sum_{v=1}^{\infty} \lambda_v^r.$$

These are cumbersome to calculate by hand, but can be easily obtained using a computer algebra system like Wolfram Mathematica since  $\sum_{v=1}^{\infty} \lambda_v^r$  is simply the trace of the  $r$ -th power of the linear operator induced by the kernel  $\tilde{k}$ ; see [Code 1](#).

In practice I would advise the reader to obtain an approximation to the finite sample distribution using the Monte Carlo method: Samples under  $H_0$  for any  $p \in (0, 1)$  are quickly obtainable (cf. [Section C.3](#)) and all of the U-statistics in [Lemma 10](#) can be quickly calculated (cf. [C.4](#)).

<sup>4</sup>The translation and scaling due to the nuisance parameter  $p$  can easily be incorporated after the approximation of the distribution.

```

K[x_, y_] := 3*(x^2 + y^2) - 6*Max[x, y] + 2
Kt[x_, y_] :=
4/5 - 3/2*(x^4 + y^4) + 3*(x^3 + y^3) - 6*(x^2 + y^2) +
9*(x + y - x*y)*x*y + 3*Abs[x - y]^3
F[n_] := With[{m = Floor[(n - 2)/2]},

With[{m1 = If[OddQ[n], m + 1, m]},

Inactive[Integrate][(4/10 + 9/2 s[1]^2 * s[m1 + 1]^2) *

Product[Kt[s[i], s[i + 1]], {i, 1, m}] *

If[OddQ[n], K[s[m1], s[m1 + 1]], 1], ###] & @@

Table[{s[k], 0, 1}, {k, 1, m1 + 1}]]
Integrate[K[x, x], {x, 0, 1}]
1
Integrate[K[x, y]*K[y, x], {x, 0, 1}, {y, 0, 1}]
2
—
5
Activate[F[3]]
8
—
35
Activate[F[4]]
24
—
175
Activate[F[5]]
32
—
385
Activate[F[6]]
44 224
—
875 875
Activate[F[7]]
768
—
25 025
Activate[F[8]]
1 388 928
—
74 449 375

```

Code 1: Calculation of the trace of powers of the linear operator induced by the kernel  $\tilde{k}$  in Lemma 11 in Wolfram Mathematica. For  $n > 8$  this becomes quite slow.

### 4.2.5. FEATURE SELECTION USING FALSE DISCOVERY RATE CONTROL

When performing feature selection based on some ranking of the features (e.g. in terms of the observed AUC) it is challenging to decide how many features to select: Selecting only very few features might drop very important features (e.g. degrading classification performance) whereas selecting too many might not reduce the dimensionality of the problem sufficiently and/or keep many irrelevant features. If marginal  $p$ -values can be obtained (which is the case for the AUC and  $\xi_n$  as well as their resampled counterparts) this trade-off can be tackled by using a procedure to control the false discovery rate (FDR) when performing multiple comparisons. This section gives a short introduction to FDR control and discusses how to evaluate the stability of the proposed feature selection method, which, given a sample  $Y \in \mathbb{R}^{n \times p}$  and  $X \in \{0, 1\}^n$ , consists of two steps:

4

1. Calculate marginal test statistics and  $p$ -values, that is the  $j$ -th  $p$ -value (for  $j = 1, \dots, p$ ) is calculated from the sample  $(Y_{\cdot, j}, X)$  consisting of the  $j$ -th column of  $Y$  and the random vector  $X$ .
2. Select features (that is a  $I \subset [p]$ ) according to the Benjamin-Yekutieli procedure, controlling the FDR at the desired level  $\alpha$ .

When simultaneously testing multiple hypotheses a type-I error can be made for each hypothesis. Therefore it is, in general, not sufficient to control the type-I error at the nominal significance level for each of the tests. Instead  $V$ , the (random) number of incorrectly rejected null hypotheses, needs to be controlled. Controlling the probability  $\mathbb{P}(V \geq 1)$  is called family-wise error rate (FWER) control. For large scale problems, methods controlling the FWER are usually too strict. A good alternative is controlling the FDR, which is the expected value of the ratio  $V/R$  (where the random variable  $R$  is the total number of rejected null hypotheses). In [BY01] a procedure, that is usually called the Benjamini-Yekutieli (BY) procedure, is proposed which controls the FDR under arbitrary dependencies between the hypotheses under consideration. Since the dependency structure is unknown for the application considered in this paper, this is the procedure that will be utilized. Of course the proposed approach works with other methods for FDR control as well, when these are appropriate.

When calculating the test statistics and  $p$ -values in step 1 it is necessary to decide how  $I^{(1)}, \dots, I^{(l)}$  are drawn. If, for each a marginal, an independent sample is drawn the procedure is called independent-component bootstrap. It has been argued by [HM09] that this can be reasonable even if the component vectors are not independent and this approach was successfully applied by [NBD21] to estimate the proportion of true null hypotheses under dependency. Nonetheless using the same sample for all marginals seems more appropriate for the goal of this paper:

- It turns out that, when calculating (4.6) or (4.7) for large data sets, much of the computational effort is spent on generating pseudo-random numbers. Thus it is desirable to use as few pseudo-random numbers as possible.
- From a theoretical point of view, using different samples could lead to different values of the test statistic and the estimated  $p$ -values for very similar marginal samples. This is obviously undesirable and therefore it is preferable to use the same approximation  $\tilde{T}_{(m,n)}^{\text{AUC}, \ell}$  (or  $\tilde{T}_{(m,n)}^{\xi, \ell}$ ) for all of the simultaneous tests.

Furthermore it is necessary choose an appropriate  $m$ . The heuristic proposed in [BS08] suggested, for the examples discussed in the following section, that choosing  $m$  very small is inappropriate, which is unsurprising since the number of values the resampled test statistic can take is very



small if  $m$  is very small. Since, by the results of the previous sections, the variance of the AUC and  $\xi$  goes to zero if  $n_0, n_1 \rightarrow \infty$  it might as first seem like it would be a good idea to choose an  $m$  on the same order of magnitude as the sample size  $n$ . However, as  $\ell \rightarrow \infty$  the variance of e.g.  $\tilde{T}_{(m,n)}^{\xi,\ell}$  depends only on the covariance between  $Z_1 := \xi_{\tilde{n}_0, \tilde{n}_1}((X_{I_1}^{(1)}, Y_{I_1}^{(1)}), \dots, (X_{I_m}^{(1)}, Y_{I_m}^{(1)}))$  and  $Z_2 := \xi_{\tilde{n}_0, \tilde{n}_1}((X_{I_1}^{(2)}, Y_{I_1}^{(2)}), \dots, (X_{I_m}^{(2)}, Y_{I_m}^{(2)}))$  since

$$\text{Var}(\tilde{T}_{(m,n)}^{\xi,\ell}) = \ell^{-1} \text{Var}(V_1) + \frac{\ell-1}{\ell} \text{Cov}(V_1, V_2).$$

This covariance is, however, increasing in  $m$  (as has been checked numerically for the examples in the next section) and therefore a small, but not extremely small choice of  $m$  seems to be best.

To allow for a parallelized implementation of the proposed approach that yields, for a given seed (and up to numerical accuracy), reproducible results, the random number generator proposed [LEc99] and implemented in C++ as described in [LEc+02] is utilized.

Inspired by the work of [MB10] the stability of the proposed feature selection method will be evaluated in terms of the (estimated) number of features that have a very high selection probability (cf. [MB10, Definition 1]). This is done by considering a partition of the sample into  $k$  disjoint subsamples  $[n] = \bigcup_{i=1}^k B_i$  (as would be done for  $k$ -fold cross validation) and the number

$$S(M_s) := \left| \bigcap_{i=1}^k M_s(B_i) \right|$$

where  $M_s(B_i)$  are the indices of the  $s$  top ranked features (based on either on the AUC or  $\xi$ ), i.e.  $|M_s(B_i)| = s$ . That is  $S(M_s) \in [s]$  counts the number of features that are selected in each of the cross validation folds.

### 4.3. APPLICATION TO MALDI MSI DATA

This section demonstrates the proposed methodology by reanalysing the data-set presented in [Kri+16] for which 5 biomarkers (that is biologically meaningful  $m/z$  values) have been identified (cf. [Kri+16, Supplementary Table 1]). This data-set was subsequently analyzed by [Bos+17], [Beh+18] and [Leu+19]. For a detailed description of the data set as well the pre-processing see either of these papers. To make a comparison between the results presented in this paper and those of [Beh+18] and [Leu+19] simple the same 4-fold (respectively 8-fold) cross-validation strategy was employed. The pre-processing was performed as described in these papers with the following changes:

- In both cases the novel multiplicative-trend normalization (which was performed on a per core basis and using a running median filter with fixed window size 301) was performed before the usual total ion count (TIC) normalization.
- For the task considered in [Beh+18] the spectra were preprocessed using "Peptide Mass Resampling" (cf. [Bos+21]) resulting in 3046  $m/z$  channels. This was done to achieve a number of features similar to that on which the analysis in [Leu+19] was based.

Table 4.1 demonstrates that the proposed method is, when combined with a `randomForest` classifier (cf. [LW02]) with standard settings, competitive with the state of the art for the reanalysed data-set. For the first task  $\ell$  was chosen very small on purpose because otherwise (as was the case for the non-resampled  $\xi$ ) all of the features would have been chosen (when controlling the

Table 4.1: Average balanced accuracy achieved over the 8 (Task [Leu+19]) / 4 (Task ADSQ (Spot)) cross validation folds. The method IsotopeNet is that of [Beh+18] and Flog\_int is one of the methods described in [Leu+19]. The other methods are those described in this paper. Resampling was performed with subsample size  $m = 50$  and  $\ell = 100$  ( $\ell = 1000$ ) for the first (second) task. All  $p$ -values used for the FDR control in the feature selection procedures were estimated using  $10^5$  Monte Carlo samples.

Method	Task	
	[Leu+19]	ADSQ (Spot) [Beh+18]
IsotopeNet		0.845
Flog_int	0.904	
AUC and Random Forest	0.927	0.870
AUC (resampled) and Random Forest	0.927	0.867
$\xi$ and Random Forest	0.926	0.848
$\xi$ (resampled) and Random Forest	0.930	0.866

FDR at  $\alpha = 0.15$ ). This suggests that the hyperparameter  $\ell$  affects the ability of the test to reject the null-hypothesis. Indeed this effect can be seen in Figure 4.4.

In Figure 4.2 it can be seen that the resampling improves the selection stability of  $\xi$  while leaving that of the AUC mostly unchanged. Without the proposed trend normalisation the stability of feature selection based on the resampled  $\xi$  and the AUC is approximately the same. The stability of both methods is improved by the proposed trend normalisation. The improvement is, however, bigger for the AUC based feature selection. The test based on  $\xi$  is suitable against all alternatives, whereas the AUC is most suitable for testing for stochastic ordering of the two populations. Thus this empirical result indicates, that the proposed normalisation improves the stochastic ordering. This is very desirable since biomarkers in MALDI IMS data should generally result in pronounced differences in the intensities at specific  $m/z$ -values. Furthermore the empirical results in Figure 4.3 suggest that it would be worthwhile to switch from using the AUC to either  $\tau^*$  or  $D$ .

#### 4.4. STATISTICAL MODELLING OF MALDI MSI DATA

Since the 1990's finite mixture models have been recognized as a powerful tool for statistical modelling, which is seeing ever more widespread use; [MP00], [MLR19]. They are a convenient and flexible way to model unknown distributional shapes. Besides this they have a natural interpretation when there is group structure in the data, or the goal is explore data for such structure:

Let  $G$  be a discrete random variable taking values in  $\{1, \dots, k\}$  and let  $X$  be the random variable of interest. Furthermore let  $(X, G)$  have a joint density  $f_{X,G}$ . This joint density can be written as

$$f_{X,G}(x, g) = f_X(x|G = g) \cdot \mathbb{P}_G(G = g)$$

and by the law of total probability the marginal density  $f_X$  is given by

$$f_X(x) = \sum_{g=1}^k f_{X,G}(x, g) = \sum_{g=1}^k f_X(x|G = g) \cdot \mathbb{P}_G(G = g). \quad (4.11)$$

By letting  $\pi_i := \mathbb{P}_G(G = i)$ ,  $f_i(x) := f_X(x|G = i)$  ( $i = 1, \dots, k$ ) (where  $\pi_i$  are called the *mixture weights*) one obtains the usual form of a finite mixture distribution:

$$f_X(x) = \sum_{i=1}^k \pi_i \cdot f_i(x) \quad (4.12)$$

Since  $\mathbb{P}_G$  is a probability measure the  $\pi_i$  satisfy  $\pi_i \in [0, 1]$  for all  $i = 1, \dots, k$  and  $\sum_{i=1}^k \pi_i = 1$ .

Equation (4.11) suggests, that sampling  $X$  can be described in terms of sampling  $G$  and  $X_j \sim f_j$ : First obtain a realization  $j$  of  $G$  and then a realization  $x_j$  of  $X_j \sim f_j$ . Then  $x_j$  will be a realization of  $X$ . This suggests, that one may assume that each observation  $x$  of  $X$  has an associated realization  $g_x$  of  $G$ . The probability that  $g_x = j$  can be calculated according to Bayes' Rule

$$f_X(x|G=j) = \frac{\mathbb{P}_G(G=j|X=x) \cdot f_X(x)}{\mathbb{P}_G(G=j)}$$

which can be rewritten as

$$\mathbb{P}_G(G=j|X=x) = \frac{f_X(x|G=j) \cdot \mathbb{P}_G(G=j)}{f_X(x)} = \frac{f_j(x) \cdot \pi_j}{\sum_{i=1}^k \pi_i \cdot f_i(x)}. \quad (4.13)$$

The  $g_x$  (even if they are not observable) play an important role in estimating the mixture weight since they can be used to greatly simplify the Maximum-Likelihood (ML) estimation as detailed in the following section. Afterwards the Expectation-Maximization algorithm, which is a general method to deal with unobservable random variables that would be helpful to simplify the ML estimation, is introduced and its application to modelling MALDI spectra is discussed.

**Remark 7.** *Identifiability can be an issue with mixtures even if the component densities are identifiable [MP00, section 2.5]. The issue that components can be switched, if the densities are from the same parametric family, can usually be overcome by imposing the restriction ([MP00, p. 27])*

$$\pi_1 \leq \pi_2 \leq \dots \leq \pi_k$$

The Expectation-Maximization algorithm can be applied without taking account of this restriction.

#### 4.4.1. LIKELIHOOD OF THE WEIGHTS OF A FINITE MIXTURE

Let  $X_1 = x_1, \dots, X_n = x_n$  be an i.i.d. sample of  $X$  distributed according to the mixture (4.12), then the likelihood function is given by

$$L((\pi_1, \dots, \pi_k); (x_1, \dots, x_n)) = \prod_{i=1}^n f_X(x_i; \pi_i)$$

Maximizing the log-likelihood

$$\ln L((\pi_1, \dots, \pi_k); (x_1, \dots, x_n)) = \sum_{i=1}^n \ln \left( \sum_{j=1}^k \pi_j \cdot f_j(x_i) \right)$$

with respect to  $(\pi_1, \dots, \pi_k)$  is difficult because it involves the logarithm of a sum which depends on the  $\pi_i$ . However, if the associated group labels  $g_i$  of  $G_i$  (where  $G_i$  is, as in the previous section, the r.v. that gives the mixture component to which  $X_i$  belongs) had been observed the realizations of the auxiliary r.v.'s ( $j = 1, \dots, k, i = 1, \dots, n$ )

$$Z_{i,j}(G_i) := \begin{cases} 1 & \text{if } G_i = j \\ 0 & \text{otherwise} \end{cases}$$

would yield a *complete data* log-likelihood (cf. [MP00, Eq. (2.26)])

$$\ln L_c((\pi_1, \dots, \pi_k); (x_1, \dots, x_n, g_1, \dots, g_k))$$

$$\begin{aligned}
&= \ln \prod_{i=1}^n \sum_{j=1}^k z_{i,j} \cdot \pi_j \cdot f_j(x_i) \\
&= \ln \prod_{i=1}^n \pi_{g_i} \cdot f_{g_i}(x_i) \tag{4.14}
\end{aligned}$$

$$= \ln \prod_{i=1}^n \prod_{j=1}^k (\pi_j \cdot f_j(x_i))^{z_{i,j}} \tag{4.15}$$

$$= \sum_{i=1}^n \sum_{j=1}^k z_{i,j} \cdot (\ln \pi_j + \ln f_j(x_i)) \tag{4.16}$$

4

that is easy to maximize with respect to  $(\pi_1, \dots, \pi_k)$  (cf. Section 4.4.3) and where

- (4.14) holds since  $z_{i,j} \cdot \pi_j \cdot f_j(x_i) = 0 \cdot \pi_j \cdot f_j(x_i) = 0$  if  $j \neq g_i$  and  $z_{i,j} \cdot \pi_j \cdot f_j(x_i) = 1 \cdot \pi_j \cdot f_j(x_i) = \pi_j \cdot f_j(x_i)$  if  $j = g_i$
- and (4.15) holds since  $(\pi_j \cdot f_j(x_i))^{z_{i,j}} = (\pi_j \cdot f_j(x_i))^0 = 1$  if  $j \neq g_i$  and  $(\pi_j \cdot f_j(x_i))^{z_{i,j}} = (\pi_j \cdot f_j(x_i))^1 = \pi_j \cdot f_j(x_i)$  if  $j = g_i$ .

If  $g_i$  is not available, then the Expectation-Maximization algorithm, which is a type of coordinate ascent, can be used. It is detailed in the following section.

#### 4.4.2. THE EXPECTATION MAXIMIZATION (EM) ALGORITHM

Following [JB01] the Expectation-Maximization algorithm can be derived as follows: Let  $X$  be the observable random variable of interest and  $Z$  be the unobserved auxiliary random variable. Furthermore let  $X, Z$  have a joint density  $f_{X,Z}(\cdot; \theta_0)$ .

By Jensen's inequality it follows that, for any density  $q(z)$  over the sample space  $E_Z$  of  $Z$ , the log-likelihood of an observation  $X = x$  satisfies

$$\begin{aligned}
\ln L(\theta; x) &= \ln \int_{E_Z} f_{X,Z}(x, z; \theta) \, dz \\
&= \ln \int_{E_Z} q(z) \cdot \frac{f_{X,Z}(x, z; \theta)}{q(z)} \, dz \\
&\geq \int_{E_Z} q(z) \cdot \ln \frac{f_{X,Z}(x, z; \theta)}{q(z)} \, dz \\
&=: l(q(z), \theta, x).
\end{aligned}$$

One can now propose to maximize  $\ln L(\theta; x)$  by maximizing the lower bound  $l(q(z; x), \theta)$ . A suitable procedure is coordinate ascent in  $q$  and  $\theta$ :

- (1)  $q^{(n+1)} := \arg \max_q l(q, \theta^{(n)})$
- (2)  $\theta^{(n+1)} := \arg \max_{\theta} l(q^{(n+1)}, \theta)$

Step (1) is solved by  $q^{(n+1)}(z; x, \theta^{(n)}) = f_Z(z; \theta^{(n)} | X = x)$  since

$$l(f_Z(z; \theta^{(n)} | X = x), \theta^{(n)}) = \int_{E_Z} f_Z(z; \theta^{(n)} | X = x) \cdot \ln \frac{f_{X,Z}(x, z; \theta^{(n)})}{f_Z(z; \theta^{(n)} | X = x)} \, dz$$

$$\begin{aligned}
&= \int_{E_Z} f_Z(z; \theta^{(n)} | X = x) \cdot \ln \frac{f_{X,Z}(x, z; \theta^{(n)})}{\frac{f_{X,Z}(x, z; \theta^{(n)})}{f_X(x; \theta^{(n)})}} dz \\
&= \int_{E_Z} f_Z(z; \theta^{(n)} | X = x) \cdot \ln f_X(x; \theta^{(n)}) dz \\
&= \ln f_X(x; \theta^{(n)}) \cdot \int_{E_Z} f_Z(z; \theta^{(n)} | X = x) dz \\
&= \ln f_X(x; \theta^{(n)}) \\
&= \ln L(\theta^{(n)}; x)
\end{aligned}$$

and  $\ln L(\theta^{(n)}; x)$  is an upper bound for  $l(\cdot, \theta^{(n)})$  which has thus been maximized.

Noticing that

$$\begin{aligned}
l(q(z), \theta, x) &= \int_{E_Z} q(z) \cdot \ln \frac{f_{X,Z}(x, z; \theta)}{q(z)} dz \\
&= E_q \left[ \ln \frac{f_{X,Z}(x, z; \theta)}{q(z)} \right] \\
&= E_q [\ln f_{X,Z}(x, z; \theta)] - E_q [\ln q(z)]
\end{aligned}$$

it follows that, since the second term does not depend on  $\theta$ , step (2) can equivalently maximize  $E_q [\ln f_{X,Z}(x, z; \theta)]$ .

In summary an extremum of  $L(\theta)$  can be found by repeating:

**E step** Calculate  $E_q [\ln f_{X,Z}(x, z; \theta)]$  where  $q = f_Z(z; \theta^{(n)} | X = x)$

**M step** Find  $\theta^{(n+1)} = \arg \max_{\theta} E_q [\ln f_{X,Z}(x, z; \theta)]$

The function

$$\ln L_c(\theta; x, z) := \ln f_{X,Z}(x, z; \theta)$$

is usually called the *complete data log-Likelihood*.

#### 4.4.3. ESTIMATION OF MIXTURE WEIGHTS USING THE EM ALGORITHM

Let  $\theta := (\pi_1, \dots, \pi_k)$ ,  $\theta^{(m)} := (\pi_1^{(m)}, \dots, \pi_k^{(m)})$  and (cf. [MP00, Eq. (2.29)])

$$\begin{aligned}
\langle z_{i,j}^{(m)} \rangle &:= E_{\theta^{(m)}} [Z_{i,j} | X_i = x_i] \\
&= 1 \cdot \mathbb{P}_{\theta^{(m)}} (Z_{i,j} = 1 | X_i = x_i) + 0 \cdot \mathbb{P}_{\theta^{(m)}} (Z_{i,j} = 0 | X_i = x_i) \\
&= \mathbb{P}_{\theta^{(m)}} (Z_{i,j} = 1 | X_i = x_i) = \mathbb{P}_{\theta^{(m)}} (G_i = j | X_i = x_i) \\
&\stackrel{(4.13)}{=} \frac{f_j(x_i) \cdot \pi_j^{(m)}}{\sum_{l=1}^k \pi_l^{(m)} \cdot f_l(x_i)}.
\end{aligned} \tag{4.17}$$

Then the  $(m+1)$ -th **E step** is given by

$$Q(\theta; \theta^{(m)}) = E_{\theta^{(m)}} [\ln L_c(\theta; (x_1, \dots, x_n, g_1, \dots, g_k)) | X_1 = x_1, \dots, X_n = x_n]$$

$$\begin{aligned}
&= \sum_{i=1}^n \sum_{j=1}^k E_{\theta^{(m)}} [z_{i,j} | X_1 = x_1, \dots, X_n = x_n] \cdot (\ln \pi_j + \ln f_j(x_i)) \\
&= \sum_{i=1}^n \sum_{j=1}^k \langle z_{i,j}^{(m)} \rangle \cdot (\ln \pi_j + \ln f_j(x_i))
\end{aligned}$$

The **M step** is a solution  $\theta^{(m+1)}$  of

$$\theta^{(m+1)} \in \arg \max_{\theta \in [0,1]^k} Q(\theta; \theta^{(m)}) \quad (4.18)$$

subject to the constraint  $\sum_{j=1}^k \pi_j^{(m+1)} = 1$ .

The local extrema of  $Q$  can be found by considering the Lagrangian

$$\Lambda(\theta, \lambda) := \sum_{i=1}^n \sum_{j=1}^k \langle z_{i,j}^{(m)} \rangle \cdot (\ln \pi_j + \ln f_j(x_i)) + \lambda \cdot \left( \sum_{j=1}^k \pi_j - 1 \right)$$

Solutions of

$$\nabla_{\pi_1, \dots, \pi_k} \Lambda(\theta, \lambda) = 0, \quad \sum_{j=1}^k \pi_j = 1$$

yield necessary conditions for a local extremum of the constrained optimization problem (4.18).

The gradient of  $\Lambda$  is given by ( $j = 1, \dots, k$ )

$$\frac{\partial \Lambda(\theta, \lambda)}{\partial \pi_j} = \sum_{i=1}^n \frac{\langle z_{i,j}^{(m)} \rangle}{\pi_j} + \lambda$$

This is zero (for any  $j = 1, \dots, k$ ) if and only if

$$\pi_j = \frac{\sum_{i=1}^n \langle z_{i,j}^{(m)} \rangle}{-\lambda} \quad (4.19)$$

Inserting this into the constraint  $\sum_{j=1}^k \pi_j = 1$  results in

$$1 = \sum_{j=1}^k \frac{\sum_{i=1}^n \langle z_{i,j}^{(m)} \rangle}{-\lambda}$$

which is solved by

$$\lambda = - \sum_{j=1}^k \sum_{i=1}^n \langle z_{i,j}^{(m)} \rangle$$

Substituting this into (4.19) yields

$$\pi_j^{(m+1)} := \pi_j = \frac{\sum_{i=1}^n \langle z_{i,j}^{(m)} \rangle}{\sum_{i=1}^n \sum_{j=1}^k \langle z_{i,j}^{(m)} \rangle} = \frac{\sum_{i=1}^n \langle z_{i,j}^{(m)} \rangle}{n} \quad (4.20)$$

Since all  $\langle z_{i,j}^{(m)} \rangle$  are positive (4.20) is positive and less than one. Thus, if (4.20) is a local maximum of (4.18) (which needs to be checked), the **M step** is given by the update rule (4.20); cf. [MP00, Eq. (2.32)].

#### 4.4.4. PROPOSED MIXTURE MODEL AND INITIAL ESTIMATION

To gain some more insights into the statistical properties of MALDI MSI I propose a marginal model in this section. It is based on the idea that the observations are rescaled ion counts which are sampled according to an (unknown) probability, which is related to the amount of analyte. This seems sensible in light of the results by [Bae+12, Table 1], which demonstrate that the total number of ions stays approximately constant while ions from the analyte increase when the analyte concentration is increased.

Since the observations can be from different types of (cancerous) tissue and there may also be technical variation I propose to use a mixture of binomials that is scaled with some (unknown) factor  $\alpha$  since only a signal (roughly) proportional to the number of ions is measured. Specifically for  $n$  MALDI MSI spectra  $X \in \mathbb{R}_+^{n \times m}$  it is assumed that  $a_i := \alpha n_i = \|X_{i,\cdot}\|_1 = \sum_{j=1}^m X_{i,j}$  and that  $\frac{X_{i,j}}{\alpha} \sim \text{Bin}(n_i, p)$  with  $\alpha$ ,  $n_i$  and  $p$  unknown. As the following will only deal with single marginal model I will, for notational convenience, fix an (arbitrary)  $j \in [m]$  and write  $X_i$  instead of  $X_{i,j}$ . Thus the observations will be  $(X_1, a_1), \dots, (X_n, a_n)$  and they are assumed to be independent. The goal is to estimate the parameters  $(\pi_1, p_1), \dots, (\pi_K, p_K)$  from these observations, the relationship  $a_i = \alpha n_i$  and the mixture model

$$\tilde{X}_i := \frac{X_i}{\alpha} \sim \sum_{k=1}^K \pi_k \cdot f_{\text{Bin}}(x; n_i, p_k) \quad (4.21)$$

Parameter estimation under this model is difficult because the observations are not identically distributed.

To get good initial parameter estimates, the method of moments estimator proposed by [Bli64] is a good starting point. It is based on solving a system of equations involving the first  $2K - 1$  population factorial moments (cf. [Bli64, Eq. (3)])

$$F_k := \frac{1}{m} \sum_{i=1}^m \frac{\tilde{X}_i(\tilde{X}_i - 1) \cdots (\tilde{X}_i - k + 1)}{n_i(n_i - 1) \cdots (n_i - k + 1)},$$

which can, however, unfortunately not be calculated using only  $(X_i, a_i)$ . But, since  $n_i$  and  $\tilde{X}_i$  can be assumed to be large in comparison to the number of mixture components  $K$  and thus the first  $2K - 1$  factorial moments can be approximated by

$$F_k \approx \tilde{F}_k := \frac{1}{m} \sum_{i=1}^m \left( \frac{\tilde{X}_i}{n_i} \right)^k = \frac{1}{m} \sum_{i=1}^m \left( \frac{X_i}{a_i} \right)^k \quad (4.22)$$

which can be calculated directly from the observations. Using a polynomial root finder (cf. [SG12]) the methodology described in [Bli64] is easy to implement when using the approximation (4.22); cf. Code 2.

#### 4.4.5. IMPROVEMENT OF THE INITIAL ESTIMATES USING EM STEPS

The binomial densities in (4.21) can be approximated by  $f_{\text{Bin}}(\cdot; n_i, p_k) \approx \phi(\cdot | a_i \cdot p_k, a_i \cdot \alpha p_k(1 - p_k))$ . Letting  $\mu_k := p_k$  and  $\sigma_k^2 := \alpha p_k(1 - p_k)$  this allows for overdispersion and the EM-steps are, apart from the presence of the  $a_i$ 's, very close to the steps for a standard gaussian mixture:

**E step** Calculate

$$\gamma_k(X_i) = \frac{\pi_k \cdot \phi(X_i | \mathbf{a}_i \cdot \mu_k, \mathbf{a}_i \cdot \sigma_k^2)}{\sum_{k=1}^K \pi_k \cdot \phi(X_i | \mathbf{a}_i \cdot \mu_k, \mathbf{a}_i \cdot \sigma_k^2)}$$

```

1 import PolynomialRoots
2
3 function bin_mix_mom_approx(x,a,K)
4     r = K;
5
6     F = map(k -> begin
7         sum((x ./ a) .^ k)/size(x,1)
8     end,1:(2*r-1))
9
10    phi = -1 .* F[r:2*r-1];
11
12    D = Array{Float64,2}(undef,(r,r));
13    D[1,1] = 1;
14    D[1,2:end] = F[1:r-1];
15    for i=2:r
16        D[i,1:r] = F[i-1:r+i-2]
17    end
18
19    beta = D\phi;
20
21    p = (sort(abs.(PolynomialRoots.roots([beta; 1]))));
22
23    alpha = map(i -> begin
24        idx = setdiff(1:r,i);
25        s = 0;
26        for k=0:r-1
27            f = 1
28            if k<r-1
29                f = F[r-1-k];
30            end
31            s += (-1)^k * f * sum(map(idx -> begin
32                prod(p[idx])
33            end,Combinatorics.combinations(idx,k)));
34        end
35        s/prod(p[i] .- p[idx])
36    end,1:r)
37
38    alpha[alpha .< 0] .= 0.0;
39    alpha ./= sum(alpha);
40
41    return (alpha,p)
42 end

```

Code 2: Binomial mixture estimation using the method of moments proposed by [Bli64] and the approximation (4.22).



**M step** Update the parameters according to

$$\begin{aligned}\mu_k &= \frac{\sum_{i=1}^n \gamma_k(X_i) \cdot X_i}{\sum_{i=1}^n \gamma_k(X_i) \cdot \mathbf{a}_i} \\ (\sigma_k)^2 &= \frac{\sum_{i=1}^n \frac{\gamma_k(X_i)}{\mathbf{a}_i} \cdot (X_i - \mathbf{a}_i \mu_k)^2}{\sum_{i=1}^n \gamma_k(X_i)} \\ \pi_k &= \frac{\sum_{i=1}^n \gamma_k(X_i)}{n}\end{aligned}$$

Since the likelihood for gaussian mixture model is not bounded (cf. [RI02]) it is necessary to penalise the variance. This can be easily incorporated into the M-step by changing the update rule for the variance (cf. [RI02, Eqns. (9), (21)]):

$$(\sigma_k)^2 = \frac{2\alpha + \sum_{i=1}^n \frac{\gamma_k(X_i)}{\mathbf{a}_i} \cdot (X_i - \mathbf{a}_i \mu_k)^2}{2\beta + \sum_{i=1}^n \gamma_k(X_i)}$$

Initial estimates for  $\mu_k = p_k$  and  $\pi_k$  can be obtained using the method described in the previous section. The variances should then be chosen small enough that the components of the mixture do not overlap a lot.

Figure 4.5 compares the proposed modelling to the kernel density estimator (with and without L1 (TIC) normalization) by considering the difference between the densities for the two different classes for the known biomarker "Keratin type II cytoskeletal 7" (that is a biologically meaningful m/z value, cf. [Kri+16, Supplementary Table 1]) and for an m/z value "Actin" which is known not to be a biomarker. The result is, that the difference at the biomarker stays similar whereas the difference at the non-biomarker location is smaller.

## 4.5. CONCLUSION

This chapter introduced a trend normalization for MALDI MSI data and an approach to feature selection that is based on the subsampling bootstrap and FDR control. It furthermore presented the finite sample distribution of the correlation coefficient  $\xi_n$  recently proposed by [Cha20] under the setting relevant for this chapter. Then an application to MALDI mass spectroscopy data of different correlation coefficients and the proposed normalization was presented. For the properties of the resampled  $\xi_n$  the results of this application suggest, that the number of subsamples  $\ell$  plays the role of a hyper-parameter that controls the ability of the test to reject the null-hypothesis. It would therefore be interesting to, in future research, investigate how this is related to the effect size for specific alternatives and if there is a principled way to calibrate  $\ell$  such that the test becomes less sensitive towards effect sizes smaller than some minimal effect. The empirical results also suggest that using  $\tau_n^*$  or  $D_n$  instead of the often used AUC to perform feature selection for MALDI MSI data could lead to better classification procedures.

Furthermore it could be interesting to consider methods to control the resampling risk introduced by Monte Carlo approximation of the  $p$ -values, that is the risk that for some hypothesis the test decision is different from the one based on the theoretical  $p$ -value. Towards this end [Gan09] proposed a sequential implementation of Monte Carlo tests with uniformly bounded resampling risk, which could be applied to estimate the  $p$ -values required by the method proposed in this paper. However, convergence has been observed to be extremely slow and therefore the application is unfortunately not straightforward.

Finally some results regarding a marginal statistical model for MALDI observations were presented.

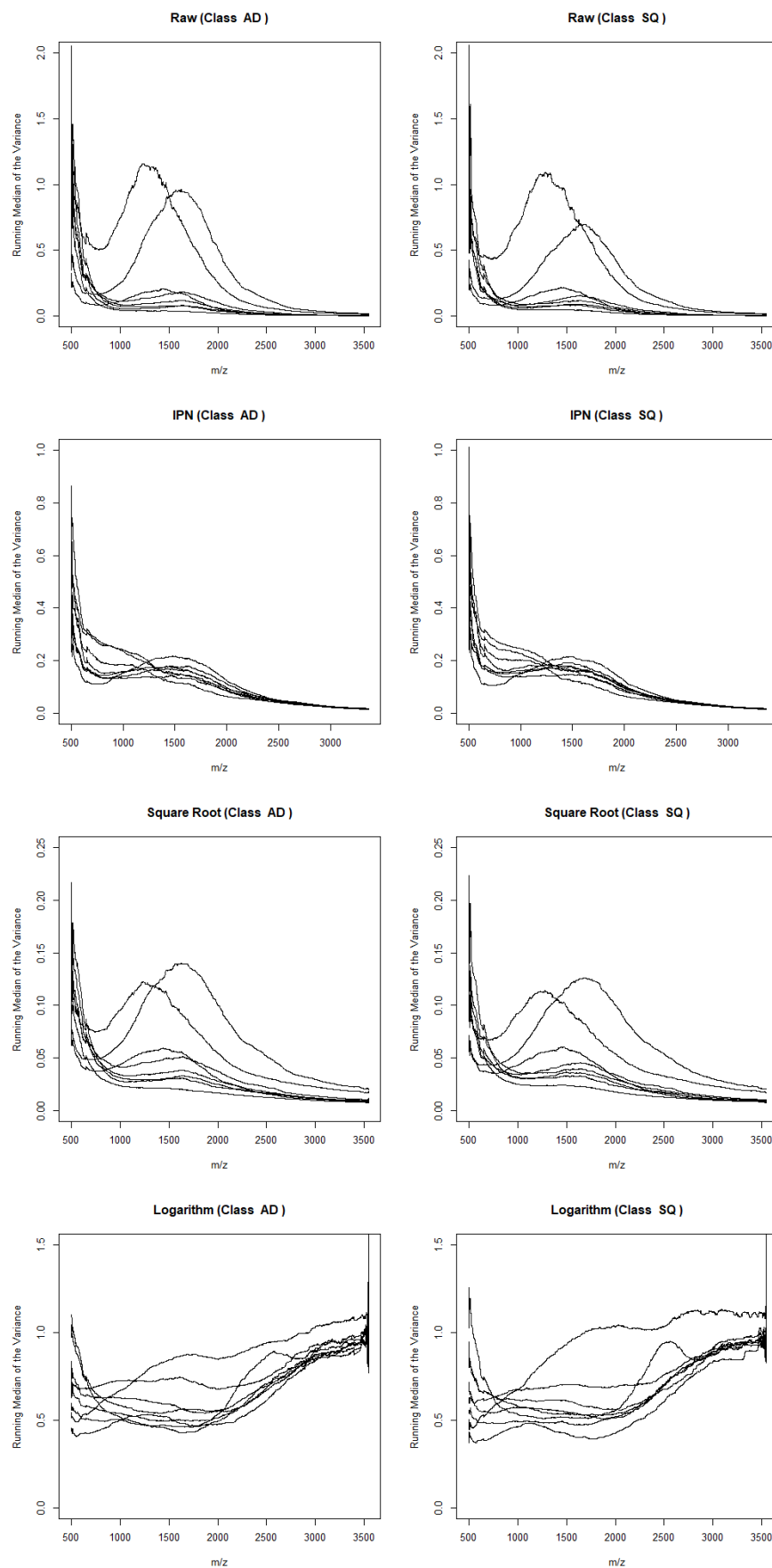


Figure 4.1: The running median (window size of 301  $m/z$  values) of the marginal variance. Each line corresponds to one TMA of the data set of [Kri+16] and one of the classes adenocarcinoma (AD, pictures in the left column) or squamous cell carcinoma (SQ, pictures in the right column). The data set was preprocessed using the "Peptide Mass Resampling" method (cf. [Bos+21]). IPN, Square root and Logarithm indicate, that these transformations were applied additionally. Raw means that no other transformation was applied.

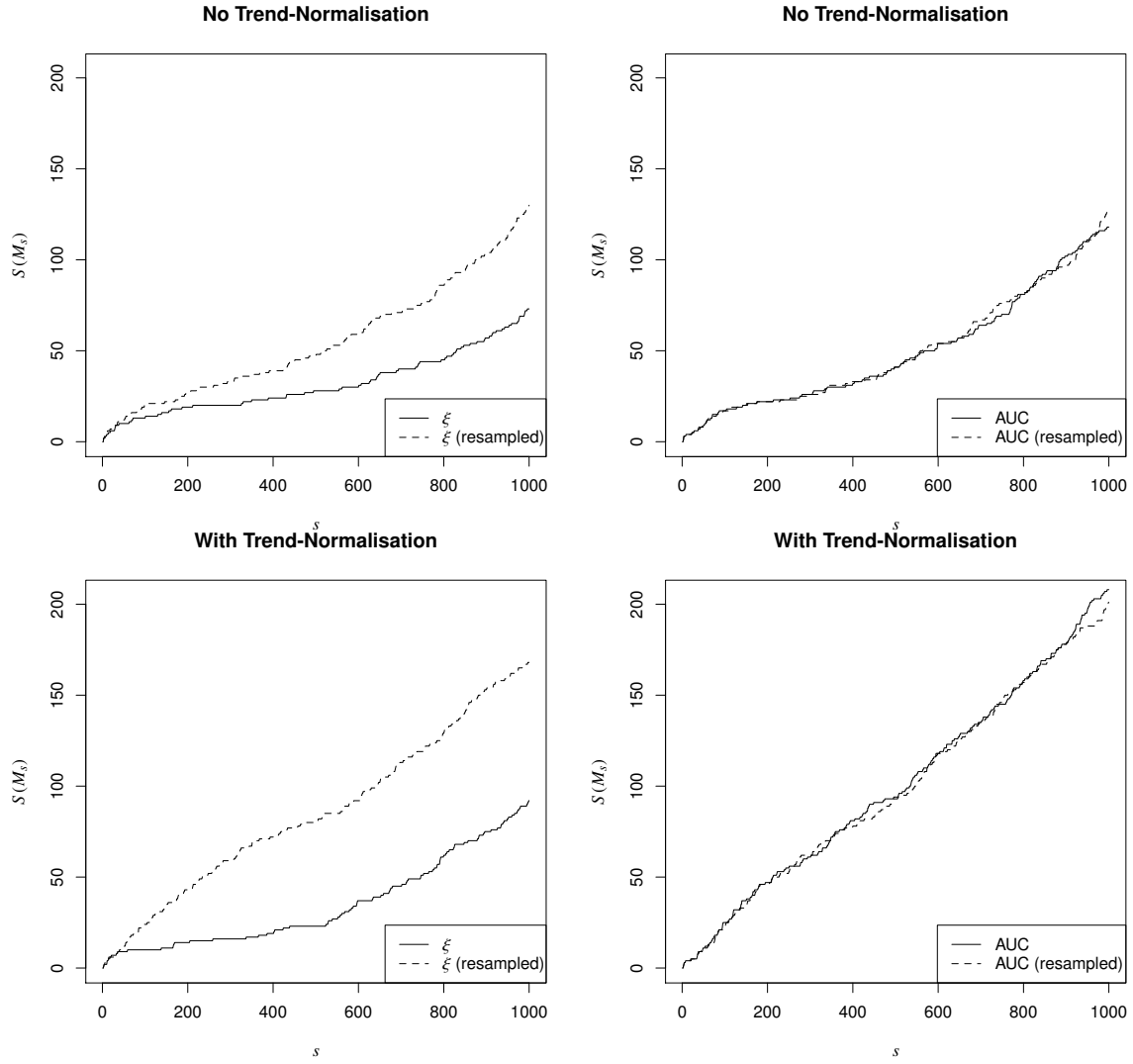


Figure 4.2: The number of stably selected features (as a function of the number of selected features  $s$ ) for the ADSQ (Spot) task of [Beh+18]. Resampling was performed with subsample size  $m = 50$  and  $\ell = 1000$  subsamples.

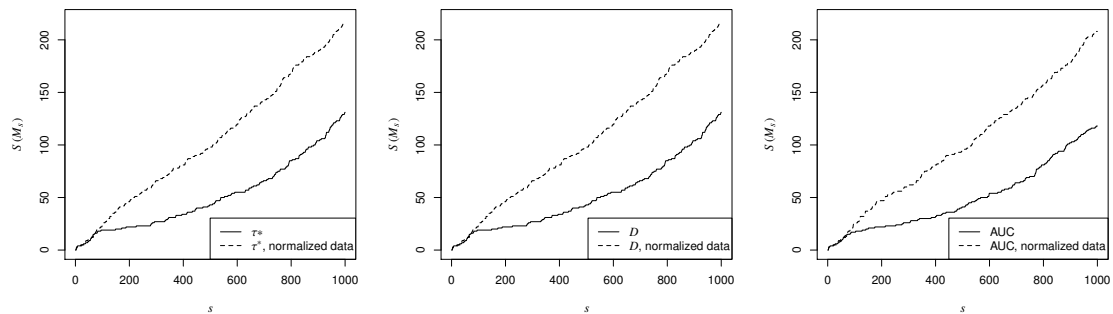


Figure 4.3: The number of stably selected features (as a function of the number of selected features  $s$ ) for the ADSQ (Spot) task of [Beh+18]. All coefficients' selection stability is positively affected by the proposed normalisation.  $D$  and  $\tau^*$  perform essentially equivalently well and slightly better than the AUC.

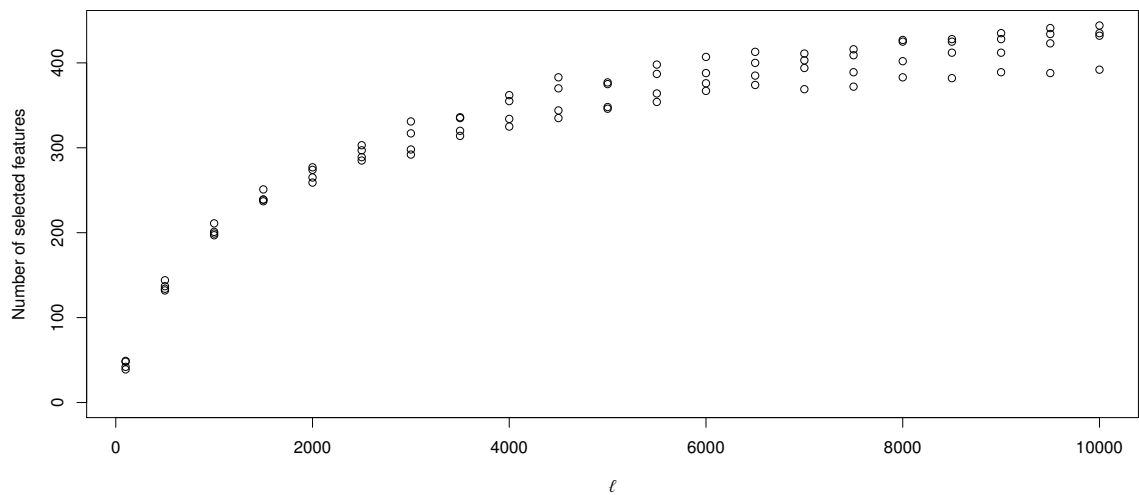


Figure 4.4: The number of features selected using the resampled version of  $\xi$  when controlling the FDR at level  $\alpha = 0.15$  (using the BY procedure) as function of the number of subsamples  $\ell$  and for the cross validation folds of the ADSQ (Spot) task of [Beh+18].

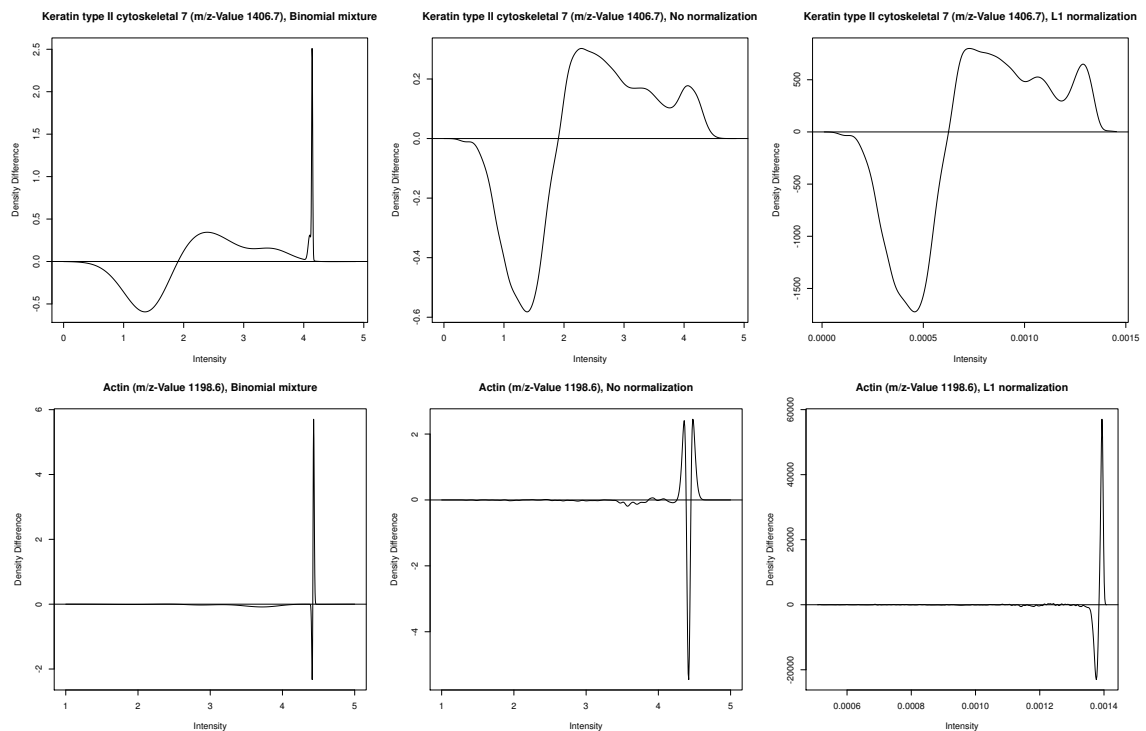


Figure 4.5: Comparison of the proposed marginal model with the usual kernel density estimator at a known biomarker location "Keratin type II cytoskeletal 7" (that is a biologically meaningful m/z value, cf. [Kri+16, Supplementary Table 1]) and for an m/z value "Actin" which is known not to be a biomarker for either of the cancer types. Depicted is the difference between the density estimated within each of the two classes (cancer types) present in the data set.

# 5

## CONCLUSION

*And ye shall know the truth and the truth shall make you free.*

John VIII-XXXII

This thesis demonstrates the multitude of applications that non-parametric statistical methods have for practically relevant and challenging tasks.

Chapter 3 describes how non-parametric methods can be utilised to improve the analysis of portfolio risk by solving the estimation problem implicit in "Reverse stress testing". This is practically very important as manual scenario selection needs to be at least supplanted with data-driven methods to avoid falling prey to human bias. I am hopeful that, in the future, it will be possible to develop fully non-parametric methods that are applicable to skew-elliptical models. Until then it will be helpful for practitioners to keep in mind that assuming an elliptical model is oftentimes not compatible with known facts about financial data. Thus it would be prudent to also sample scenarios from (a neighbourhood of) the subspace identified in Theorem 6 to guard against the effects of skewness.

Performing a stable selection of good features is important because, in many applications (including MALDI MSI), further investigation of features is quite expensive and time-consuming. Thus any improvement that can be obtained by better statistical methods, better preprocessing but also better understanding the limits of the methods utilised is of great value. Indeed the literature on feature selection is still developing quite dynamically and I believe that, apart from the methods described in this thesis, there is still a lot of potential to fruitfully apply new feature selection methods to MALDI IMS data to further the progress towards digital pathology. Specifically, since during MALDI imaging molecules break apart, a method that selects features according to conditional independence (cf. e.g. [AC21]) seems promising.

Understanding the performance and limitations of statistical methods is quite challenging, and therefore tools like the methods described in Chapter 2 that aid e.g. in the analysis of the power of statistical tests are quite helpful. While writing the paper that is reproduced in Chapter 2 I noticed how important the actual performance of the implementation of statistical methods can be, since it may otherwise be infeasible to apply them to real world problems. Thus for all methods described in this thesis I tried to provide insights into the practical and computational aspects.




## ACKNOWLEDGEMENTS

I gratefully acknowledge the support I received from the Deutsche Forschungsgemeinschaft (DFG) within the framework of RTG 2224 " $\pi^3$ : Parameter Identification - Analysis, Algorithms, Applications" during my time as PhD student.

I am especially thankful for all the help and scientific input I received from my supervisor Prof. Dr. Thorsten Dickhaus. Under his supervision I started to deepen my statistical knowledge as master student and he gave me the opportunity to continue this journey as a PhD student. He introduced me to Assoc.-Prof. Dr. Taras Bodnar, with whom I had the privilege to collaborate on a scientific publication.

Furthermore I would like to thank Prof. Dr. Werner Brannath for his helpful input as part of my scientific advisory committee.

I would also like to thank Dr. Tobias Boskamp and Dr. Delf Lachmund for introducing me to the fascinating topic of matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI MSI).

A special thank you goes to my family and friends, without whose patience and unconditional support I would not be where I am today 





## BIBLIOGRAPHY

- [AA20] Christopher J. Adcock and Adelchi Azzalini. “A Selective Overview of Skew-Elliptical and Related Distributions and of Their Applications”. In: *Symmetry* 12.1 (2020), p. 118. doi: [10.3390/sym12010118](https://doi.org/10.3390/sym12010118).
- [AC21] Mona Azadkia and Sourav Chatterjee. *A simple measure of conditional dependence*. 2021. arXiv: [1910.12327](https://arxiv.org/abs/1910.12327) [math.ST].
- [Adc10] Christopher J. Adcock. “Asset Pricing and Portfolio Selection Based on the Multivariate Extended Skew-Student-t Distribution”. In: *Annals of Operations Research* 176.1 (2010), pp. 221–234. doi: [10.1007/s10479-009-0586-4](https://doi.org/10.1007/s10479-009-0586-4).
- [Adc14] Christopher J. Adcock. “Mean–Variance–Skewness Efficient Surfaces, Stein’s Lemma and the Multivariate Extended Skew-Student Distribution”. In: *European Journal of Operational Research* 234.2 (2014), pp. 392–401. doi: [10.1016/j.ejor.2013.07.011](https://doi.org/10.1016/j.ejor.2013.07.011).
- [AEL15] Christopher J. Adcock, Martin Eling, and Nicola Loperfido. “Skewed Distributions in Finance and Actuarial Science: A Review”. In: *The European Journal of Finance* 21.13–14 (2015), pp. 1253–1281. doi: [10.1080/1351847X.2012.720269](https://doi.org/10.1080/1351847X.2012.720269).
- [AR12] Adelchi Azzalini and Giuliana Regoli. “Some Properties of Skew-Symmetric Distributions”. In: *Annals of the Institute of Statistical Mathematics* 64.4 (2012), pp. 857–879. doi: [10.1007/s10463-011-0338-5](https://doi.org/10.1007/s10463-011-0338-5).
- [Ay+17] Nihat Ay et al. *Information Geometry. A Series of Modern Surveys in Mathematics* 64. Springer International Publishing, 2017. ISBN: 978-3-319-56477-7. doi: [10.1007/978-3-319-56478-4](https://doi.org/10.1007/978-3-319-56478-4).
- [Azz13] Adelchi Azzalini. *The Skew-Normal and Related Families*. Cambridge: Cambridge University Press, 2013. ISBN: 978-1107029279. doi: [10.1017/CBO9781139248891](https://doi.org/10.1017/CBO9781139248891).
- [Azz19] Adelchi Azzalini. *The R package sn: The Skew-Normal and Related Distributions such as the Skew-t*. R package version 1.5-4. 2019. URL: <http://azzalini.stat.unipd.it/SN>.
- [BA15] Dean A. Bodenham and Niall M. Adams. “A Comparison of Efficient Approximations for a Weighted Sum of Chi-Squared Random Variables”. In: *Statistics and Computing* 26 (2015), pp. 917–928. doi: [10.1007/s11222-015-9583-4](https://doi.org/10.1007/s11222-015-9583-4).
- [Bae+12] Yong Jin Bae et al. “Degree of Ionization in MALDI of Peptides: Thermal Explanation for the Gas-Phase Ion Formation”. In: *Journal of the American Society for Mass Spectrometry* 23.8 (June 2012), pp. 1326–1335. doi: [10.1007/s13361-012-0406-y](https://doi.org/10.1007/s13361-012-0406-y).
- [Ban09] Basel Committee on Banking Supervision. *Principles for sound stress testing practices and supervision*. Bank for International Settlements, 2009. ISBN: 92-9131-784-5. URL: <https://www.bis.org/publ/bcbs155.pdf>.
- [Bar36] M. S. Bartlett. “The Square Root Transformation in Analysis of Variance”. In: *Supplement to the Journal of the Royal Statistical Society* 3.1 (1936), p. 68. doi: [10.2307/2983678](https://doi.org/10.2307/2983678).

- [BB14] Sharmodeep Bhattacharyya and Peter J. Bickel. “Adaptive Estimation in Elliptical Distributions with Extensions to High Dimensions”. preprint on webpage at [https://sites.science.oregonstate.edu/~bhattash/Site/Research\\_files/mixture\\_elliptic.pdf](https://sites.science.oregonstate.edu/~bhattash/Site/Research_files/mixture_elliptic.pdf). 2014.
- [BB89] R. B. Bapat and M. I. Beg. “Order Statistics for Nonidentically Distributed Variables and Permanents”. In: *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)* 51.1 (1989), pp. 79–93. URL: <http://www.jstor.org/stable/25050725>.
- [BC18] Arup Bose and Snigdhasu Chatterjee. *U-Statistics, Mm-Estimators and Resampling*. 75th ed. Texts and Readings in Mathematics. Springer Singapore, 2018. ISBN: 978-981-13-4756-6. doi: [10.1007/978-981-13-2248-8](https://doi.org/10.1007/978-981-13-2248-8).
- [BD14] Wicher Bergsma and Angelos Dassios. “A Consistent Test of Independence Based on a Sign Covariance Related to Kendall’s Tau”. In: *Bernoulli* 20.2 (2014), pp. 1006–1028. doi: [10.3150/13-bej514](https://doi.org/10.3150/13-bej514).
- [BD17] Taras Bodnar and Thorsten Dickhaus. “On the Simes Inequality in Elliptical Models”. In: *Annals of the Institute of Statistical Mathematics* 69.1 (2017), pp. 215–230. doi: [10.1007/s10463-015-0539-4](https://doi.org/10.1007/s10463-015-0539-4).
- [BDE03] Wolfgang Breymann, Alexandra Dias, and Paul Embrechts. “Dependence Structures For Multivariate High-Frequency Data in Finance”. In: *Quantitative Finance* 3.1 (2003), pp. 1–14. doi: [10.1080/713666155](https://doi.org/10.1080/713666155).
- [Beh+18] Jens Behrmann et al. “Deep Learning For Tumor Classification in Imaging Mass Spectrometry”. In: *Bioinformatics* 34.7 (2018), pp. 1215–1223. doi: [10.1093/bioinformatics/btx724](https://doi.org/10.1093/bioinformatics/btx724).
- [Bez+17] Jeff Bezanson et al. “Julia: A fresh approach to numerical computing”. In: *SIAM review* 59.1 (2017), pp. 65–98. doi: [10.1137/141000671](https://doi.org/10.1137/141000671).
- [BG15] Taras Bodnar and Arjun K. Gupta. “Robustness of the Inference Procedures for the Global Minimum Variance Portfolio Weights in a Skew-Normal Model”. In: *The European Journal of Finance* 21.13-14 (2015), pp. 1176–1194. doi: [10.1080/1351847X.2012.696073](https://doi.org/10.1080/1351847X.2012.696073).
- [BH95] Yoav Benjamini and Yosef Hochberg. “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing”. In: *Journal of the royal statistical society. Series B (Methodological)* 57.1 (1995), pp. 289–300. doi: [10.1111/j.2517-6161.1995.tb02031.x](https://doi.org/10.1111/j.2517-6161.1995.tb02031.x).
- [BKR61] J. R. Blum, J. Kiefer, and M. Rosenblatt. “Distribution Free Tests of Independence Based on the Sample Distribution Function”. In: *The Annals of Mathematical Statistics* 32.2 (1961), pp. 485–498. doi: [10.1214/aoms/1177705055](https://doi.org/10.1214/aoms/1177705055).
- [Bla+10] Gilles Blanchard et al. “ $\mu$ TOSS - Multiple hypothesis testing in an open software system”. In: *Proceedings of the First Workshop on Applications of Pattern Analysis*. Ed. by Tom Diethe, Nello Cristianini, and John Shawe-Taylor. Vol. 11. Proceedings of Machine Learning Research. Cumberland Lodge, Windsor, UK: PMLR, Sept. 2010, pp. 12–19. URL: <http://proceedings.mlr.press/v11/blanchard10a.html>.
- [Bla+14] Gilles Blanchard et al. “On Least Favorable Configurations for Step-Up-Down Tests”. In: *Statistica Sinica* 24.1 (2014), pp. 1–23. doi: [10.5705/ss.2011.205](https://doi.org/10.5705/ss.2011.205).
- [Bli64] W. R. Blischke. “Estimating the Parameters of Mixtures of Binomial Distributions”. In: *Journal of the American Statistical Association* 59.306 (June 1964), pp. 510–528. doi: [10.1080/01621459.1964.10482176](https://doi.org/10.1080/01621459.1964.10482176).

- [BMP19] Taras Bodnar, Stepan Mazur, and Nestor Parolya. “Central Limit Theorems for Functionals of Large Sample Covariance Matrix and Mean Vector in Matrix-Variate Location Mixture of Normal Distributions”. In: *Scandinavian Journal of Statistics* 46.2 (2019), pp. 636–660. doi: [10.1111/sjos.12383](https://doi.org/10.1111/sjos.12383).
- [Bod+19] Taras Bodnar et al. “Tests for the Weights of the Global Minimum Variance Portfolio in a High-Dimensional Setting”. In: *IEEE Transactions on Signal Processing* 67.17 (2019), pp. 4479–4493. doi: [10.1109/TSP.2019.2929964](https://doi.org/10.1109/TSP.2019.2929964).
- [Bod09] Taras Bodnar. “An Exact Test on Structural Changes in the Weights of the Global Minimum Variance Portfolio”. In: *Quantitative Finance* 9.3 (2009), pp. 363–370. doi: [10.1080/14697680802446748](https://doi.org/10.1080/14697680802446748).
- [Bod16] Dean Bodenham. *momentchi2: Moment-Matching Methods for Weighted Sums of Chi-Squared Random Variables*. R package version 0.1.5. 2016. URL: <https://CRAN.R-project.org/package=momentchi2>.
- [Bog07] Vladimir I. Bogachev. *Measure Theory*. Vol. 1. Springer Berlin Heidelberg, 2007. ISBN: 978-3-540-34513-8. doi: [10.1007/978-3-540-34514-5](https://doi.org/10.1007/978-3-540-34514-5).
- [Bos+17] Tobias Boskamp et al. “A New Classification Method for MALDI Imaging Mass Spectrometry Data Acquired on Formalin-Fixed Paraffin-Embedded Tissue Samples”. In: *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1865.7 (2017), pp. 916–926. doi: [10.1016/j.bbapap.2016.11.003](https://doi.org/10.1016/j.bbapap.2016.11.003).
- [Bos+21] Tobias Boskamp et al. “Cross-Normalization of MALDI Mass Spectrometry Imaging Data Improves Site-to-Site Reproducibility”. In: *Analytical Chemistry* 93.30 (July 2021), pp. 10584–10592. doi: [10.1021/acs.analchem.1c01792](https://doi.org/10.1021/acs.analchem.1c01792).
- [Box76] George E. P. Box. “Science and Statistics”. In: *Journal of the American Statistical Association* 71.356 (Dec. 1976), pp. 791–799. doi: [10.1080/01621459.1976.10480949](https://doi.org/10.1080/01621459.1976.10480949).
- [BP09] Prabir Burman and Wolfgang Polonik. “Multivariate Mode Hunting: Data Analytic Tools with Measures of Significance”. In: *Journal of Multivariate Analysis* 100.6 (2009), pp. 1198–1218. doi: [10.1016/j.jmva.2008.10.015](https://doi.org/10.1016/j.jmva.2008.10.015).
- [Bra10] Michael W. Brandt. “Chapter 5 - Portfolio Choice Problems”. In: *Handbook of Financial Econometrics: Tools and Techniques*. Vol. 1. Handbooks in Finance. North-Holland, 2010, pp. 269–336. doi: [10.1016/B978-0-444-50897-3.50008-0](https://doi.org/10.1016/B978-0-444-50897-3.50008-0).
- [BS08] Peter J. Bickel and Anat Sakov. “On the Choice of  $m$  in the  $m$  out of  $n$  Bootstrap and Confidence Bounds for Extrema”. In: *Statistica Sinica* 18.3 (2008), pp. 967–985. URL: <http://www.jstor.org/stable/24308525>.
- [BY01] Yoav Benjamini and Daniel Yekutieli. “The Control of the False Discovery Rate in Multiple Testing Under Dependency”. In: *Annals of Statistics* 29.4 (2001), pp. 1165–1188. doi: [10.1214/aos/1013699998](https://doi.org/10.1214/aos/1013699998).
- [Cap12] Antonella Capitanio. *On the Canonical Form of Scale Mixtures of Skew-Normal Distributions*. 2012. arXiv: [1207.0797](https://arxiv.org/abs/1207.0797).
- [CGT03] John T. Chen, Arjun K. Gupta, and Cas G. Troskie. “The Distribution of Stock Returns When the Market is Up”. In: *Communications in Statistics-Theory and Methods* 32.8 (2003), pp. 1541–1558. doi: [10.1081/STA-120022244](https://doi.org/10.1081/STA-120022244).
- [CH13] Jiahua Chen and Yi Huang. “Finite-Sample Properties of the Adjusted Empirical Likelihood”. In: *Journal of Nonparametric Statistics* 25.1 (2013), pp. 147–159. doi: [10.1080/10485252.2012.738906](https://doi.org/10.1080/10485252.2012.738906).

- [Cha20] Sourav Chatterjee. “A New Coefficient of Correlation”. In: *Journal of the American Statistical Association* (2020), pp. 1–21. doi: [10.1080/01621459.2020.1758115](https://doi.org/10.1080/01621459.2020.1758115).
- [Con01] Rama Cont. “Empirical Properties of Asset Returns: Stylized Facts and Statistical Issues”. In: *Quantitative Finance* 1.2 (2001), pp. 223–236. doi: [10.1080/713665670](https://doi.org/10.1080/713665670).
- [CR16] Jiahao Chen and Jarrett Revels. “Robust benchmarking in noisy environments”. In: *arXiv e-prints* (Aug. 2016). arXiv: [1608.04295](https://arxiv.org/abs/1608.04295) [cs.PF].
- [Dei+11] Sören-Oliver Deininger et al. “Normalization in MALDI-TOF imaging datasets of proteins: practical considerations”. In: *Analytical and Bioanalytical Chemistry* 401.1 (Apr. 2011), pp. 167–181. doi: [10.1007/s00216-011-4929-z](https://doi.org/10.1007/s00216-011-4929-z).
- [Dhr13] Phoebus J. Dhrymes. *Mathematics for Econometrics*. 4th ed. Springer, 2013. ISBN: 978-1461481447. doi: [10.1007/978-1-4614-8145-4](https://doi.org/10.1007/978-1-4614-8145-4).
- [Dic14] Thorsten Dickhaus. *Simultaneous Statistical Inference: With Applications in the Life Sciences*. Springer, Berlin, Heidelberg, 2014. ISBN: 978-3-642-45181-2. doi: [10.1007/978-3-642-45182-9](https://doi.org/10.1007/978-3-642-45182-9).
- [Dic18] Thorsten Dickhaus. *Theory of Nonparametric Tests*. Springer International Publishing, 2018. ISBN: 978-3-319-76314-9. doi: [10.1007/978-3-319-76315-6](https://doi.org/10.1007/978-3-319-76315-6).
- [DR15] Sylvain Delattre and Etienne Roquain. “New Procedures Controlling the False Discovery Proportion via Romano-Wolf’s Heuristic”. In: *The Annals of Statistics* 43.3 (2015), pp. 1141–1177. doi: [10.1214/14-AOS1302](https://doi.org/10.1214/14-AOS1302).
- [DY02] Adrian A. Drăgulescu and Victor M. Yakovenko. “Probability Distribution of Returns in the Heston Model with Stochastic Volatility”. In: *Quantitative Finance* 2.6 (2002), pp. 443–453. doi: [10.1080/14697688.2002.0000011](https://doi.org/10.1080/14697688.2002.0000011).
- [Edd13] Dirk Eddelbuettel. *Seamless R and C++ Integration with Rcpp*. Vol. 64. Use R! Springer, 2013. ISBN: 978-1-4614-6868-4. doi: [10.1007/978-1-4614-6868-4](https://doi.org/10.1007/978-1-4614-6868-4).
- [Elt+14] Edwin J. Elton et al. *Modern Portfolio Theory and Investment Analysis*. 9th ed. John Wiley & Sons, 2014.
- [FDR09] Helmut Finner, Thorsten Dickhaus, and Markus Roters. “On the False Discovery Rate and an Asymptotically Optimal Rejection Curve”. In: *Annals of Statistics* 37.2 (2009), pp. 596–618. doi: [10.1214/07-AOS569](https://doi.org/10.1214/07-AOS569).
- [FG18] Helmut Finner and Veronika Gontscharuk. “Two-Sample Kolmogorov-Smirnov-Type Tests Revisited: Old and New Tests in Terms of Local Levels”. In: *Annals of Statistics* 46.6A (2018), pp. 3014–3037. doi: [10.1214/17-AOS1647](https://doi.org/10.1214/17-AOS1647).
- [FGD12] H. Finner, V. Gontscharuk, and T. Dickhaus. “False Discovery Rate Control of Step-Up-Down Tests with Special Emphasis on the Asymptotically Optimal Rejection Curve”. In: *Scandinavian Journal of Statistics* 39.2 (2012), pp. 382–397. doi: [10.1111/j.1467-9469.2012.00791.x](https://doi.org/10.1111/j.1467-9469.2012.00791.x).
- [FK15] Mark D. Flood and George G. Korenko. “Systematic Scenario Selection: Stress Testing and the Nature of Uncertainty”. In: *Quantitative Finance* 15.1 (2015), pp. 43–59. doi: [10.1080/14697688.2014.926018](https://doi.org/10.1080/14697688.2014.926018).
- [FP10] Michael P. Fay and Michael A. Proschan. “Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules”. In: *Statistics Surveys* 4.0 (2010), pp. 1–39. doi: [10.1214/09-ss051](https://doi.org/10.1214/09-ss051).
- [FZ06] José A. Ferreira and Aeilko Zwinderman. “Approximate Sample Size Calculations with Microarray Data: An Illustration”. In: *Statistical Applications in Genetics and Molecular Biology* 5.1 (2006). doi: [10.2202/1544-6115.1227](https://doi.org/10.2202/1544-6115.1227).

- [Gan09] Axel Gandy. “Sequential Implementation of Monte Carlo Tests With Uniformly Bounded Resampling Risk”. In: *Journal of the American Statistical Association* 104.488 (2009), pp. 1504–1511. doi: [10.1198/jasa.2009.tm08368](https://doi.org/10.1198/jasa.2009.tm08368).
- [GDG04] Graciela González-Farías, J. Armando Domínguez-Molina, and Arjun K. Gupta. “The Closed Skew-Normal Distribution”. In: *Skew-Elliptical Distributions and their Applications: A Journey Beyond Normality*. Ed. by Marc G. Genton. Chapman and Hall/CRC, 2004, pp. 25–42. doi: [10.1201/9780203492000](https://doi.org/10.1201/9780203492000).
- [GKK15] Paul Glasserman, Chulmin Kang, and Wanmo Kang. “Stress Scenario Selection by Empirical Likelihood”. In: *Quantitative Finance* 15.1 (2015), pp. 25–41. doi: [10.1080/14697688.2014.926019](https://doi.org/10.1080/14697688.2014.926019).
- [GLF15] Veronika Gontscharuk, Sandra Landwehr, and Helmut Finner. “The Intermediates Take It All: Asymptotics of Higher Criticism Statistics and a Powerful Alternative Based on Equal Local Levels”. In: *Biometrical Journal* 57.1 (2015), pp. 159–180. doi: [10.1002/bimj.201300255](https://doi.org/10.1002/bimj.201300255).
- [GLF16] Veronika Gontscharuk, Sandra Landwehr, and Helmut Finner. “Goodness of Fit Tests in Terms of Local Levels with Special Emphasis on Higher Criticism Tests”. In: *Bernoulli* 22.3 (2016), pp. 1331–1363. doi: [10.3150/14-BEJ694](https://doi.org/10.3150/14-BEJ694).
- [Glu+08a] Deborah Glueck et al. “Exact calculations of average power for the Benjamini-Hochberg procedure”. In: *The International Journal of Biostatistics* 4.1 (2008). doi: [10.2202/1557-4679.1103](https://doi.org/10.2202/1557-4679.1103).
- [Glu+08b] Deborah H. Glueck et al. “Fast Computation by Block Permanents of Cumulative Distribution Functions of Order Statistics from Several Populations”. In: *Communications in Statistics—Theory and Methods* 37.18 (2008), pp. 2815–2824. doi: [10.1080/03610920802001896](https://doi.org/10.1080/03610920802001896).
- [GN16] Evarist Giné and Richard Nickl. *Mathematical Foundations of Infinite-Dimensional Statistical Models*. Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, 2016. ISBN: 978-1-107-04316-9. doi: [10.1017/cbo9781107337862](https://doi.org/10.1017/cbo9781107337862).
- [Gov12] Board of Governors of the Federal Reserve System. *SR 12-7: Supervisory Guidance on Stress Testing for Banking Organizations with More Than \$10 Billion in Total Consolidated Assets*. Last Accessed: 2019-04-04. 2012. URL: <https://www.federalreserve.gov/supervisionreg/srletters/sr1207.htm>.
- [Gry21] Karol Gryszka. “From Biased Coin to Any Discrete Distribution”. In: *Periodica Mathematica Hungarica* 83 (2021), pp. 71–80. doi: [10.1007/s10998-020-00363-w](https://doi.org/10.1007/s10998-020-00363-w).
- [GVB13] Arjun K. Gupta, Tamas Varga, and Taras Bodnar. *Elliptically Contoured Models in Statistics and Portfolio Theory*. 2nd ed. Springer, 2013. ISBN: 978-1-4614-8153-9. doi: [10.1007/978-1-4614-8154-6](https://doi.org/10.1007/978-1-4614-8154-6).
- [Ham+13] Toshimitsu Hamasaki et al. “Sample Size Determination for Clinical Trials with Co-Primary Outcomes: Exponential Event Times”. In: *Pharmaceutical Statistics* 12.1 (2013), pp. 28–34. doi: [10.1002/pst.1545](https://doi.org/10.1002/pst.1545).
- [Hes93] Steven L. Heston. “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”. In: *The Review of Financial Studies* 6.2 (1993), pp. 327–343. doi: [10.1093/rfs/6.2.327](https://doi.org/10.1093/rfs/6.2.327).
- [HM09] Peter Hall and Hugh Miller. “Using the Bootstrap to Quantify the Authority of an Empirical Ranking”. In: *The Annals of Statistics* 37.6B (2009), pp. 3929–3959. doi: [10.1214/09-aos699](https://doi.org/10.1214/09-aos699).



- [Hoe48] Wassily Hoeffding. "A Non-Parametric Test of Independence". In: *The Annals of Mathematical Statistics* 19.4 (1948), pp. 546–557. doi: [10.1214/aoms/1177730150](https://doi.org/10.1214/aoms/1177730150).
- [Hor19] Joel L. Horowitz. "Bootstrap Methods in Econometrics". In: *Annual Review of Economics* 11.1 (2019), pp. 193–224. doi: [10.1146/annurev-economics-080218-025651](https://doi.org/10.1146/annurev-economics-080218-025651).
- [HY13] Zhiqiu Hu and Rong-Cai Yang. "A New Distribution-Free Approach to Constructing the Confidence Region for Multiple Parameters". In: *PLOS ONE* 8.12 (2013), pp. 1–13. doi: [10.1371/journal.pone.0081179](https://doi.org/10.1371/journal.pone.0081179).
- [Ive10] Daniel H. Iversen. "Closed-Skew Distributions: Simulation, Inversion and Parameter Estimation". MA thesis. Norwegian University of Science and Technology, 2010. URL: [https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/258728/351897\\_FULLTEXT01.pdf](https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/258728/351897_FULLTEXT01.pdf).
- [Izm18] Grant Izmirlan. *Average Power and  $\lambda$ -power in Multiple Testing Scenarios when the Benjamini-Hochberg False Discovery Rate Procedure is used*. 2018. arXiv: [1801.03989](https://arxiv.org/abs/1801.03989).
- [JB01] Michael I. Jordan and Christopher M. Bishop. "Chapter 10 - The EM algorithm". In: *Introduction to Graphical Models*. 2001. URL: <http://people.csail.mit.edu/yks/documents/classes/mlbook/pdf/chapter10.pdf>.
- [Joh70] Roy A. Johnson. "Atomic and nonatomic measures". In: *Proceedings of the American Mathematical Society* 25.3 (1970), pp. 650–650. doi: [10.1090/s0002-9939-1970-0279266-8](https://doi.org/10.1090/s0002-9939-1970-0279266-8).
- [JR18] Claude-Pierre Jeannerod and Siegfried M. Rump. "On Relative Errors of Floating-point Operations: Optimal Bounds and Applications". In: *Mathematics of Computation* 87.310 (2018), pp. 803–819. doi: [10.1090/mcom/3234](https://doi.org/10.1090/mcom/3234).
- [Jun05] Sin-Ho Jung. "Sample Size for FDR-control in Microarray Data Analysis". In: *Bioinformatics* 21.14 (2005), pp. 3097–3104. doi: [10.1093/bioinformatics/bti456](https://doi.org/10.1093/bioinformatics/bti456).
- [Kab66] D. G. Kabe. "On the exact distribution of the Fisher-Behrens'-Welch statistic". In: *Metrika* 10.1 (Dec. 1966), pp. 13–15. doi: [10.1007/bf02613414](https://doi.org/10.1007/bf02613414).
- [Kar67] Stepan Karamardian. "Strictly Quasi-Convex (Concave) Functions and Duality in Mathematical Programming". In: *Journal of Mathematical Analysis and Applications* 20.2 (1967), pp. 344–358. doi: [https://doi.org/10.1016/0022-247X\(67\)90095-9](https://doi.org/10.1016/0022-247X(67)90095-9).
- [Kea] R. Baker Kearfott. "Interval Computations: Introduction, Uses, and Resources". URL: <https://interval.louisiana.edu/preprints/survey.pdf>.
- [Knu98] Donald E. Knuth. *The Art of Computer Programming, Volume 2: Seminumerical Algorithms*. 3rd ed. Addison-Wesley, 1998. ISBN: 0201896842.
- [Kop+15] Yaacov Kopeliovich et al. "Robust Risk Estimation and Hedging: A Reverse Stress Testing Approach". In: *The Journal of Derivatives* 22.4 (2015), pp. 10–25. doi: [10.3905/jod.2015.22.4.010](https://doi.org/10.3905/jod.2015.22.4.010).
- [KR12] Rekha P. Kulkarni and Akshay S. Rane. "Asymptotic Expansions for Approximate Eigenvalues of Integral Operators with Nonsmooth Kernels". In: *Numerical Functional Analysis and Optimization* 33.4 (2012), pp. 415–440. doi: [10.1080/01630563.2011.650811](https://doi.org/10.1080/01630563.2011.650811).

- [Kri+16] Mark Kriegsmann et al. “Reliable Entity Subtyping in Non-small Cell Lung Cancer by Matrix-assisted Laser Desorption/Ionization Imaging Mass Spectrometry on Formalin-fixed Paraffin-embedded Tissue Specimens”. In: *Molecular & Cellular Proteomics* 15.10 (2016), pp. 3081–3089. doi: [10.1074/mcp.m115.057513](https://doi.org/10.1074/mcp.m115.057513).
- [LD75] Erich L. Lehmann and Howard J. D’Abrera. *Nonparametrics: Statistical Methods Based on Ranks*. Holden-Day Series in Probability and Statistics. Holden-Day, 1975. ISBN: 0-8162-4996-6.
- [LEc+02] Pierre L’Ecuyer et al. “An Object-Oriented Random-Number Package with Many Long Streams and Substreams”. In: *Operations Research* 50.6 (2002), pp. 1073–1075. doi: [10.1287/opre.50.6.1073.358](https://doi.org/10.1287/opre.50.6.1073.358).
- [LEc99] Pierre L’Ecuyer. “Good Parameters and Implementations for Combined Multiple Recursive Random Number Generators”. In: *Operations Research* 47.1 (1999), pp. 159–164. doi: [10.1287/opre.47.1.159](https://doi.org/10.1287/opre.47.1.159).
- [Lee90] A. J. Lee. *U-statistics: Theory and Practice*. Statistics: Textbooks and Monographs 110. Boca Raton: CRC Press, 1990. ISBN: 9780824782535.
- [Leu+19] Johannes Leuschner et al. “Supervised non-negative matrix factorization methods for MALDI imaging applications”. In: *Bioinformatics* 35.11 (2019), pp. 1940–1947. doi: [10.1093/bioinformatics/bty909](https://doi.org/10.1093/bioinformatics/bty909).
- [LPB00] Bruce G. Lindsay, Ramani S. Pilla, and Prasanta Basak. “Moment-Based Approximations of Distributions Using Mixtures: Theory and Applications”. In: *Annals of the Institute of Statistical Mathematics* 52.2 (2000), pp. 215–230. doi: [10.1023/a:1004105603806](https://doi.org/10.1023/a:1004105603806).
- [LR05] E. L. Lehmann and J. P. Romano. *Testing Statistical Hypotheses*. Springer New York, 2005. doi: [10.1007/0-387-27605-x](https://doi.org/10.1007/0-387-27605-x).
- [LR17] Marko Lange and Siegfried Rump. “Faithfully Rounded Floating-point Computations”. Preprint, available at [www.ti3.tuhh.de/paper/rump/LaRu2017b.pdf](http://www.ti3.tuhh.de/paper/rump/LaRu2017b.pdf). 2017.
- [LV03] Zinoviy M. Landsman and Emiliano A. Valdez. “Tail Conditional Expectations for Elliptical Distributions”. In: *North American Actuarial Journal* 7.4 (2003), pp. 55–71. doi: [10.1080/10920277.2003.10596118](https://doi.org/10.1080/10920277.2003.10596118).
- [LW02] Andy Liaw and Matthew Wiener. “Classification and Regression by randomForest”. In: *R News* 2.3 (2002), pp. 18–22. URL: <https://CRAN.R-project.org/doc/Rnews/>.
- [Mar52] Harry Markowitz. “Portfolio Selection”. In: *The Journal of Finance* 7.1 (1952), pp. 77–91. doi: [10.2307/2975974](https://doi.org/10.2307/2975974).
- [MB10] Nicolai Meinshausen and Peter Bühlmann. “Stability Selection”. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 72.4 (2010), pp. 417–473. doi: [10.1111/j.1467-9868.2010.00740.x](https://doi.org/10.1111/j.1467-9868.2010.00740.x).
- [Med02] European Medicines Agency / Committee for Medicinal Products for Human Use. “Points to Consider on Multiplicity Issues in Clinical Trials”. available at [https://www.ema.europa.eu/en/documents/scientific-guideline/points-consider-multiplicity-issues-clinical-trials\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/points-consider-multiplicity-issues-clinical-trials_en.pdf). 2002.
- [MFE05] Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative Risk Management*. Princeton Series in Finance. Princeton University Press, 2005. ISBN: 9780691122557.
- [MLR19] Geoffrey J. McLachlan, Sharon X. Lee, and Suren I. Rathnayake. “Finite Mixture Models”. In: *Annual Review of Statistics and Its Application* 6.1 (2019), pp. 355–378. doi: [10.1146/annurev-statistics-031017-100325](https://doi.org/10.1146/annurev-statistics-031017-100325).

- [MN17] Amit Moscovich and Boaz Nadler. “Fast Calculation of Boundary Crossing Probabilities for Poisson Processes”. In: *Statistics & Probability Letters* 123 (2017), pp. 177–182. doi: [10.1016/j.spl.2016.11.027](https://doi.org/10.1016/j.spl.2016.11.027).
- [MNS16] Amit Moscovich, Boaz Nadler, and Clifford Spiegelman. “On the Exact Berk-Jones Statistics and their  $p$ -value Calculation”. In: *Electronic Journal of Statistics* 10.2 (2016), pp. 2329–2354. doi: [10.1214/16-EJS1172](https://doi.org/10.1214/16-EJS1172).
- [MP00] G. J. McLachlan and D. Peel. *Finite Mixture Models*. 2000.
- [NBD21] André Neumann, Taras Bodnar, and Thorsten Dickhaus. “Estimating the Proportion of True Null Hypotheses Under Dependency: A Marginal Bootstrap Approach”. In: *Journal of Statistical Planning and Inference* 210 (2021), pp. 76–86. doi: [10.1016/j.jspi.2020.04.011](https://doi.org/10.1016/j.jspi.2020.04.011).
- [Not+01] Daniel A. Notterman et al. “Transcriptional Gene Expression Profiles of Colorectal Adenoma, Adenocarcinoma, and Normal Tissue Examined by Oligonucleotide Arrays”. In: *Cancer Research* 61.7 (2001), pp. 3124–3130. URL: <https://pubmed.ncbi.nlm.nih.gov/11306497/>.
- [Owe01] Art B. Owen. *Empirical Likelihood*. 1st. Monographs on Statistics and Applied Probability. Chapman and Hall/CRC, 2001. doi: [10.1201/9781420036152](https://doi.org/10.1201/9781420036152).
- [PM20] Bernhard Pfaff and Alexander McNeil. *QRM: Provides R-Language Code to Examine Quantitative Risk Management Concepts*. R package version 0.4-20. 2020. URL: <https://CRAN.R-project.org/package=QRM>.
- [PS10] Belinda Phipson and Gordon K. Smyth. “Permutation P-values Should Never Be Zero: Calculating Exact P-values When Permutations Are Randomly Drawn”. In: *Statistical Applications in Genetics and Molecular Biology* 9.1 (2010). doi: [10.2202/1544-6115.1585](https://doi.org/10.2202/1544-6115.1585).
- [R C20] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria, 2020. URL: <https://www.R-project.org/>.
- [RI02] Andrea Ridolfi and Jérôme Idier. *Penalized Maximum Likelihood Estimator for Normal Mixtures*. Tech. rep. École Polytechnique Fédérale de Lausanne, 2002. URL: <https://pagesperso.ls2n.fr/~idier-j/pub/ridolfi02.pdf>.
- [RV11] Etienne Roquain and Fanny Villers. “Exact Calculations for False Discovery Proportion with Application to Least Favorable Configurations”. In: *The Annals of Statistics* 39.1 (2011), pp. 584–612. doi: [10.1214/10-AOS847](https://doi.org/10.1214/10-AOS847).
- [SB07] Stephen Senn and Frank Bretz. “Power and Sample Size when Multiple Endpoints are Considered”. In: *The Journal of Applied Statistics in the Pharmaceutical Industry* 6.3 (2007), pp. 161–170. doi: [10.1002/pst.301](https://doi.org/10.1002/pst.301).
- [SB18] Charlie Sharpsteen and Cameron Bracken. *tikzDevice: R Graphics Output in LaTeX Format*. R package version 0.12. 2018. URL: <https://CRAN.R-project.org/package=tikzDevice>.
- [Sch+19] Erik Schnetter et al. *eschnett/SIMD.jl: Explicit SIMD vectorization in Julia*. 2019. doi: [10.5281/ZENODO.3355421](https://doi.org/10.5281/ZENODO.3355421).
- [Sch20] Jonathan von Schroeder. *Stable Feature Selection with Applications to MALDI Imaging Mass Spectrometry Data*. 2020. arXiv: [2006.15077](https://arxiv.org/abs/2006.15077).
- [SDH20] Hongjian Shi, Mathias Drton, and Fang Han. “On the Power of Chatterjee Rank Correlation”. In: (2020). arXiv: [2008.11619](https://arxiv.org/abs/2008.11619).



- [SG12] J. Skowron and A. Gould. *General Complex Polynomial Root Solver and Its Further Optimization for Binary Microlenses*. 2012. arXiv: [1203.1034](#).
- [Shu17] Tomer Shushi. “Skew-Elliptical Distributions with Applications in Risk Theory”. In: *European Actuarial Journal* 7.1 (2017), pp. 277–296. doi: [10.1007/s13385-016-0144-9](#).
- [SNT17] Khamis K. Said, Wei Ning, and Yubin Tian. “Likelihood Procedure for Testing Changes in Skew Normal Model with Applications to Stock Returns”. In: *Communications in Statistics-Simulation and Computation* 46.9 (2017), pp. 6790–6802. doi: [10.1080/03610918.2016.1212067](#).
- [Spr07] Peter Sprent. *Applied nonparametric statistical methods*. Boca Raton: Chapman & Hall/CRC, 2007. ISBN: 978-1-4398-9401-9.
- [Stu97] Gerold Studer. “Maximum Loss for Measurement of Market Risk”. Last Accessed: 2019-04-04. PhD thesis. ETH Zurich, 1997. URL: <http://www2.risklab.ch/ftp/papers/ThesisGeroldStuder.pdf>.
- [SW09] Galen R. Shorack and Jon A. Wellner. *Empirical Processes with Applications to Statistics*. Vol. 59. Classics In Applied Mathematics. Society for Industrial and Applied Mathematics, 2009. ISBN: 978-0-89871-684-9. doi: [10.1137/1.9780898719017](#).
- [Tha01] J. Tharrats. *Structure of Binary Sequences*. 2001. arXiv: [math/0109182](#).
- [TP19] Phontita Thiuthad and Nabendu Pal. “Point Estimation of the Location Parameter of a Skew-Normal Distribution: Some Fixed Sample and Asymptotic Results”. In: *Journal of Statistical Theory and Practice* 13.2 (2019). doi: [10.1007/s42519-018-0033-4](#).
- [Tra+19] Pascal Traccucci et al. *A Triptych Approach for Reverse Stress Testing of Complex Portfolios*. 2019. arXiv: [1906.11186](#).
- [vT20] Jonathan von Schroeder and Dickhaus. Thorsten. “Efficient Calculation of the Joint Distribution of Order Statistics”. In: *Computational Statistics & Data Analysis* 144 (2020). doi: [10.1016/j.csda.2019.106899](#).
- [VTS17] Zsófia Varga, Yu C. Tsang, and Júlia Singer. “A Simple Procedure to Estimate the Optimal Sample Size in Case of Conjunctive Coprimary Endpoints”. In: *Biometrical Journal* 59.4 (2017), pp. 626–635. doi: [10.1002/bimj.201500231](#).
- [WDL15] Luca Weihs, Mathias Drton, and Dennis Leung. *Efficient Computation of the Bergsma-Dassios Sign Covariance*. 2015. arXiv: [1504.00964](#).
- [Wic16] Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. 2nd ed. Use R! Springer-Verlag New York, 2016. ISBN: 978-3-319-24277-4. URL: <http://ggplot2.org>.
- [Wil+19] Agata Wilk et al. “On Stability of Feature Selection Based on MALDI Mass Spectrometry Imaging Data and Simulated Biopsy”. In: *Advances in Intelligent Systems and Computing*. Ed. by Józef Korbicz. Springer International Publishing, 2019, pp. 82–93. doi: [10.1007/978-3-030-29885-2\\_8](#).
- [Wil38] Samuel S. Wilks. “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”. In: *The Annals of Mathematical Statistics* 9.1 (1938), pp. 60–62. doi: [10.1214/aoms/1177732360](#).
- [Wil45] Frank Wilcoxon. “Individual Comparisons by Ranking Methods”. In: *Biometrics Bulletin* 1.6 (Dec. 1945), p. 80. doi: [10.2307/3001968](#).
- [Wol+05] Witold E Wolski et al. “Transformation and other factors of the peptide mass spectrometry pairwise peak-list comparison process”. In: *BMC Bioinformatics* 6.1 (Nov. 2005). doi: [10.1186/1471-2105-6-285](#).

- [Yu09] Guan Yu. “Variance stabilizing transformations of Poisson, binomial and negative binomial distributions”. In: *Statistics & Probability Letters* 79.14 (2009), pp. 1621–1629. ISSN: 0167-7152. doi: [10.1016/j.spl.2009.04.010](https://doi.org/10.1016/j.spl.2009.04.010).
- [Zal02] C. Zălinescu. *Convex Analysis in General Vector Spaces*. World Scientific, 2002. doi: [10.1142/5021](https://doi.org/10.1142/5021).
- [Zha+16] Jin E. Zhang et al. “The Skewness Implied in the Heston Model and Its Application”. In: *Journal of Futures Markets* 37.3 (2016), pp. 211–237. doi: [10.1002/fut.21801](https://doi.org/10.1002/fut.21801).
- [Zho+14] Xiaoping Zhou et al. “Smooth Monotone Covariance for Elliptical Distributions and Applications in Finance”. In: *Quantitative Finance* 14.9 (2014), pp. 1555–1571. doi: [10.1080/14697688.2014.911949](https://doi.org/10.1080/14697688.2014.911949).

# A

## APPENDIX FOR CHAPTER 2

### A.1. WHEN IS THE RESULT FAITHFULLY ROUNDED?

By [LR17, Theorem 4.2] we can, by calculating  $k = k(n_1, n_2)$  according to Equation (11), figure out for which values of  $n_1$  and  $n_2$  faithful rounding is guaranteed. This can be done by examining the evaluation tree (cf. [LR17, Definition 2.2]) of our concrete implementation. The result will be faithfully rounded if no under- or overflow occurs, and  $k \leq 2^{26} - 2$  (when utilising double precision floating point numbers). For our concrete implementation we obtained  $k(n_1, n_2) = n_1 \cdot n_2 + 8 \cdot (n_1 + n_2) - 7$ , provided that  $n_1 + n_2 \geq 2$ . Thus (assuming that no over- or underflow occurs) the result is guaranteed to be faithfully rounded if  $n_1, n_2 \leq 8184$ . Our implementation could be, in terms of  $k$ , significantly improved by using binary summation. By this we mean arranging the  $n$  numbers that are to be summed in a binary tree (which has depth  $O(\log(n))$ ) and performing the summation by successively replacing the nodes (whose children are leafs) by the sum of their children until the root of the tree contains the result. Obviously this tree does not actually have to be built, see e.g. the function `binary_sum` in <https://github.com/jvschroeder/PairArithmetic/blob/master/PairArithmetic.hpp>. For example, for  $n_1 = n_2 = 400$  we obtain  $k(400, 400) = 166,398$ , while the corresponding number of  $k$  in the case of binary summation would equal 17,421. The latter improvement however comes at an additional computational cost, and may be considered mostly of theoretical interest since the calculation for  $n_1 = n_2 = 400$  already takes approximately ten minutes. For  $n_1 = n_2 = 800$  the calculation already takes more than four hours.

## A

## A.2. MULTIPLE HYPOTHESIS TESTING

This section gives a short overview of the definitions and results from multiple testing necessary for Section 2.6. The basic definition is that of a multiple test. It operates on a vector  $p$ -values  $\mathbf{p} = (p_1, \dots, p_m)^\top$  and is a measurable mapping  $\varphi : [0, 1]^m \rightarrow \mathcal{P}([m])$ , where hypothesis  $H_i$  is rejected iff  $i \in \varphi(\mathbf{p})$ . Under  $FM(m, m_0, F)$  denote by  $M_0 \equiv m_0$  a constant random variable. Then the (random) number of rejections of the multiple test  $\varphi$  is given by  $R(\varphi, \mathbf{p}) := |\varphi(\mathbf{p})|$ , and  $V(\varphi, \mathbf{p}) := |\varphi(\mathbf{p}) \cap [M_0]|$  is the (random) number of false rejections (type I errors).

In the following we will consider step-up procedures  $\varphi = \text{SU}_{\mathbf{t}}$  with critical values  $\mathbf{t} = (t_1, \dots, t_m)^\top \in (0, 1)^m$  such that  $t_1 \leq \dots \leq t_m$ . The corresponding decision rule can be written as

$$\text{SU}_{\mathbf{t}}(\mathbf{p}) := [\max(\{0\} \cup \{i \in [m] : p_{i:m} \leq t_i\})], \text{ where } [0] := \emptyset.$$

Summarizing results of [RV11], the joint distribution of  $R$  and  $V$  for any step-up procedure  $\text{SU}_{\mathbf{t}}$  has the following properties.

**Lemma 12.** Let  $0 \leq j \leq k \leq m$ .

(i) Under the unconditional model  $RM(m, \pi_0, F)$  it holds that

$$\begin{aligned} \mathbb{P}_{m, \pi_0, F}(V(\text{SU}_{\mathbf{t}}, \mathbf{p}) = j, R(\text{SU}_{\mathbf{t}}, \mathbf{p}) = k) \\ = \binom{m}{k} \binom{k}{j} \tilde{\pi}_0^j (1 - \tilde{\pi}_0)^{k-j} G(t_k)^k \Psi_{m-k, 0}^{\text{Un}[0, 1], F}(1 - G(t_m), \dots, 1 - G(t_{k+1})) \end{aligned}$$

where  $\tilde{\pi}_0 := \pi_0 t_k / G(t_k)$  and  $G(t) := \pi_0 t + (1 - \pi_0)F(t)$ .

(ii) Under the conditional model  $FM(m, m_0, F)$  it holds that

$$\begin{aligned} \mathbb{P}_{m, m_0, F}(V(\text{SU}_{\mathbf{t}}, \mathbf{p}) = j, R(\text{SU}_{\mathbf{t}}, \mathbf{p}) = k) \\ = \binom{m_0}{j} \binom{m - m_0}{k - j} t_k^j (F(t_k))^{k-j} \Psi_{m-k-(m_0-j), m_0-j}^{\text{Un}[0, 1], \bar{F}}(1 - t_m, \dots, 1 - t_{k+1}), \end{aligned}$$

where  $0 \leq j \leq m_0$  and  $\bar{F}(t) := 1 - F(t)$ .

Combining Lemma 12 with the previously discussed efficient evaluation of  $\Psi$  it is possible to calculate various summary statistics pertaining to the joint distribution of  $(V, R, M_0)$  under the above models:

**Definition 4.**

(a) The FDR of  $\text{SU}_{\mathbf{t}}(\mathbf{p})$  is given by the expectation of the false discovery proportion (FDP) of  $\text{SU}_{\mathbf{t}}(\mathbf{p})$ , which is given by

$$\text{FDP}(\text{SU}_{\mathbf{t}}, \mathbf{p}) := \frac{V(\text{SU}_{\mathbf{t}}, \mathbf{p})}{R(\text{SU}_{\mathbf{t}}, \mathbf{p}) \vee 1}.$$

(b) Considering the number of correct rejections  $R(\text{SU}_{\mathbf{t}}, \mathbf{p}) - V(\text{SU}_{\mathbf{t}}, \mathbf{p})$  the average power of  $\text{SU}_{\mathbf{t}}(\mathbf{p})$  is given by

$$\text{Pow}_{\text{avg}}(\text{SU}_{\mathbf{t}}) := \mathbb{E} \left[ \frac{R(\text{SU}_{\mathbf{t}}, \mathbf{p}) - V(\text{SU}_{\mathbf{t}}, \mathbf{p})}{m - M_0} \right], \quad (\text{A.1})$$

where the convention  $\frac{0}{0} = 0$  is utilized and where  $\mathbb{E} = \mathbb{E}_{m, \pi_0, F}$  (under  $RM(m, \pi_0, F)$ ) or  $\mathbb{E} = \mathbb{E}_{m, m_0, F}$  (under  $FM(m, m_0, F)$ ), respectively.

(c) The  $\lambda$ -power is the probability of rejecting at least  $\lambda \cdot (m - M_0)$  of the false hypotheses:

$$\text{Pow}_\lambda(SU_t) := \mathbb{P} \left( \frac{R(SU_t, \mathbf{p}) - V(SU_t, \mathbf{p})}{m - M_0} \geq \lambda \right) \quad (\text{A.2})$$

where, again, the convention  $\frac{0}{0} = 0$  is utilised and where  $\mathbb{P} = \mathbb{P}_{m, \pi_0, F}$  (under  $RM(m, \pi_0, F)$ ) or  $\mathbb{P} = \mathbb{P}_{m, m_0, F}$  (under  $FM(m, m_0, F)$ ), respectively.

## A

## A.3. FURTHER PROOFS AND ALGORITHMS

**Proof of Lemma 3.** Let  $c_X(t) := |\{U_i \leq t | i \in X\}|$ ,  $m_1 \in [n_1]$ ,  $m_2 \in [n_2]$  and  $m := m_1 + m_2$ . Then, mimicking the approach of [SW09, p. 367 ff.] and [Bla+14, Proposition 1], it holds that

$$\begin{aligned}
1 - \Psi(m_1, m_2) &:= 1 - \mathbb{P}(U_{1:m} \leq b_1, \dots, U_{m:m} \leq b_m) \\
&= 1 - \mathbb{P}\left(\bigcap_{k=1}^m c_{[m]}(b_k) \geq k\right) \\
&= \mathbb{P}(\exists k \in [m] : c_{[m]}(b_k) = k - 1) \\
&= \sum_{k=1}^m \mathbb{P}\left([c_{[m]}(b_k) = k - 1] \cap \left[\bigcap_{j=1}^{k-1} c_{[m]}(b_j) \geq j\right]\right) \\
&= \sum_{k=0}^{m-1} \mathbb{P}\left([c_{[m]}(b_{k+1}) = k] \cap \left[\bigcap_{j=1}^k c_{[m]}(b_j) \geq j\right]\right) \\
&= \sum_{k=0}^{m-1} \sum_{\substack{X \subset [m] \\ |X|=k}} \mathbb{P}\left(\left[\bigcap_{i \in [m] \setminus X} U_i > b_{k+1}\right] \cap \left[\bigcap_{j=1}^k c_X(b_j) \geq j\right]\right) \\
&= \sum_{\substack{0 \leq k_1 \leq m_1 \\ 0 \leq k_2 \leq m_2 \\ k_1 + k_2 < m}} \sum_{\substack{X \subset [m] \\ |X|=k_1+k_2 \\ |X \cap [m_1]|=k_1}} \mathbb{P}\left(\bigcap_{i \in [m] \setminus X} U_i > b_{k_1+k_2+1}\right) \cdot \mathbb{P}\left(\bigcap_{j=1}^{k_1+k_2} c_X(b_j) \geq j\right) \\
&= \sum_{\substack{0 \leq k_1 \leq m_1 \\ 0 \leq k_2 \leq m_2 \\ k_1 + k_2 < m}} \binom{m_1}{k_1} \binom{m_2}{k_2} (1 - b_{k_1+k_2+1})^{m_1-k_1} \cdot (1 - F(b_{k_1+k_2+1}))^{m_2-k_2} \\
&\quad \times \mathbb{P}(U_{1:(k_1+k_2)} \leq b_1, \dots, U_{(k_1+k_2):(k_1+k_2)} \leq b_{k_1+k_2})
\end{aligned}$$

Since this holds for any  $m_1 \in [n_1]$ ,  $m_2 \in [n_2]$  it follows that

$$\Psi(m_1, m_2) = 1 - \sum_{\substack{0 \leq k_1 \leq m_1 \\ 0 \leq k_2 \leq m_2 \\ k_1 + k_2 < m}} M_{k_1, k_2}^{(m_1, m_2)} \cdot \Psi(k_1, k_2)$$

where  $M$  is given by (2.2).

The recursions for  $M$  follow from the definition of the binomial coefficient and routine calculations.  $\square$

**Proof of Lemma 5.** Let  $X_1, \dots, X_{n_1} \sim \text{Uni}[0, 1]$  and  $X_{n_1+1}, \dots, X_{n_1+n_2} \sim F$  be jointly stochastically independent. Let for  $0 \leq i_1 \leq n_1$  and  $0 \leq i_2 \leq n_2$

$$\Psi(i_1, i_2) := \mathbb{P}(X_{1:M} \leq b_1, \dots, X_{i:M} \leq b_{i_1+i_2}),$$

where  $M := [i_1] \cup \{n_1 + j | j \in [i_2]\}$ ,  $[0] := \emptyset$  and  $X_{i:M}$  denotes the  $i$ -th order statistic of  $(X_j)_{j \in M}$ , and  $(b_i)_{i \in [n]}$  is an increasing sequence with values in  $[0, 1]$ . To simplify the notation let  $b_0 := 0$ .

To mimic the proof-technique of [SW09, p. 362 ff.]

$$Q_{i_1, i_2}(m) := \mathbb{P}\left(\bigcap_{j=1}^{i_1+i_2} [X_{j:(i_1+i_2)} \leq b_j] \cap [X_{j:(i_1+i_2)} \leq b_m]\right). \quad (\text{A.3})$$

Denote by  $\bar{M} = \bar{M}(i) := [i] \setminus M$  the complement of  $M = M(i)$ . Furthermore let  $i := i(i_1, i_2) = i_1 + i_2$ . As in [SW09, p. 364 ff.] we begin by partitioning the probability space based on how many of the order statistics fall into the interval  $(b_{m-1}, b_m]$ :

$$Q_{i_1, i_2}(m) = \sum_{k=0}^i \sum_{\substack{M \subset [i] \\ |M|=k}} \mathbb{P} \left( \bigcap_{j=1}^k [X_{j:M} \leq b_j \cap X_{j:M} \leq b_{m-1}] \right. \\ \left. \cap \bigcap_{j=1}^{i-k} [X_{j:\bar{M}} \leq b_{k+j} \cap b_{m-1} < X_{j:\bar{M}} \leq b_m] \right)$$

We then continue (as in [Bla+14, Proposition 1]) by letting  $k := k_1 + k_2$  and conditioning on the number of random variables distributed according to  $\text{Uni}[0, 1]$  and  $F$ , respectively:

$$\dots = \sum_{k_1=0}^{i_1} \sum_{k_2=0}^{i_2} \sum_{\substack{M \subset [i] \\ |M|=k_1+k_2 \\ |M \cap [i_1]|=k_1}} \mathbb{P} \left( \bigcap_{j=1}^k [X_{j:M} \leq b_j \cap X_{j:M} \leq b_{m-1}] \right) \\ \times \mathbb{P} \left( \bigcap_{j=1}^{i-k} [X_{j:\bar{M}} \leq b_{k+j} \cap b_{m-1} < X_{j:\bar{M}} \leq b_m] \right)$$

Using (A.3) we can further simplify this:

$$\dots = \sum_{k_1=0}^{i_1} \sum_{k_2=0}^{i_2} \sum_{\substack{M \subset [i] \\ |M|=k_1+k_2 \\ |M \cap [i_1]|=k_1}} Q_{k_1, k_2}(m-1) \\ \times \mathbb{P} \left( \bigcap_{j=1}^{i-k} [X_{j:\bar{M}} \leq b_{k+j} \cap b_{m-1} < X_{j:\bar{M}} \leq b_m] \right)$$

If  $m-1 \geq k+1$  the intersection event

$$\{X_{1:\bar{M}} \leq b_{k+1} \cap b_{m-1} < X_{1:\bar{M}} \leq b_m\}$$

is empty because the  $b_i$  are an increasing sequence. If  $m-1 \geq k+1$  it is equal to  $\{b_{m-1} < X_{1:\bar{M}} \leq b_m\}$  otherwise. Therefore we can further simplify:

$$\dots = \sum_{\substack{0 \leq k_1 \leq i_1 \\ 0 \leq k_2 \leq i_2 \\ m-1 \leq k_1+k_2}} \sum_{\substack{M \subset [i] \\ |M|=k_1+k_2 \\ |M \cap [i_1]|=k_1}} Q_{k_1, k_2}(m-1) \times \mathbb{P} \left( \bigcap_{j=1}^{i-k} b_{m-1} < X_{j:\bar{M}} \leq b_m \right) \\ = \sum_{\substack{0 \leq k_1 \leq i_1 \\ 0 \leq k_2 \leq i_2 \\ m-1 \leq k_1+k_2}} \binom{i_1}{k_1} \binom{i_2}{k_2} Q_{k_1, k_2}(m-1) \times (b_m - b_{m-1})^{i_1-k_1} \\ \times (F(b_m) - F(b_{m-1}))^{i_2-k_2} \tag{A.4}$$

From the definition it follows immediately that  $Q_{0,0}(0) = 1$  and  $Q_{i_1,i_2}(1) = b_1^{i_1} \cdot F(b_1)^{i_2}$  and thus for  $m > 1$  the quantity  $Q$  can be calculated by (A.4). Furthermore for all  $i_1 \in [n_1], i_2 \in [n_2]$  it holds that

$$\Psi(i_1, i_2) = \mathbb{P} \left( \bigcap_{j=1}^i [X_{j,i} \leq b_j \cap X_{j,i} \leq b_i] \right) = Q_{i_1,i_2}(i_1 + i_2)$$

which completes the proof. Consequently one needs to calculate  $Q_{i_1,i_2}(m)$  for  $\forall (i_1, i_2) \in [n_1] \times [n_2]$  and  $m \leq i_1 + i_2 \leq n$ .  $\square$

**Proof of Lemma 6.** Counting the number of operations in the loops of Algorithm 2 it follows that

$$\begin{aligned} \# \text{Operations} &< \sum_{m_1=0}^{n_1} \left[ 7 + \sum_{m_2=0}^{n_2} \sum_{k_1=0}^{m_1} \left[ 6 + \sum_{k_2=0}^{m_2} 10 \right] \right] = \dots \\ &= \frac{5n_1^2 n_2^2 + 21n_1^2 n_2 + 15n_1 n_2^2 + 63n_1 n_2}{2} \\ &\quad + 8n_1^2 + 5n_2^2 + 31n_1 + 21n_2 + 23 \end{aligned}$$

holds. For the space complexity simply note that, to use the recursions for  $M_{k_1,k_2}^{(m_1,m_2)}$ , we need to keep track of at most  $M_{k_1,k_2}^{(m_1,m_2-1)}$  and  $M_{k_1,k_2}^{(m_1-1,m_2)}$  (which is pessimistic - cf. algorithm 2).

For Steck's recursion first notice that, using exponentiation by squaring, one can calculate  $a^n$  in  $O(\log_2(n))$  multiplications (where  $n \in \mathbb{N}$ , cf. [Knu98, p. 462, Algorithm A]). Thus the first row and last column of  $M^{(m_1,m_2)}$  (cf. equations (2.5) (2.6)) can be calculated in

$$\begin{aligned} &O \left( \sum_{j=1}^{m_2-1} [\log_2(m_1) + \log_2(m_2 - j)] + \sum_{j=1}^{m_1-1} \log_2(m_1 - j) \right) \\ &= O \left( m_2 \log_2(m_1) + \sum_{j=1}^{m_2} \log_2(j) + \sum_{j=1}^{m_1} \log_2(j) \right) \\ &\subset O(m_2 \log_2(m_1) + m_1 \log_2(m_1) + m_2 \log_2(m_2)) \end{aligned}$$

Thus to calculate all the coefficient matrices we need at most

$$\begin{aligned} &O \left( \sum_{m_1=1}^{n_1} \sum_{m_2=1}^{n_2} [m_1 m_2 + m_2 \log_2(m_1) + m_1 \log_2(m_1) + m_2 \log_2(m_2)] \right) \\ &\subset O(n_1^2 n_2^2 + n_1^2 n_2 \log_2(n_1) + n_1 n_2^2 \log_2(n_2)) \end{aligned}$$

arithmetic operations (since we can calculate  $a(m_1, j_1), a(m_2, j_2)$  for  $j_1 \in [m_1], j_2 \in [m_2]$  in  $O(m_1 + m_2)$  using (2.4)). It remains to notice that (2.3) needs at most  $O(m_1 m_2)$  arithmetic operations. For the space complexity simply note that we do not need to keep track of the previous coefficient matrices.

For Noe's recursion first note that, for every  $m \in \mathbb{N}$ , we can calculate  $a^{(m),1}(j_1), a^{(m),2}(j_2)$  for  $j_1 \in [i_1], j_2 \in [i_2]$  in  $O(j_1 + j_2)$ . Furthermore, using (2.4), the binomial coefficients in (2.7) can be calculated in  $O(i_1 + i_2)$ . Thus  $M^{i_1,i_2}$  can be calculated in  $O(i_1 i_2)$ . Thus  $Q_{i_1,i_2}(m)$  (assuming the necessary  $Q_{\cdot, \cdot}(m-1)$  have already been calculated) is  $O(i_1 i_2)$ . Therefore the overall computational complexity is at most

$$O \left( (n_1 + n_2) \cdot \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} i_1 i_2 \right) = O(n_1^2 n_2^2 (n_1 + n_2))$$

For the space complexity simply note, again, that we do not need to keep track of the previous coefficient matrices.  $\square$



**Proof of Lemma 7.** By Noe's recursion (cf. lemma 5) the probability  $\Psi(i_1, i_2)$  can be obtained by evaluating a polynomial of degree  $i_1 + i_2$  with only positive coefficients at  $\mathbf{x} \in \mathbb{R}^{2n}$  (where  $\mathbf{x}$  is given by (2.9)). It is therefore sufficient to show that the statement is true for such polynomials when applied to non-negative arguments. To provide a concise proof we utilise interval arithmetic (cf. [Kea]). Due to linearity (since all coefficients and inputs are non-negative) it is sufficient to show the claim for monomials  $p(x_1, \dots, x_{2n}) := \prod_{i=1}^{2n} x_i^{a_i}$  with  $a_i \in \mathbb{N}$ ,  $\sum_{i=1}^{2n} a_i \leq i_1 + i_2$ . Since  $0 < 1 - \varepsilon < 1$  and  $1 + \varepsilon > 1$  it follows that

$$\tilde{x}_i^{a_i} \in x_i^{a_i} \cdot ((1 - 2\varepsilon)^{a_i}, (1 + 2\varepsilon)^{a_i})$$

which implies

$$p(\tilde{x}_1, \dots, \tilde{x}_{2n}) \in p(x_1, \dots, x_{2n}) \cdot ((1 - 2\varepsilon)^{i_1+i_2}, (1 + 2\varepsilon)^{i_1+i_2})$$

due to [Kea, Equation (4)].  $\square$

**Proof of Lemma 12.** By [RV11, Theorem 3.1] under the unconditional model  $RM(m, \pi_0, F)$  and for a step-up procedure  $SU_t$ :

$$\begin{aligned} & \mathbb{P}(V(SU_t, p) = j, R(SU_t, \mathbf{p}) = k) \\ = & \mathbb{P}(V(SU_t, \mathbf{p}) = j | R(SU_t, \mathbf{p}) = k) \cdot \mathbb{P}(R(SU_t, \mathbf{p}) = k) \\ = & \binom{m}{k} \binom{k}{j} \tilde{\pi}_0^j (1 - \tilde{\pi}_0)^{k-j} G(t_k)^k \tilde{\Psi}_{m-k}(t_m, \dots, t_{k+1}) \end{aligned}$$

where

$$\begin{aligned} \tilde{\Psi}_{m-k}(t_m, \dots, t_{k+1}) &:= \Psi_{m-k,0}^{\text{Uni}[0,1],F}(1 - G(t_m), \dots, 1 - G(t_{k+1})) \\ \tilde{\pi}_0 &:= \frac{\pi_0 t_k}{G(t_k)} \\ G(t) &:= \pi_0 t + (1 - \pi_0)F(t) \end{aligned}$$

and  $\mathbb{P}$  denotes  $\mathbb{P}_{m,\pi_0,F}$ .

Furthermore under the conditional model  $FM(m, m_0, F)$  and for a step-up procedure  $SU_t$  it holds (by [RV11, Section 5.3]) that (where  $\bar{F}(t) := 1 - F(1 - t)$ ):

$$\begin{aligned} & \mathbb{P}(V(SU_t, \mathbf{p}) = j, R(SU_t, \mathbf{p}) = k) \\ = & \binom{m_0}{j} \binom{m - m_0}{k - j} t_k^j (F(t_k))^{k-j} \Psi_{m-k-(m_0-j), m_0-j}^{\text{Uni}[0,1],\bar{F}}(1 - t_m, \dots, 1 - t_{k+1}) \end{aligned}$$

where  $\mathbb{P}$  denotes  $\mathbb{P}_{m,m_0,F}$ .  $\square$

**Algorithm 1** Bolshev Recursion

---

```

1: procedure BOLSHEV( $b$ )
2:    $b \leftarrow 1 - b$ 
3:    $s \in \mathbb{R}^n$ 
4:    $s_1 \leftarrow 0$ 
5:   for  $k = 2, \dots, n$  do
6:      $v \leftarrow 1$ 
7:     for  $j = 1, \dots, k - 1$  do
8:        $v \leftarrow v - s_j$ 
9:        $s_j \leftarrow \frac{s_j \cdot b_j^k}{k - (j-1)}$ 
10:    end for
11:     $s_k \leftarrow k \cdot v \cdot b_k$ 
12:  end for
13:  return  $1 - \sum_{i=1}^n s_i$ 
14: end procedure

```

---

**Algorithm 2** Efficient Generalized Bolshev Recursion

---

```

1: procedure GENERALIZEDBOLSHEV( $n_1 \in \mathbb{N}, n_2 \in \mathbb{N}, v^{(1)} \in (0, 1)^{n_1+n_2}, v^{(2)} \in (0, 1)^{n_1+n_2}$ )
2:    $r \in \mathbb{R}^{(n_1+1) \times (n_2+1)}$ 
3:    $\forall i \in [n_1 + 1], j \in [n_2 + 1] : r_{i,j} \leftarrow 1$ 
4:    $M \leftarrow r$ 
5:    $M^{(0)} \in \mathbb{R}^{n_1+1}$ 
6:    $\forall i \in [n_1 + 1] : M_i^{(0)} \leftarrow 1$ 
7:   for  $m_1 = 0, \dots, n_1$  do
8:     for  $m_2 = 0, \dots, n_2$  do
9:       for  $k_1 = 0, \dots, m_1$  do
10:        for  $k_2 = 0, \dots, m_2$  do
11:          if  $k_1 < m_1 \vee k_2 < m_2$  then
12:             $r_{m_1+1, m_2+1} \leftarrow r_{m_1+1, m_2+1} - M_{k_1+1, k_2+1} \cdot r_{k_1+1, k_2+1}$ 
13:          end if
14:          if  $m_2 < n_2$  then
15:             $M_{k_1+1, k_2+1} \leftarrow M_{k_1+1, k_2+1} \cdot \frac{m_2+1}{m_2+1-k_2} \cdot (1 - v_{k_1+k_2+1}^{(2)})$ 
16:          end if
17:        end for
18:        if  $m_2 < n_2 \wedge k_1 < m_1$  then
19:           $M_{k_1+1, m_2+2} \leftarrow M_{k_1+2, m_2+1} \cdot \frac{k_1+1}{m_1-k_1} \cdot (1 - v_{k_1+m_2+2}^{(1)})$ 
20:        end if
21:      end for
22:    end for
23:    if  $m_1 < n_1$  then
24:      for  $k_1 = 0, \dots, m_1$  do
25:         $M_{k_1+1}^{(0)} \leftarrow M_{k_1+1}^{(0)} \cdot \frac{m_1+1}{m_1+1-k_1} \cdot (1 - v_{k_1+1}^{(1)})$ 
26:         $M_{k_1+1, 1} \leftarrow M_{k_1+1}^{(0)}$ 
27:      end for
28:    end if
29:  end for
30:  return  $r_{n_1+1, n_2+1}$ 
31: end procedure

```

---

# B

## APPENDIX FOR CHAPTER 3

### B.1. CONDITIONAL DENSITY

Let  $X$  be a  $\mathbb{R}^d$  valued random variable with (Lebesgue-) density  $f$ . Then for any (Lebesgue-) measurable  $A \subset \mathbb{R}^d$  it holds, by definition, that

$$\begin{aligned}\int_A f(x|c^T X \geq \ell) dx &= \mathbb{P}(X \in A | c^T X \geq \ell) \\ &= \frac{\mathbb{P}(X \in A \cap c^T X \geq \ell)}{\mathbb{P}(c^T X \geq \ell)} \\ &= \frac{1}{\mathbb{P}(c^T X \geq \ell)} \cdot \int_{A \cap B} f(x) dx,\end{aligned}$$

where  $B := \{x \in \mathbb{R}^d | c^T x \geq \ell\}$ . Hence, it follows that

$$f(x|c^T X \geq \ell) = \frac{1}{\mathbb{P}(c^T X \geq \ell)} \begin{cases} 0 & \text{if } c^T x < \ell, \\ f(x) & \text{otherwise,} \end{cases}$$

i.e. the conditional density is zero on the complement of  $B$  and proportional to  $f$  on  $B$ . Thus finding the maximizer of the conditional density is equivalent to finding a maximizer of the unconditional density over  $B$ .

### B.2. PROOFS

*Proof of Lemma 8.* For the normal distribution it holds that  $g(u) := \exp(-\frac{u}{2})$  and since

$$g'(u) = -\frac{1}{2} \exp(-\frac{u}{2}) < 0$$

the generator is strictly decreasing.

For the Student-t distribution it holds that  $g(u) := \left(1 + \frac{u}{c_p}\right)^{-p}$  for  $c_p \in \mathbb{R}$ ,  $p > \frac{1}{2}$ . Since it holds that

$$g'(u) = -\frac{p\left(1 + \frac{u}{c_p}\right)^{-1-p}}{c_p}$$

and, noting that the integrability condition of definition 2 is only fulfilled for  $c_p > 0$ , it is immediate that the generator is strictly decreasing since  $g' < 0$ .

For the Logistic distribution it holds that  $g(u) := \frac{e^{-u}}{(1+e^{-u})^2}$ . By the quotient rule

$$\begin{aligned} g'(u) &= \frac{-e^{-u}(1+e^{-u})^2 + 2e^{-u}(1+e^{-u})}{(1+e^{-u})^4} \\ &= \frac{e^{-u}(1+e^{-u}) \cdot (-1 - e^{-u} + 2)}{(1+e^{-u})^4} \\ &= \frac{e^{-u}(1+e^{-u}) \cdot (1 - e^{-u})}{(1+e^{-u})^4} \end{aligned}$$

Since  $1 < e^{-u}$  for  $u > 0$  it follows that  $g'(u) < 0$  on  $\mathbb{R}_+$  and therefore  $g$  is strictly decreasing on  $\mathbb{R}_+$ .

For the Exponential distribution it holds that  $g(u) := \exp(-ru^s)$  for  $r, s > 0$ . Furthermore

$$g'(u) = -\exp(-ru^s) r s u^{s-1}$$

and thus for  $u > 0$  it holds that  $g'(u) < 0$  and consequently  $g$  is strictly decreasing on  $\mathbb{R}_+$ .  $\square$

### B.2.1. REVERSE STRESS TESTING IN ELLIPTICAL MODELS

*Proof of Theorem 2.* Define  $a = x - \mu$ . Then (3.2) can be written by

$$\begin{aligned} \arg \max_a \quad & g(a^T \Sigma^{-1} a) \\ \text{subject to} \quad & c^T a \geq \ell - c^T \mu \end{aligned} \tag{B.1}$$

Let

$$\Omega_v = \{a : c^T a = v\} \quad \text{for } v \in [\ell - c^T \mu, \infty).$$

Since  $g$  is a decreasing function, the solution of (B.1) is obtained by the following optimization problem:

$$\min_{v \geq \ell - c^T \mu} \min_{a \in \Omega_v} a^T \Sigma^{-1} a$$

If  $\ell \leq c^T \mu$ , then  $v = 0$  is admissible which implies  $\mathbf{0} = a \in \Omega_0$  is admissible and therefore the minimum of  $a^T \Sigma^{-1} a$  is attained at  $a = \mathbf{0}$  independently of  $v$  and consequently  $x^*(\ell) = \mu$  is the only solution of (3.2). Otherwise, using that finding the minimum of  $a^T \Sigma^{-1} a$  over  $\Omega_v$  is a quadratic optimization problem under a single linear equality constraint we straightforwardly get its solution as

$$a_v = v \frac{\Sigma c}{c^T \Sigma c}$$

or

$$x^*(\ell) = \mu + v \frac{\Sigma c}{c^T \Sigma c} \tag{B.2}$$

and therefore

$$\min_{a \in \Omega_v} a^T \Sigma^{-1} a = \frac{v^2}{c^T \Sigma c}.$$

Hence,

$$\min_{v \geq \ell - c^T \mu} \min_{a \in \Omega_v} a^T \Sigma^{-1} a = \min_{v \geq \ell - c^T \mu} \frac{v^2}{c^T \Sigma c} = \frac{(\ell - c^T \mu)^2}{c^T \Sigma c}$$

for  $v = \ell - c^T \mu$  which together with (B.2) leads to the statement of the theorem.  $\square$

### B.2.2. REVERSE STRESS TESTING IN SKEW-ELLIPTICAL MODELS

**Lemma 13.** Let  $f$  be the density of a skew-elliptical distribution in the sense of (3.4) and in  $\mathbb{R}^d$  for  $d \geq 2$ . Furthermore let  $M$  be a non-singular matrix. If  $X$  is a random vector with density  $f$ , then the density of  $Y := M(X - \mu)$  is given by

$$f(y; \mu, \Sigma, \lambda) = 2 \cdot \tilde{f}(y; \mathbf{0}, M\Sigma M^T) \cdot F(\lambda^T M^{-1}y),$$

that is  $Y$ , is also skew-elliptically distributed with parameters  $\mathbf{0}$ ,  $M\Sigma M^T$ , and  $(M^{-1})^T \lambda$ .

*Proof.* The statement of the lemma follows directly from the application of the change-of-variables formula which is applied first to the density  $f(x; \mu, \Sigma, \lambda)$  and then to  $\tilde{f}(x; \mu, \Sigma)$ .  $\square$

**Lemma 14.** Write  $X = (X_1, X_2^T)^T$ , where  $X_1$  takes values in  $\mathbb{R}$ . Write analogously  $x = (x_1, x_2^T)^T$  for a realization of  $X$ . Assume that the density of  $X$  is given by

$$f(x; \mathbf{0}, \Omega, \lambda) = 2 \cdot \tilde{f}(x; \mathbf{0}, \Omega) \cdot F(x_1)$$

where  $\tilde{f}$  is the density of an elliptical distribution with density generator  $\tilde{g}$  and  $F$  is a cumulative distribution function of a univariate distribution. Then, conditionally on  $X_1$ ,  $X_2$  is elliptically contoured distributed with location parameter  $\omega_{11}^{-1} \cdot x_1 \cdot \Omega_{21}$ , dispersion matrix  $\Omega_{22} - \omega_{11}^{-1} \cdot \Omega_{21} \Omega_{12}$ , and density generator given by

$$\check{g}(\cdot | x_1) \propto \tilde{g}(\omega_{11}^{-1} x_1^2 + \cdot),$$

where  $\check{g}(\cdot | x_1)$  is a decreasing function as soon as  $\tilde{g}$  is decreasing.

*Proof of Lemma 14.* Let  $\Omega$  and  $B = \Omega^{-1}$  be partitioned as

$$\Omega = \begin{pmatrix} \omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \quad (\text{B.3})$$

where  $B_{12} = B_{21}^T$  and  $b_{11} > 0$ . Moreover, using the formula for the inverse of the partitioned matrix we get

$$B_{22} = \left( \Omega_{22} - \frac{\Omega_{21} \Omega_{12}}{\omega_{11}} \right)^{-1}, \quad B_{22}^{-1} B_{21} = -\frac{\Omega_{21}}{\omega_{11}}, \quad b_{11} - B_{12} B_{22}^{-1} B_{21} = \omega_{11}^{-1}.$$

Then, the application of

$$\begin{aligned} x^T \Omega^{-1} x &= x^T B x = b_{11} x_1^2 + 2B_{21}^T x_2 x_1 + x_2^T B_{22} x_2 \\ &= (x_2 + B_{22}^{-1} B_{21} x_1)^T B_{22} (x_2 + B_{22}^{-1} B_{21} x_1) + (b_{11} - B_{12} B_{22}^{-1} B_{21}) x_1^2 \\ &= \omega_{11}^{-1} x_1^2 + (x_2 - \omega_{11}^{-1} \Omega_{21} x_1)^T \left( \Omega_{22} - \frac{\Omega_{21} \Omega_{12}}{\omega_{11}} \right)^{-1} (x_2 - \omega_{11}^{-1} \Omega_{21} x_1) \end{aligned}$$

leads to

$$\tilde{f}(x_1, x_2; \mathbf{0}, \Omega, \lambda) \propto F(x_1) \times \tilde{g} \left( \omega_{11}^{-1} x_1^2 + (x_2 - \omega_{11}^{-1} \Omega_{21} x_1)^T \left( \Omega_{22} - \frac{\Omega_{21} \Omega_{12}}{\omega_{11}} \right)^{-1} (x_2 - \omega_{11}^{-1} \Omega_{21} x_1) \right)$$

from which the conditional distribution of  $X_2$  given  $X_1$  follows as provided in the statement of the lemma.  $\square$

In the following we assume that  $\lambda \neq \mathbf{0}$ , since otherwise the results of Theorem 2 can be applied. Let  $\lambda = (\lambda_1, \lambda_2^T)^T$  with  $\lambda_1 \neq 0$  which can be assumed without loss of generality, since otherwise the components of the vector  $x - \mu$  can be rearranged. We define

$$M = \begin{pmatrix} \lambda_1 & \lambda_2^T \\ \mathbf{0}_{d-1,1} & I_{d-1} \end{pmatrix} \quad (\text{B.4})$$

and applying [Dhr13, Proposition 2.31, p.45] to  $M$  yields

$$M^{-1} = \begin{pmatrix} \lambda_1^{-1} & -\lambda_1^{-1} \cdot \lambda_2^T \\ \mathbf{0}_{d-1,1} & I_{d-1} \end{pmatrix}$$

from which it follows that, for  $c = (c_1, c_2^T)^T$ ,

$$(m_1, m_2^T)^T = (M^T)^{-1} c = \frac{c_1}{\lambda_1} \cdot \begin{pmatrix} 1 \\ -\lambda_2 \end{pmatrix} + \begin{pmatrix} 0 \\ c_2 \end{pmatrix},$$

i.e.  $m_1 = c_1/\lambda_1$  and  $m_2 = -c_1/\lambda_1 \cdot \lambda_2 + c_2$ . Define  $\Omega = M \Sigma M^T$  and consider the representation

$$\Sigma = \begin{pmatrix} \sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

of the dispersion matrix. Furthermore, let

$$\Sigma_1 = \begin{pmatrix} \sigma_{11} \\ \Sigma_{21} \end{pmatrix} \quad \text{and} \quad \Sigma_2 = \begin{pmatrix} \Sigma_{12} \\ \Sigma_{22} \end{pmatrix}.$$

For symmetric  $\Sigma$ , it follows that

$$\Omega = \begin{pmatrix} \omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix},$$

where  $\omega_{11} = \lambda^T \Sigma \lambda$ ,  $\Omega_{22} = \Sigma_{22}$  and  $\Omega_{21}^T = \Omega_{12} = \lambda^T \Sigma_2$ . In preparation of the proof of Theorem 7 we let

$$k = m_1 + \omega_{11}^{-1} m_2^T \Omega_{21} = \frac{c_1}{\lambda_1} \cdot \left[ 1 - \frac{\lambda_2^T \Sigma_2^T \lambda}{\lambda^T \Sigma \lambda} \right] + \frac{c_2^T \Sigma_2^T \lambda}{\lambda^T \Sigma \lambda} = \frac{c^T \Sigma^T \lambda}{\lambda^T \Sigma \lambda}$$

$$A = \Omega_{22} - \omega_{11}^{-1} \Omega_{21} \Omega_{12}.$$

**Theorem 7.** Let  $f$  be the density of a skew-elliptical distribution (in  $\mathbb{R}^d$  for  $d \geq 2$ ) in the sense of Definition 3, with decreasing density generator  $\tilde{g}$ . Let  $\omega_{11}, \Omega_{21}, k, M, m_2$  and  $A$  be defined as previously.

Then, a global maximum of the constrained optimization problem

$$\begin{aligned} & \underset{x}{\text{maximize}} && f(x; \mu, \Sigma, \lambda) \\ & \text{subject to} && c^T x \geq \ell \end{aligned} \quad (\text{B.5})$$

is given by

$$x^*(\ell) = \mu + M^{-1} \begin{pmatrix} z_{1;\max}(\ell) \\ z_{2;\max}(\ell) \end{pmatrix},$$

where

$$z_{1;\max}(\ell) \in \begin{cases} \arg \max_{kz_1 \geq \ell - c^T \mu} F(z_1) \cdot \tilde{g}(\omega_{11}^{-1} z_1^2) & \text{if } c = \eta \lambda, \\ \arg \max_{z_1} F(z_1) \cdot \tilde{g}(\omega_{11}^{-1} z_1^2 + k_1 \cdot (\max\{0, \ell - c^T \mu - kz_1\})^2) & \text{otherwise,} \end{cases} \quad (\text{B.6})$$

$$z_{2;max}(\ell) = \omega_{11}^{-1} \Omega_{21} z_{1;max}(\ell) + k_1 \cdot \max\{0, \ell - c^T \mu - k z_{1;max}(\ell)\} \cdot A m_2,$$

$$\text{and } k_1 = \begin{cases} 0 & \text{if } c = \eta \lambda, \\ (m_2^T A m_2)^{-1} & \text{otherwise.} \end{cases}$$

*Proof.* We consider the following transformation given by

$$z = M(x - \mu),$$

where  $M$  is given in (B.4). Then, from Lemma 13, the density of  $z = (z_1, z_2)^T$  is given by

$$\begin{aligned} f(z; \mathbf{0}, \Omega, \lambda) &= 2 \cdot \tilde{f}(z; \mathbf{0}, \Omega) \cdot F(z_1) \\ &\propto F(z_1) \times \tilde{g}\left(\omega_{11}^{-1} z_1^2 + (z_2 - \omega_{11}^{-1} \Omega_{21} z_1)^T A^{-1} (z_2 - \omega_{11}^{-1} \Omega_{21} z_1)\right), \end{aligned}$$

where  $\Omega$  is partitioned as in (B.3).

Furthermore, the condition  $c^T x \geq \ell$  is equivalent to  $((M^T)^{-1} c)^T z \geq \ell - c^T \mu$  or  $m_2^T z_2 \geq \ell - c^T \mu - m_1 z_1$ . It follows that

$$\max_{x: c^T x \geq \ell} f(x; \mu, \Sigma, \lambda)$$

is equivalent to

$$\max_{z_1} \max_{z_2: m_2^T z_2 \geq \ell - c^T \mu - m_1 z_1} F(z_1) \cdot \tilde{g}\left(\omega_{11}^{-1} z_1^2 + (z_2 - \omega_{11}^{-1} \Omega_{21} z_1)^T A^{-1} (z_2 - \omega_{11}^{-1} \Omega_{21} z_1)\right),$$

and by Lemma 14 the objective function is, for every fixed  $z_1$ , the density of an elliptical distribution with location parameter  $\tilde{\mu} = \omega_{11}^{-1} \Omega_{21} z_1$  and dispersion matrix  $\tilde{\Sigma} = \Omega_{22} - \omega_{11}^{-1} \Omega_{21} \Omega_{12}$ . If  $c = \eta \lambda$  then  $m_2 = 0$  and the optimizer is  $z_2 = \omega_{11}^{-1} \Omega_{21} z_1$ . Otherwise let  $\tilde{c} = m_2^T$  and  $\tilde{\ell} = \ell - c^T \mu - m_1 z_1$ . Then, by Theorem 2, it follows that, for  $v = z_2 - \omega_{11}^{-1} \Omega_{21} z_1$ ,

$$\begin{aligned} z_{1;max} &= \arg \max_{z_1} \max_{z_2: m_2^T z_2 \geq \tilde{\ell}} F(z_1) \cdot \tilde{g}\left(\omega_{11}^{-1} z_1^2 + v^T A^{-1} v\right) \\ &= \arg \max_{z_1} F(z_1) \cdot \tilde{g}\left(\omega_{11}^{-1} z_1^2 + \frac{(\max\{0, \ell - c^T \mu - k z_1\})^2}{m_2^T A m_2}\right) \end{aligned}$$

where  $k = m_1 + \omega_{11}^{-1} m_2^T \Omega_{21}$  which is attained at

$$z_{2;max} = \omega_{11}^{-1} \Omega_{21} z_{1;max} + \max\{0, \ell - c^T \mu - k z_{1;max}\} \cdot \frac{A m_2}{m_2^T A m_2}.$$

Finally, using the inverse transformation we get  $x_{max}$  as stated in the theorem.  $\square$

To derive a more explicit representation of  $x^*(\ell)$  one can verify the following three identities:

$$\begin{aligned} M^{-1} \begin{pmatrix} 1 \\ \omega_{11}^{-1} \Omega_{21} \end{pmatrix} &= \frac{\Sigma \lambda}{\lambda^T \Sigma \lambda} \\ M^{-1} \begin{pmatrix} 0 \\ A m_2 \end{pmatrix} &= \Sigma c - \frac{\lambda^T \Sigma c}{\lambda^T \Sigma \lambda} \Sigma \lambda \\ m_2^T A m_2 &= c^T \left( \Sigma c - \frac{\lambda^T \Sigma c}{\lambda^T \Sigma \lambda} \Sigma \lambda \right) \end{aligned}$$

These imply the following corollary.

**Corollary 4.** Under the assumptions of Theorem 7 it holds that

$$x^*(\ell) = \mu + \begin{cases} \eta^{-1} \cdot z_{1;\max}(\ell) \cdot \frac{\Sigma c}{c^T \Sigma c} & \text{if } c = \eta \lambda, \\ z_{1;\max}(\ell) \cdot \frac{\Sigma \lambda}{\lambda^T \Sigma \lambda} + \max\{0, \ell - c^T \mu - k \cdot z_{1;\max}(\ell)\} \frac{v}{c^T v} & \text{otherwise,} \end{cases}$$

where  $v = \Sigma c - (\lambda^T \Sigma c)/(\lambda^T \Sigma \lambda) \cdot \Sigma \lambda$  and  $k = (\lambda^T \Sigma c)/(\lambda^T \Sigma \lambda)$ .

### B.2.3. REVERSE STRESS TESTING IN SKEW-NORMAL MODELS

In light of Theorem 7 it is interesting to consider, in which models a unique solution  $z_{1;\max}(\ell)$  of (B.6) is guaranteed to exist. In this section, we will show this to be true for the skew-normal distribution, i.e. for the case where  $f$  is a normal density and  $F$  is the cumulative distribution function of the univariate normal distribution, which we will denote by  $\Phi$ .

The following lemma is the first step.

**Lemma 15.** Consider the optimization problem

$$\arg \max_q g(\eta_1 q^2 + \eta_2 q + \eta_3) \cdot \Phi(q),$$

where  $\eta_1, \eta_2, \eta_3 \in \mathbb{R}$  with  $\eta_1 > 0$  and where  $g(u) = \exp(-u/2)$ . The objective function is unimodal and therefore the optimization problem always has a unique solution  $q^*$  which satisfies  $\phi(-\eta_2/(2\eta_1))/\Phi(-\eta_2/(2\eta_1)) \geq q^* + \eta_2/(2\eta_1) \geq 0$ .

*Proof.* First note that the objective function  $\varphi(q) := g(\eta_1 q^2 + \eta_2 q + \eta_3) \cdot \Phi(q)$  looks very similar to the density of a univariate skew-normal distribution; cf. [Azz13, Equation (2.1)]. It is indeed possible to generalize [Azz13, Proposition 2.6] to functions like  $\varphi$ : Showing that the second derivative of  $\log \varphi(q)$  is still strictly negative will imply log-concavity which in turn will imply that  $\varphi$  has a unique mode which is the desired result; cf. [Azz13, Proposition 2.6]. Let  $\zeta_1, \zeta_2$  be given as in [Azz13, Equation (2.20)]. Then it holds that

$$\begin{aligned} \frac{\partial^2}{\partial q^2} \log \varphi(q) &= \frac{\partial^2}{\partial q^2} \left[ -\frac{1}{2} (\eta_1 q^2 + \eta_2 q + \eta_3) + \log(\Phi(q)) \right] \\ &= \frac{\partial}{\partial q} \left[ -\eta_1 q - \frac{\eta_2}{2} \right] + \zeta_2(q) \\ &= -\eta_1 - \zeta_1(q) [q + \zeta_1(q)] \end{aligned}$$

The first summand is negative by the assumption  $\eta_1 > 0$  and by [Azz13, Equation (2.21)] the second summand is negative, too. Thus  $\frac{\partial^2}{\partial q^2} \log \varphi(q) < 0$ . Furthermore from the first order conditions it follows that

$$\begin{aligned} 0 &= \frac{\partial}{\partial q} \log \varphi(q^*) = -\eta_1 q^* - \frac{\eta_2}{2} + \frac{\phi(q^*)}{\Phi(q^*)} \\ &\geq -\eta_1 q^* - \frac{\eta_2}{2} \end{aligned}$$

from which the lower bound follows directly. Furthermore this implies (since  $x \mapsto \phi(x)/\Phi(x)$  is strictly decreasing)

$$0 = -\eta_1 q^* - \frac{\eta_2}{2} + \frac{\phi(q^*)}{\Phi(q^*)} \leq -\eta_1 q^* - \frac{\eta_2}{2} + \frac{\phi\left(-\frac{\eta_2}{2\eta_1}\right)}{\Phi\left(-\frac{\eta_2}{2\eta_1}\right)}$$

which implies  $q^* \leq -\frac{\eta_2}{2\eta_1} + \frac{\phi\left(-\frac{\eta_2}{2\eta_1}\right)}{\Phi\left(-\frac{\eta_2}{2\eta_1}\right)}$ . □



An immediate consequence of the preceding lemma is the following corollary:

**Corollary 5.** *In the skew-normal case there are at most two solutions to (B.5).*

The next step is to characterize the solution for large losses  $\ell$ .

**Corollary 6.** *Let  $f$  be a skew-normal density. Then there exists  $L$  such that, for all  $\ell \geq L$  and  $c, \lambda$  not collinear, the solution  $z_{1;\max}(\ell)$  of (B.6) is unique.*

*Proof.* Since  $\tilde{g}$  is monotonically decreasing and  $k_1 > 0$  it is sufficient to show that there exists  $L \in \mathbb{R}$  such that, for all  $\ell \geq L$ , the solution  $q^*$  of

$$\begin{aligned} q^*(\ell) &= \arg \max_q F(z_1) \cdot \tilde{g}(\omega_{11}^{-1} q^2 + k_1 \cdot (\ell - c^T \mu - k q)^2) \\ &= \arg \max_q F(z_1) \cdot \tilde{g}((\omega_{11}^{-1} + k_1 k^2) q^2 - 2k_1 k (\ell - c^T \mu) q + (\ell - c^T \mu)^2) \end{aligned}$$

(which is unique by Lemma 15) satisfies  $\ell - c^T \mu - k q^*(\ell) \geq 0$ . Consider the case  $k > 0$  (the case  $k < 0$  is analogous). Then we need to show that  $q^*(\ell) \leq (\ell - c^T \mu)/k$  which is implied (cf. Lemma 15) by  $\phi(-\frac{\eta_2}{2\eta_1})/\Phi(-\frac{\eta_2}{2\eta_1}) - \frac{\eta_2}{2\eta_1} \leq (\ell - c^T \mu)/k$  where  $-\frac{\eta_2}{2\eta_1} = \frac{k_1 k (\ell - c^T \mu)}{\omega_{1,1}^{-1} + k_1 k^2} \xrightarrow{\ell \rightarrow \infty} \infty$ . Since  $\phi(x)/\Phi(x) \xrightarrow{x \rightarrow \infty} 0$  it is sufficient to show that  $\frac{k_1 k}{\omega_{1,1}^{-1} + k_1 k^2} < \frac{1}{k}$ . But this follows immediately from the fact that  $\omega_{11} > 0$  and  $k_1 > 0$ :

$$\frac{k_1 k}{\omega_{1,1}^{-1} + k_1 k^2} = \frac{1}{k} \cdot \frac{k_1 k^2}{\omega_{1,1}^{-1} + k_1 k^2} < \frac{1}{k} \cdot \frac{k_1 k^2}{k_1 k^2} = \frac{1}{k}$$

□

**Theorem 8** ([Cap12, Proposition 7]). *Let  $f$  be the density of a skew-normal distribution. Then the unique mode  $M_0$  of  $f$  is given by  $M_0 := \mu + (k^* / \sqrt{\lambda^T \Sigma \lambda}) \cdot \Sigma \lambda$  where  $k^*$  is the unique solution of*

$$0 = z \cdot \Phi(\sqrt{\lambda^T \Sigma \lambda} \cdot z) - \sqrt{\lambda^T \Sigma \lambda} \cdot \phi(\sqrt{\lambda^T \Sigma \lambda} \cdot z). \quad (\text{B.7})$$

Due to Theorem 8 the previous results can be strengthened if one notices that, if,

$$x^*(\ell) = \mu + b_1 \Sigma c + b_2 \Sigma \lambda$$

we can eliminate the constraint  $c^T x^*(\ell) = \ell$  by letting  $b_1 := (\ell - c^T \mu - b_2 \cdot c^T \Sigma \lambda) / (c^T \Sigma c)$ . Therefore, by taking the derivative of the objective (as a function of  $b_2$ ), the following corollary to the previous results can be easily proven:

**Corollary 7.** *Let  $f$  be the density of a skew-normal distribution. Then, the global maximum of the constrained optimization problem*

$$\begin{aligned} &\underset{x}{\text{maximize}} && f(x; \mu, \Sigma, \lambda) \\ &\text{subject to} && c^T x \geq \ell \end{aligned} \quad (\text{B.8})$$

is given by

$$x^*(\ell) = \mu + k_1 \cdot \Sigma \lambda + k_2 \cdot \Sigma c,$$

where  $k_1, k_2 \in \mathbb{R}$  depend on the population parameters. If the mode of  $f$  is feasible then  $k_1 = k^* / \sqrt{\lambda^T \Sigma \lambda}$  (where  $k^*$  is the unique solution of (B.7)) and  $k_2 = 0$ . If it is infeasible and  $\lambda^T \Sigma c = 0$  then

$k_1 = 0$  and  $k_2 = (\ell - c^T \mu) / (c^T \Sigma c)$ . Otherwise  $k_1 = b^*$  and  $k_2 = (\ell - c^T \mu - b^* \cdot c^T \Sigma \lambda) / (c^T \Sigma c)$  where  $b^* = 0$  if there exists  $\eta \in \mathbb{R}$  such that  $c = \eta \lambda$ . Otherwise  $b^*$  is the unique solution of

$$0 = \frac{\phi(a_1 \cdot b + a_3)}{\Phi(a_1 \cdot b + a_3)} - b, \quad (\text{B.9})$$

where  $a_1 = \lambda^T \Sigma \lambda - (c^T \Sigma \lambda)^2 / (c^T \Sigma c)$  and  $a_3 = (\ell - c^T \mu) \cdot (\lambda^T \Sigma c) / (c^T \Sigma c)$ . The solution  $b^*$  of (B.9) satisfies  $b^* \in [0, \phi(a_3) / \Phi(a_3)]$ .

The derivative of  $q \mapsto \phi(q) / \Phi(q)$  is given in [Azz13, Equation (2.20)]; see the quantity  $\zeta_2$  there. This can be e.g. utilized to apply Newton's method to (B.9).

#### B.2.4. REVERSE STRESS TESTING IN SKEW-T MODELS

It is possible to define a skew-t distribution by choosing  $\tilde{f}$  in Definition 3 as a multivariate Student's t-density. However, through conditioning, one obtains the more popular formulation (cf. [Azz13, Section 6])

$$f(z; \mu, \Sigma, \nu, \lambda) := 2 \cdot t_d(z - \mu; \Sigma, \nu) \cdot T_1 \left( \lambda^T (z - \mu) \sqrt{\frac{\nu + d}{\nu + Q(z - \mu; \Sigma)}}; \nu + d \right) \quad (\text{B.10})$$

of a  $d$ -dimensional skew-t distribution with  $\nu$  degrees of freedom and where  $t_d$  denotes a  $d$ -dimensional Student's t density,  $T_1$  is the distribution function of the univariate Student's t distribution and  $Q(z; \Sigma) := z^T \Sigma^{-1} z$ . For notational convenience we have chosen to absorb the scale vector into the dispersion matrix  $\Sigma$  and the skewness parameter  $\lambda$  (cf. [Azz13, Equation (6.23)]). Since (B.10) is not (in the sense of Definition 3) a skew-elliptical density it is not possible to apply Theorem 7. We therefore give a separate proof of a similar result which is based on specific properties of the skew-t distribution. For the proof of the result we will need the following lemma:

**Lemma 16.** For  $\nu > 1$  it holds that

$$0 < h(z) := \nu \cdot T_1(z; \nu) + z \cdot t_1(z; \nu)$$

for all  $z \in \mathbb{R}$  and  $h$  is a strictly increasing function.

*Proof of Lemma 16.* The derivative

$$\begin{aligned} h'(z) &= \nu \cdot t_1(z; \nu) + t_1(z; \nu) + z \cdot t_1'(z; \nu) \\ &= (\nu + 1) \cdot t_1(z; \nu) + z \cdot t_1'(z; \nu) \end{aligned}$$

is greater than zero if and only if

$$\nu + 1 > -z \frac{t_1'(z; \nu)}{t_1(z; \nu)} = (\nu + 1) \cdot \frac{z^2}{\nu + z^2}$$

which holds since  $\frac{z^2}{\nu + z^2} < 1$  and therefore  $h$  is strictly increasing.

For  $z > 0$  it is obvious that  $h(z) > 0$ . For  $z \leq 0$  notice that  $\lim_{z \rightarrow -\infty} T_1(z; \nu) = 0$  and by applying L'Hospital's rule it follows that

$$\lim_{z \rightarrow -\infty} z \cdot t_1(z; \nu) = \lim_{z \rightarrow -\infty} z \left( 1 + \frac{z^2}{\nu} \right)^{-\frac{\nu+1}{2}} = 0$$

since  $\frac{\nu+1}{2} > 1$ . Thus  $\lim_{z \rightarrow -\infty} h(z) = 0$  and because  $h(0) = \frac{\nu}{2} > 0$  and we have proven that  $h$  is strictly increasing it follows that  $h(z) > 0$ .  $\square$

Since the unique mode is the solution of the reverse stress testing problem if it is feasible, we will also make use of the following result of [Cap12]:

**Theorem 9** ([Cap12, Proposition 8]). *Let  $f$  be the density of a skew-t distribution in the sense of Equation (B.10). Then the unique mode  $M_0$  of  $f$  is given by  $M_0 := \mu + (k^* / \sqrt{\lambda^T \Sigma \lambda}) \cdot \Sigma \lambda$ , where  $k^*$  is the unique root of*

$$k \mapsto k(v+d)^{\frac{1}{2}} T_1(w(k); v+d) - t_1(w(k); v+d) v \sqrt{\lambda^T \Sigma \lambda} (v+k^2)^{-\frac{1}{2}} \quad (\text{B.11})$$

and  $w(k) = \sqrt{\lambda^T \Sigma \lambda} \cdot k \cdot ((v+d)/(v+k^2))^{\frac{1}{2}}$ .

**Theorem 10.** *Let  $f$  be the density of a skew-t distribution (in  $\mathbb{R}^d$  for  $d \geq 2$ ) in the sense of Equation (B.10).*

*Then, the global maximum of the constrained optimization problem*

$$\begin{aligned} & \underset{x}{\text{maximize}} && f(x; \mu, \Sigma, v, \lambda) \\ & \text{subject to} && c^T x \geq \ell \end{aligned} \quad (\text{B.12})$$

*is given by*

$$x^*(\ell) = \mu + k_1 \cdot \Sigma \lambda + k_2 \cdot \Sigma c,$$

*where  $k_1, k_2 \in \mathbb{R}$  are constants that depend on the population parameters. If the mode of  $f$  is feasible then  $k_1 = k^* / \sqrt{\lambda^T \Sigma \lambda}$  (where  $k^*$  is the unique solution of (B.11)) and  $k_2 = 0$ . If it is infeasible and  $\lambda^T \Sigma c = 0$ , then  $k_1 = 0$  and  $k_2 = (\ell - c^T \mu) / (c^T \Sigma c)$ . Otherwise  $k_1 = b^*$  and  $k_2 = (\ell - c^T \mu - b^* \cdot c^T \Sigma \lambda) / (c^T \Sigma c)$ , where  $b^* = 0$  if there exists  $\eta \in \mathbb{R}$  such that  $c = \eta \lambda$ . Otherwise  $b^*$  is the unique solution of*

$$b \cdot \sqrt{v+d} \cdot T_1(w_1(b); v+d) = \frac{a_2 + v - a_3 b}{(v + a_1 \cdot b^2 + a_2)^{1/2}} \cdot t_1(w_1(b); v+d), \quad (\text{B.13})$$

*where  $a_1 = \lambda^T \Sigma \lambda - (c^T \Sigma \lambda)^2 / (c^T \Sigma c)$ ,  $a_2 = (\ell - c^T \mu)^2 / (c^T \Sigma c)$ ,  $a_3 = (\ell - c^T \mu) \cdot (\lambda^T \Sigma c) / (c^T \Sigma c)$ , and  $w_1(b) = (a_1 \cdot b + a_3) \cdot ((v+d)/(v + a_1 \cdot b^2 + a_2))^{1/2}$ . The solution  $b^*$  of (B.13) satisfies*

$$b^* \in \begin{cases} \left[0, \frac{a_2 + v}{a_3}\right] & \text{if } a_3 > 0, \\ \left[0, \frac{1}{2} \cdot \sqrt{\frac{a_3^2}{a_1} + a_2 + v} \cdot \sqrt{1 + \frac{1}{(v+d)^2} + 2 \cdot (v+d)}\right] & \text{otherwise.} \end{cases}$$

**Proof of Theorem 10.** By [Cap12, Proposition 8] there exists a unique mode  $M_0$  of  $f$ . If  $c^T M_0 \geq \ell$ , i.e.  $M_0$  is feasible, then by definition  $x^*(\ell) = M_0$ . If  $M_0$  is not feasible, then for any feasible point  $x \in \mathbb{R}^d$  the function  $g(t) := f(M_0 + (x - M_0) \cdot t; \mu, \Sigma, v, \lambda)$  is a strictly decreasing function on  $[0, 1]$  because the level sets of  $f$  are convex (cf. [Azz13, p. 190]). Furthermore  $x$  and  $M_0$  are on different sides of the hyperplane  $H := \{x \in \mathbb{R}^d \mid c^T x = \ell\}$  and therefore there exists  $t_0 \in (0, 1]$  such that  $x_0 := M_0 + (x - M_0) \cdot t_0 \in H$ . Thus  $f(x_0; \mu, \Sigma, v, \lambda) \geq f(x; \mu, \Sigma, v, \lambda)$  (with equality only if  $x_0 \in H$ ) and we can conclude that a solution to (B.12) must be an element of  $H$  if  $M_0$  is infeasible.

To simplify the remaining proof we utilize the canonical form (cf. [Cap12, Section 2]). To this end let  $C^T C = \Sigma$  and  $P = \begin{pmatrix} P_1 & P_2 \end{pmatrix}$  an orthogonal Matrix with  $P_1 = \frac{C \lambda}{\sqrt{\lambda^T \Sigma \lambda}}$ . Applying the transformation  $x \mapsto (C^{-1} P)^T (x - \mu)$  transforms the density into one that is proportional to

$$\tilde{f}(z; v, \alpha) = \left[1 + \frac{\tilde{Q}(z)}{v}\right]^{-\frac{(v+d)}{2}} \times T_1\left(\alpha \cdot z_1 \cdot \left(\frac{v+d}{v + \tilde{Q}(z)}\right)^{1/2}; v+d\right)$$

where  $\tilde{Q}(z) = z^T z$  and  $\alpha = \sqrt{\lambda^T \Sigma \lambda} > 0$ . The condition is transformed into  $c^T (\mu + C^T P) z = \ell$  which is equivalent to  $\tilde{c}^T z = \tilde{\ell}$  where  $\tilde{c} := P^T C c$  and  $\tilde{\ell} = \ell - c^T \mu$ .

If  $\tilde{c}_1 = 0$  then for  $z$  feasible the vector  $(-z_1 \quad z_2)$  is feasible, too. Thus the maximizer will have  $z_1 \geq 0$  and since, for any fixed  $z_1 \geq 0$ ,  $z_2 \mapsto \tilde{f}((z_1 \quad z_2); \nu, \alpha)$  is proportional to an elliptical density with strictly decreasing generator, location parameter 0 and dispersion matrix  $I_d = \text{diag}(1, \dots, 1)$  we can apply Theorem 2 to obtain  $x^*(\ell) = \mu + C^T P \frac{\tilde{\ell}}{\tilde{c}^T \tilde{c}} \tilde{c} = \mu + \frac{\ell - c^T \mu}{c^T \Sigma c} \Sigma c$ .

It remains to analyze the more interesting case  $\tilde{c}_1 \neq 0$ . In this case we can eliminate the constraint by setting  $z_1(z_2) := \frac{\tilde{\ell} - \tilde{c}_2^T z_2}{\tilde{c}_1}$ . Thus we now need to, unconditionally, maximize

$$g(y; \nu, \alpha) = \left[ 1 + \frac{\tilde{Q}(y)}{\nu} \right]^{-\frac{(\nu+d)}{2}} \times T_1 \left( \alpha \cdot z_1(y) \cdot \left( \frac{\nu+d}{\nu + \tilde{Q}(y)} \right)^{1/2}; \nu + d \right)$$

over  $y \in \mathbb{R}^{d-1}$  and where we have redefined  $\tilde{Q}(y) := y^T y + z_1(y)^2$ . The partial derivatives of  $g$  are given by

$$\begin{aligned} \frac{\partial g}{\partial y_i} &= \frac{-(\nu+d)}{2\nu} \times \frac{\partial \tilde{Q}(y)}{\partial y_i} \times \left[ 1 + \frac{\tilde{Q}(y)}{\nu} \right]^{-\frac{(\nu+d)}{2}-1} \times T_1(\alpha z_1(y) \cdot w(y); \nu + d) \\ &\quad + \left[ 1 + \frac{\tilde{Q}(y)}{\nu} \right]^{-\frac{(\nu+d)}{2}} \times t_1(\alpha z_1(y) \cdot w(y); \nu + d) \times \left[ \alpha \frac{\partial z_1(y)}{\partial y_i} w(y) - \alpha z_1(y) \frac{\partial \tilde{Q}(y)}{\partial y_i} \frac{w(y)}{2(\nu + \tilde{Q}(y))} \right] \end{aligned}$$

where  $w(y) := \left( \frac{\nu+d}{\nu + \tilde{Q}(y)} \right)^{1/2}$ ,  $\frac{\partial \tilde{Q}(y)}{\partial y_i} = 2y_i - \frac{2z_1(y)\tilde{c}_{i+1}}{\tilde{c}_1}$  and  $\frac{\partial z_1(y)}{\partial y_i} = -\frac{\tilde{c}_{i+1}}{\tilde{c}_1}$ . The derivatives are zero iff

$$0 = -\frac{\partial \tilde{Q}(y)}{\partial y_i} h(\alpha z_1(y) \cdot w(y); \nu + d) + 2g(y) \frac{\partial z_1(y)}{\partial y_i} [\nu + \tilde{Q}(y)]$$

which is equivalent to

$$y_i \cdot h(\alpha z_1(y) \cdot w(y); \nu + d) = \frac{\tilde{c}_{i+1}}{\tilde{c}_1} \times \left[ z_1(y) \times h(\alpha z_1(y) \cdot w(y); \nu + d) - g(y) [\nu + \tilde{Q}(y)] \right]$$

where  $h$  is as in Lemma 16 and

$$g(y) := \alpha \cdot w(y) \cdot t_1(\alpha z_1(y) \cdot w(y); \nu + d)$$

does not depend on  $i$ . Thus the necessary condition for optimality of  $y^*$  implies, since  $h > 0$  by Lemma 16, that there exists  $k^* \in \mathbb{R}$  such that  $y^* = k^* \cdot \tilde{c}_2$  where we partition  $\tilde{c}$  as  $\tilde{c} = (\tilde{c}_1 \quad \tilde{c}_2)$  with  $\tilde{c}_1 \in \mathbb{R}$ . We conclude that there exists  $k^* \in \mathbb{R}$  such that

$$\begin{aligned} x^*(\ell) &= \mu + C^T P \left[ k^* \cdot \tilde{c} + \begin{pmatrix} \frac{\ell - c^T \mu - k^* \cdot (c^T \Sigma c - \tilde{c}_1^2)}{\tilde{c}_1} - k^* \cdot \tilde{c}_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right] \\ &= \mu + k^* \cdot \Sigma c + \Sigma \lambda \cdot \left[ v_1(k^*) - \frac{k^* \cdot \lambda^T \Sigma c}{\alpha^2} \right] \end{aligned}$$

where  $v_1(k) := \frac{\ell - c^T \mu - k \cdot (c^T \Sigma c - \alpha^{-2} (\lambda^T \Sigma c)^2)}{\lambda^T \Sigma c}$  and we utilized the identities  $\tilde{c}_1 = \alpha^{-1} \lambda^T \Sigma c$  and  $\tilde{c}_2 = c^T \Sigma c - (\tilde{c}_1)^2$ . Equivalently there exist  $b_1, b_2 \in \mathbb{R}$  such that

$$x^*(\ell) = \mu + b_1 \Sigma c + b_2 \Sigma \lambda$$

and we can eliminate the constraint by letting  $b_1 := \frac{\ell - c^T \mu - b_2 \cdot c^T \Sigma \lambda}{c^T \Sigma c}$  which implies that  $x^*(\ell)$  is of the form  $x^*(\ell, b) = \mu + \frac{\ell - c^T \mu}{c^T \Sigma c} \Sigma c + b \cdot \left( \Sigma \lambda - \frac{c^T \Sigma \lambda}{c^T \Sigma c} \Sigma c \right)$ . Plugging  $x^*(\ell, b)$  into the original problem it follows that  $b$  is a maximizer of

$$\tilde{g}(b) = \left[ 1 + \frac{a_1 \cdot b^2 + a_2}{\nu} \right]^{\frac{-(\nu+d)}{2}} \times T_1 \left( (a_1 \cdot b + a_3) \cdot \left( \frac{\nu+d}{\nu + a_1 \cdot b^2 + a_2} \right)^{1/2}; \nu+d \right)$$

where  $a_1 = \lambda^T \Sigma \lambda - \frac{(c^T \Sigma \lambda)^2}{c^T \Sigma c}$ ,  $a_2 = \frac{(\ell - c^T \mu)^2}{c^T \Sigma c}$  and  $a_3 = (\ell - c^T \mu) \cdot \frac{\lambda^T \Sigma c}{c^T \Sigma c}$ . Since the skew-t density is strictly quasi-concave (cf. [AR12, p. 874]) it follows from standard results from quasi-convex programming (cf. [Kar67, Theorem 2.6]) that any local maximum on the convex set  $H \cap \{\mu + a \Sigma c + b \Sigma \lambda | a, b \in \mathbb{R}\}$  is a global maximum. It remains to show the existence of local maximum. From the Cauchy-Schwarz inequality it follows that  $a_1 \geq 0$  since  $\Sigma$  is positive definite and therefore  $x \mapsto x^T \Sigma x$  is an inner product. A quick calculation verifies that the derivative of  $\tilde{g}$  is zero if and only if

$$b \cdot \sqrt{\nu+d} \cdot T_1(w_1(b); \nu+d) = \frac{a_2 + \nu - a_3 b}{(\nu + a_1 \cdot b^2 + a_2)^{1/2}} \cdot t_1(w_1(b); \nu+d)$$

where  $w_1(b) = (a_1 \cdot b + a_3) \cdot \left( \frac{\nu+d}{\nu + a_1 \cdot b^2 + a_2} \right)^{1/2}$ . This equation has at least one solution: For  $b = 0$  the left hand side is equal to zero whereas the right hand side is strictly bigger than zero. Furthermore the left hand side is unbounded as  $b \rightarrow \infty$  whereas the right hand side converges to  $-\frac{a_3}{\sqrt{a_1}} \cdot t_1(\sqrt{a_1} \cdot (\nu+d); \nu+d)$ . Since both sides are continuous functions they must intersect at least once by the intermediate value theorem. For  $a_3 \geq 0$  and  $a_1 > 0$  it furthermore follows that the solution must be an element of  $[0, \frac{a_2 + \nu}{a_3}]$ . For  $a_3 \leq 0$  it follows from

$$\max_{b \geq 0} \frac{a_2 + \nu - a_3 b}{(\nu + a_1 \cdot b^2 + a_2)^{1/2}} = \sqrt{\frac{a_3^2}{a_1} + a_2 + \nu}$$

and Lemma 17 that the solution must be an element of

$$\left[ 0, \frac{1}{2} \cdot \frac{\sqrt{\frac{a_3^2}{a_1} + a_2 + \nu}}{\sqrt{\nu+d}} \cdot \sqrt{\nu+d + \frac{1}{\nu+d} + 2} \right] = \left[ 0, \frac{1}{2} \cdot \sqrt{\frac{a_3^2}{a_1} + a_2 + \nu} \cdot \sqrt{1 + \frac{1}{(\nu+d)^2} + 2 \cdot (\nu+d)} \right].$$

For  $a_1 = 0$  we can simply choose  $b = 0$ . □

To justify the upper bound in the previous theorem we need the following Lemma:

**Lemma 17.** For  $\nu > 0$  denote by  $T_1(\cdot; \nu)$  and  $t_1(\cdot; \nu)$  the cumulative distribution function and the probability density function of a univariate Student's t-distribution with  $\nu$  degrees of freedom. Then for all  $x \in \mathbb{R}$  it holds that  $\frac{t_1(x; \nu)}{T_1(x; \nu)} \leq \frac{1}{2} \sqrt{\nu + \frac{1}{\nu} + 2}$ .

*Proof.* The first order optimality condition is

$$0 = \frac{\partial}{\partial x} \frac{t_1(x; \nu)}{T_1(x; \nu)} = \frac{(T_1(x; \nu) \cdot \frac{\partial}{\partial x} t_1(x; \nu)) - (t_1(x; \nu))^2}{(T_1(x; \nu))^2}$$

which is equivalent to

$$\frac{t_1(x; \nu)}{T_1(x; \nu)} = \frac{\frac{\partial}{\partial x} t_1(x; \nu)}{t_1(x; \nu)} = -\frac{x \cdot (1 + \nu)}{\nu + x^2}$$

and the claim follows from

$$\max_x -\frac{x \cdot (1 + \nu)}{\nu + x^2} = \frac{1}{2} \sqrt{\nu + \frac{1}{\nu} + 2}.$$

B

□

### B.2.5. CONFIDENCE REGION FOR ELLIPTICAL MODELS

To derive an empirical likelihood confidence region (cf. [Owe01]) for the scenario obtained under the setting of Theorem 2, which is given by

$$x^*(\ell) = \mu + \frac{\ell - c^T \mu}{c^T \Sigma c} \Sigma c,$$

we need to estimate the  $2d$  population parameters in  $\mu$  and  $\Sigma c$ . To apply the "smooth function of means" approach (cf. [Owe01, Section 3.2]) we need to write these as smooth functions of expectations of suitable random variables. Denote by  $X$  a random variable distributed according to an elliptical density with parameters  $\mu, \Sigma$  and let  $m := (\mu^T, (\mu^{(1)})^T)$  where  $\mu^{(1)} := \mathbb{E}[(c^T X) \cdot X]$  and  $\mu^{(2)} := \mathbb{E}[c^T X] = c^T \mu$ . Then it holds that

$$x^*(\ell) = x^*(\ell, m) = \mu + \frac{\ell - c^T \mu}{c^T (\mu^{(1)} - \mu^{(2)} \cdot \mu)} \times (\mu^{(1)} - \mu^{(2)} \cdot \mu),$$

and it follows that the gradients of  $x^*$  with respect to the components of  $m$  are given by

$$\begin{aligned} \frac{\partial x^*(\ell, m)}{\partial \mu_i^{(1)}} &= \frac{\ell - c^T \mu}{c^T (\mu^{(1)} - \mu^{(2)} \cdot \mu)} \times \left[ e_i - \frac{c_i \cdot (\mu^{(1)} - \mu^{(2)} \cdot \mu)}{c^T (\mu^{(1)} - \mu^{(2)} \cdot \mu)} \right] \\ \frac{\partial x^*(\ell, m)}{\partial \mu_i} &= \left( \frac{c^T (\mu^{(1)} - \mu^{(2)} \cdot \mu)}{\ell - c^T \mu} - \mu^{(2)} \right) \cdot \frac{\partial x^*(\ell, m)}{\partial \mu_i^{(1)}}. \end{aligned}$$

This implies that the Jacobian of  $x^*$  wrt.  $m$  is given by

$$\frac{\partial x^*(\ell, m)}{\partial m} = \left( \frac{\ell - c^T \mu}{c^T v} \times M \quad \left( 1 - c^T \mu \cdot \frac{\ell - c^T \mu}{c^T v} \right) \times M \right), \quad (\text{B.14})$$

where  $M := I_d - (c^T v)^{-1} v c^T$  and  $v := \mu^{(1)} - c^T \mu \cdot \mu$ . The latter Jacobian has the same rank as  $M$  which has either rank  $d$  or rank  $d - 1$ . Since

$$Mv = v - (c^T v)^{-1} v c^T v = v - v = 0$$

and  $v \neq 0$  (since  $v = \Sigma c$  and  $\Sigma$  has (as a covariance matrix of a non-degenerate distribution) full rank) it follows that the Jacobian has rank  $d - 1$ . Due to the following corollary to [Owe01, Theorem 3.3], which follows directly from a strong law of large numbers and the continuous mapping theorem, we can construct an asymptotic  $1 - \alpha$  confidence region for  $x^*(\ell, \mu)$ :

**Corollary 8.** Let  $n \in \mathbb{N}$  denote the (fixed) sample size and let  $X_i \in \mathbb{R}^p$  ( $i = 1, \dots, n$ ) be independent with common distribution function  $F_0$ , mean  $\mu_0$ , full-rank variance Matrix  $V_0$  and let  $h : \mathbb{R}^p \rightarrow \mathbb{R}^d$  (with  $1 \leq d \leq p$ ) be a mapping that is differentiable with Jacobian  $\partial h / \partial \mu$  of rank  $q > 0$ .

If additionally  $h$  is continuously differentiable in  $\theta_0 := h(\mu_0)$ , then the plug-in estimate of the confidence interval  $C_{r,n}^{(3)}$  from [Owe01, Theorem 3.3] is an asymptotic  $1 - \alpha$  confidence interval for  $\theta_0$  and it is given by

$$\left\{ h(\hat{\mu}) + \sum_{i=1}^n w_i \left( \frac{\partial h}{\partial \mu}(\hat{\mu}) X_i \right) \middle| \prod_{i=1}^n n w_i \geq r, w_i \geq 0, \sum_{i=1}^n w_i = 1 \right\},$$

where  $\hat{\mu} := \frac{1}{n} \sum_{i=1}^n X_i$  and  $r = \exp\left(-\frac{c_\alpha}{2}\right)$  where  $c_\alpha$  is the  $(1 - \alpha)$  quantile of the  $\chi^2$ -distribution with  $q$  degrees of freedom.

### B.3. RESULTS ON METHOD OF MOMENT ESTIMATION FOR THE SKEW-NORMAL AND SKEW-T DISTRIBUTION

Method of moment estimation for some of the parameters of the skew-normal and skew-t distribution can be based on the following result:

**Lemma 18.** *If  $X$  follows a skew-normal distribution with location parameter  $\xi$ , skewness parameter  $\alpha$  and dispersion matrix  $\Omega$  or a skew-t distribution with the additional parameter  $\nu > 2$ , then*

$$\begin{aligned}\mu &= \mathbb{E}[X] = \xi + b \cdot \delta, \\ \Sigma &= \text{Var}(X) = a \cdot \Omega - b^2 \cdot \delta \delta^T, \\ \delta &= \left(1 + \alpha^T \Omega \alpha\right)^{-\frac{1}{2}} \cdot \Omega \alpha,\end{aligned}$$

where  $a = 1$ ,  $b = \sqrt{\frac{2}{\pi}}$  for the skew-normal case and  $a = \frac{\nu}{\nu-2}$ ,  $b = \sqrt{\frac{\nu}{\pi}} \cdot \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})}$  for the skew-t case.

*Proof.* For the skew-normal case see [Cap12, Eq. (2)]. For the skew-t case see [Azz13, Eqns. (6.25), (6.26)].  $\square$

For this paper we are mainly interested in estimators of  $\xi$ ,  $\Omega_c$  and  $\Omega_\alpha$ . These are linearly many (in the dimension of  $X$ ) parameters and one could hope that these can be estimated with linear effort. This is, however, unfortunately not the case and the method of moment estimator can even fail to exist. Furthermore we are not aware of method of moments estimator for the location parameter  $\xi$ . It can, however, be obtained by performing e.g. maximum likelihood estimation for each marginal (using e.g. the `sn` package (cf. [Azz19] and so can  $a$  for the skew-t distribution), details of the estimation of the location parameter in the univariate skew-normal family are discussed by [TP19]). Denote such estimators by  $\hat{\xi}$  and  $\hat{a}$ . Then, by Lemma 18, we have for the other parameters of interest the plug-in estimates

$$\widehat{\Omega_c} := \hat{a}^{-1} \cdot \left( \widehat{\Sigma_c} + (\hat{\mu} - \hat{\xi})(\hat{\mu} - \hat{\xi})^T c \right), \quad (\text{B.15})$$

$$\gamma \cdot \widehat{\Omega_\alpha} := b^{-1} \cdot (\hat{\mu} - \hat{\xi}), \quad (\text{B.16})$$

where  $\hat{\mu} := n^{-1} \sum_{i=1}^n X_i$  and  $\widehat{\Sigma_c} = \left[ n^{-1} \sum_{i=1}^n (c^T X_i) \cdot X_i \right] - \left[ n^{-1} \sum_{i=1}^n c^T X_i \right] \cdot \left[ n^{-1} \sum_{i=1}^n X_i \right]$  are estimators from an i.i.d. sample  $X_1, \dots, X_n$  of  $X$  and where  $\gamma := \alpha^T \Omega \alpha$ . Furthermore we have that

$$\alpha = (1 + \gamma)^{\frac{1}{2}} \cdot \left( \Sigma + b^2 \cdot \delta \delta^T \right)^{-1} \cdot \left( \frac{\mu - \xi}{b} \right)$$

and letting  $\beta := \left( \Sigma + b^2 \cdot \delta \delta^T \right)^{-1} \cdot \left( \frac{\mu - \xi}{b} \right)$  it holds that  $\text{sign}(\alpha_i) = \text{sign}(\beta_i)$  for all  $1 \leq i \leq d$ . If  $\beta_i \neq 0$  then it is simple to verify that the above equation implies that  $\frac{\alpha_i}{\beta_i} \cdot \beta = \alpha$  from which it follows that

$$\alpha_i = \left( \beta_i^2 + \alpha_i^2 \beta^T \Omega \beta \right)^{\frac{1}{2}}.$$

## B

Replacing  $\beta$  by its plug-in estimator  $\hat{\beta}$  (where  $\Sigma$  is replaced by the empirical covariance matrix) it follows that there exists a solution  $\alpha$  if and only if  $\hat{\beta}^T \hat{\Omega} \hat{\beta} < 1$  and in this case it holds that

$$\alpha \approx \hat{\alpha} = \begin{cases} 0 & \text{if } \hat{\beta} = 0 \\ \frac{\hat{\beta}}{\sqrt{1 - \hat{\beta}^T \hat{\Omega} \hat{\beta}}} & \text{if } \hat{\beta}^T \hat{\Omega} \hat{\beta} < 1 \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Hence, the plug-in estimator is not always well defined and necessitates the calculation of a matrix inverse. However, as detailed above, we can estimate  $\Omega_c$  and  $\Omega_\alpha$  up to a multiplicative constant by estimating  $3d$  (i.e. linear in the dimension) parameters. The multiplicative constant  $\gamma$  can, however, only be found by estimating  $O(d^2)$  population parameters and the method of moments estimator is not guaranteed to exist. By Theorem 4 and Theorem 5 we know that the optimizer (under a skew-normal or skew-t model) is an element of the two-dimensional affine subspace

$$\{\xi + a_1 \cdot \Omega_c + a_2 \cdot \Omega_\alpha \mid a_1, a_2 \in \mathbb{R}\} \quad (\text{B.17})$$

and therefore marginal estimation of the location parameters together with (B.15) and (B.16) can be utilized to estimate this subspace from which interesting stress scenarios can be selected. Furthermore for the scenario according to (3.6) it holds that

$$\begin{aligned} x^*(\ell, \Sigma) &:= \mu + \left( \frac{\ell - c^T \mu}{c^T \Sigma c} \right) \Sigma c \\ &= \xi + b \cdot \delta + \left( \frac{\ell - c^T \mu}{c^T \Sigma c} \right) [a \cdot \Omega_c - b^2 \cdot (\delta^T c) \cdot \delta] \\ &= \xi + \left[ b - b^2 \cdot \left( \frac{\ell - c^T \mu}{c^T \Sigma c} \right) \cdot (\delta^T c) \right] \cdot \delta + a \cdot \left( \frac{\ell - c^T \mu}{c^T \Sigma c} \right) \cdot \Omega_c \\ &= \xi + (1 + \alpha^T \Omega \alpha)^{-\frac{1}{2}} \cdot \left[ b - b^2 \cdot \left( \frac{\ell - c^T \mu}{c^T \Sigma c} \right) \cdot (\delta^T c) \right] \cdot \Omega \alpha + a \cdot \left( \frac{\ell - c^T \mu}{c^T \Sigma c} \right) \cdot \Omega c \end{aligned}$$

which implies that, if the skewness is misspecified as  $\alpha = 0$ , the plug-in estimate of the scenario will lie in the affine subspace (B.17), but will in general have a much smaller likelihood than the true reverse stress testing scenario.



# C

## APPENDIX FOR CHAPTER 4

### C.1. PROOFS

*Proof of Corollary 3.* Simply note that  $\tau$  in [Tha01] and  $\tau$  in this publication are equal if  $x_1 = x_n$  and the  $\tau$  in this publication is one smaller otherwise. Thus one only needs to distinguish these cases. This is fortunately simple since the latter kind of sequences correspond to the loops (where one also needs to consider those loops that start with a block of ones). There are (in the notation of Tharrats (2001)) exactly  $2\binom{m-1}{h-1}\binom{n-1}{h-1}$  such loops which yields the desired formula. Thus it follows that

$$\begin{aligned}\mathbb{P}(\tau(X) = x) &= \binom{n}{n_1}^{-1} \begin{cases} 2 \times \binom{n_1-1}{\frac{x+1}{2}-1} \binom{n_2-1}{\frac{x+1}{2}-1} & \text{if } x \text{ is odd} \\ \left(\frac{2n}{x} - 2\right) \times \binom{n_1-1}{\frac{x}{2}-1} \binom{n_2-1}{\frac{x}{2}-1} & \text{if } x \text{ is even} \end{cases} \\ &= \binom{n}{n_1}^{-1} \begin{cases} \frac{(x+1)^2}{2n_1n_2} \times \binom{n_1}{\frac{x+1}{2}} \binom{n_2}{\frac{x+1}{2}} & \text{if } x \text{ is odd} \\ \frac{nx - x^2}{2n_1n_2} \times \binom{n_1}{\frac{x}{2}} \binom{n_2}{\frac{x}{2}} & \text{if } x \text{ is even} \end{cases}\end{aligned}$$

which yields the desired result.  $\square$

*Proof of Lemma 9.* To show the symmetry write

$$x = m \pm a = \begin{cases} m + a & \text{if } x \geq a \\ m - a & \text{if } x < 0 \end{cases}$$

for  $x \in [2m-1]$  and  $a > 0$ . Then

$$\mathbb{P}(\tilde{\tau}_{n,n_1,n_2} = x) = \begin{cases} (m \pm a + 1)^2 \times G_{m,m}(m \pm a + 1) & \text{if } x \text{ is odd} \\ (m^2 - a^2) \times G_{m,m}(m \pm a) & \text{if } x \text{ is even} \end{cases}$$

Thus, for  $m \in \mathbb{N}, a \in \mathbb{N}$  and  $m \pm a$  odd it suffices to show that  $(m + a + 1)^2 \times G_{m,m}(m + a + 1) = (m - a + 1)^2 \times G_{m,m}(m - a + 1)$ . To verify this claim one only needs to check the identity

$$(m + a + 1) \times \binom{m}{\frac{m+a+1}{2}} = (m - a + 1) \times \binom{m}{\frac{m-a+1}{2}}.$$

To this end it is helpful to distinguish two cases:

- If  $m$  is odd (and thus  $a$  is even), then it is sufficient to verify that

$$(m+a+1) \times \binom{2m+1}{m+a+1} = (m-a+1) \times \binom{2m+1}{m-a+1}$$

for arbitrary  $m, a \in \mathbb{N}$ .

- If  $m$  is even (and thus  $a$  is odd), then it is sufficient to verify that

$$(m+a+1) \times \binom{2m}{m+a+1} = (m-a) \times \binom{2m}{m-a}$$

for arbitrary  $m, a \in \mathbb{N}$ .

Furthermore for  $m \pm a$  even it holds, that  $G_{m,m}(m+a) = G_{m,m}(m-a)$  and thus the density of  $\tau(X)$  is symmetric about  $m$  and therefore  $\mathbb{E}[\tau(X)] = m$ .

$\mathbb{E}[\tau(X)] = \frac{2n_0n_1}{n}$  holds for  $n_0 \neq n_1$  since it holds that

$$\begin{aligned} \mathbb{E}[\tau(X)] &= \sum_{x=1}^{2(n_0 \wedge n_1)} x \mathbb{P}(\tau(X) = x) \\ &= \sum_{x=1}^{n_0 \wedge n_1} 2x \mathbb{P}(\tau(X) = 2x) + (2x-1) \mathbb{P}(\tau(X) = 2x-1) \\ &= \sum_{x=1}^{n_0 \wedge n_1} G_{n_0, n_1}(2x) \times [2x(2nx - (2x)^2) + (2x-1) \cdot (2x-1+1)^2] \\ &= 4(n-1) \times \sum_{x=1}^{n_0 \wedge n_1} x^2 G_{n_0, n_1}(2x) = \frac{(n-1)}{n_0 n_1 \binom{n}{n_0}} \times \sum_{x=1}^{n_0 \wedge n_1} x^2 \binom{n_0}{x} \binom{n_1}{x} \\ &= \frac{2n_0 n_1}{n} \end{aligned}$$

where the last equality follows from the relation

$$\sum_{h=1}^{m \wedge n} h^2 \binom{m}{h} \binom{n}{h} = \frac{m^2 n^2}{(m+n)(m+n-1)} \binom{m+n}{m}$$

proved in Tharrats (2001) on p. 8.

For  $n_0 = n_1 = m$  it holds that

$$\begin{aligned} \mathbb{E}[\tau(X)^2] &= \sum_{x=1}^{2m-1} x^2 \mathbb{P}(\tau(X) = x) \\ &= \sum_{x=1}^m G_{m,m}(2x) \times [(2x)^2 \times (2x \cdot (2m) - (2x)^2) + (2x-1)^2 \times (2x)^2] \\ &= \frac{2}{m^2 \binom{2m}{m}} \times \left[ \sum_{x=1}^m (x^2 + 4(m-1)x^3) \binom{m}{x}^2 \right] \\ &= \frac{2}{m^2} \times \left[ \frac{m^3}{2(2m-1)} + 4(m-1) \frac{m^3}{4(2m-1)} (1+m) \right] \end{aligned}$$

$$= \frac{m}{2m-1} + \frac{2(m-1)m(m+1)}{2m-1} = \frac{m(2m^2-1)}{2m-1}$$

where the exact relations (13,2) and (13,3) of Tharrats (2001) were used. Thus the formula for the variance follows from  $\text{Var}(\tau(X)) = \mathbb{E}[\tau(X)^2] - m^2$ .  $\square$

*Proof of Lemma 11.* All that is necessary to apply [Lee90, Corollary 1 on p. 83] is to calculate the conditional expectation  $\psi_2((x_1, y_1), (x_2, y_2)) = \mathbb{E}[D_n((x_1, x_2, X_3, X_4, X_5), (y_1, y_2, Y_3, Y_4, Y_5))]$  and reformulate the eigenvalue problem:

Since  $D_n$  only depends on the ranks of the  $Y_i$  it can be, w.l.o.g., assumed that  $y_1, y_2 \in [0, 1]$  and that  $Y_3, Y_4, Y_5$  are i.i.d. uniform on  $[0, 1]$ .

The probability of  $A_{y_1, y_2} := \text{Rank}_{y_1, y_2, Y_3, Y_4, Y_5}(y_1, y_2) \in \{(1, 4), (1, 5), (2, 4), (2, 5)\}$  where

$$\text{Rank}_M(y_1, y_2) := (|\{y \in M : y \leq \min\{y_1, y_2\}\}|, |\{y \in M : y \leq \max\{y_1, y_2\}\}|)$$

is given by

$$\begin{aligned} P(A_{y_1, y_2}) &= 3(F_{\max} - F_{\min})^2 \cdot (1 - F_{\max}) + (F_{\max} - F_{\min})^3 \\ &\quad + 6F_{\min} \cdot (F_{\max} - F_{\min}) \cdot (1 - F_{\max}) + 3F_{\min} \cdot (F_{\max} - F_{\min})^2 \end{aligned}$$

where  $F_{\min} = \min\{y_1, y_2\}$  and  $F_{\max} = \max\{y_1, y_2\}$ . This can be simplified to

$$\mathbb{P}(A_{y_1, y_2}) = (-3 + 2F_{\min}) \cdot (F_{\min})^2 + (3 - 2F_{\max}) \cdot (F_{\max})^2.$$

Furthermore for  $B_{y_1, y_2} := \text{Rank}_{y_1, y_2, Y_3, Y_4, Y_5}(y_1, y_2) \in \{(1, 2), (4, 5)\}$  it holds that

$$\mathbb{P}(B_{y_1, y_2}) = (F_{\min})^3 + (1 - F_{\max})^3$$

and some algebraic manipulations yield

$$16 \cdot \mathbb{P}(B_{y_1, y_2}) - 8 \cdot \mathbb{P}(A_{y_1, y_2}) = 8 \cdot (3((F_{\min})^2 + (F_{\max})^2) - 6F_{\max} + 2)$$

Thus, in summary

$$\psi_2((x_1, y_1), (x_2, y_2)) = p^{2-(x_1+x_2)} \cdot (1-p)^{1+x_1+x_2} \cdot (-1)^{x_1+x_2} \cdot 8 \cdot (3((F_{\min})^2 + (F_{\max})^2) - 6F_{\max} + 2)$$

and it follows that

$$\begin{aligned} &\mathbb{E}_{(x_2, y_2)}[\psi_2((x_1, y_1), (x_2, y_2)) \cdot f(x_2, y_2)] \\ &= \sum_{x_2=0}^1 \left[ (1-p)^{1-x_2} \cdot p^{x_2} \cdot \int_0^1 \psi_2((x_1, y_1), (x_2, y_2)) \cdot f(x_2, y_2) dy_2 \right] \\ &= \sum_{x_2=0}^1 \left[ p^{2-x_1} \cdot (1-p)^{2+x_1} \cdot (-1)^{x_1+x_2} \cdot 8 \cdot \int_0^1 k(y_1, y_2) \cdot f(x_2, y_2) dy_2 \right] \\ &= 8 \cdot p^{2-x_1} \cdot (1-p)^{2+x_1} \cdot (-1)^{x_1} \cdot \sum_{x_2=0}^1 \left[ (-1)^{x_2} \cdot \int_0^1 k(y_1, y_2) \cdot f(x_2, y_2) dy_2 \right] \\ &= c(x_1) \cdot \int_0^1 k(y_1, y_2) \cdot (f(0, y_2) - f(1, y_2)) dy_2 \end{aligned}$$

for any  $f \in \mathcal{L}^2$  and where  $k(y_1, y_2) := 3((F_{\min})^2 + (F_{\max})^2) - 6F_{\max} + 2$  and  $c(x_1) := 8 \cdot p^{2-x_1} \cdot (1-p)^{2+x_1} \cdot (-1)^{x_1}$ . Since  $c(0) = \frac{p}{p-1} \cdot c(1)$  it follows that if  $f$  is an eigenfunction with eigenvalue  $\lambda_f \neq 0$ , then

$$\frac{\lambda_f}{c(1)} f(1, y_1) = \frac{\lambda_f}{c(1)} \frac{p-1}{p} f(0, y_1)$$

which implies  $f(1, y_1) = \frac{p-1}{p} f(0, y_1)$  and thus the eigenproblem can be reduced to solving

$$\lambda_f f(x_1, y_1) = c(x_1) \cdot \left( \frac{p-1}{p} - 1 \right) \int_0^1 k(y_1, y_2) \cdot f(1, y_2) dy_2 = -\frac{c(x_1)}{p} \int_0^1 k(y_1, y_2) \cdot f(1, y_2) dy_2$$

for either  $x_1 = 0$  or  $x_1 = 1$ . For  $x_1 = 1$  this yields

$$\lambda_f f(1, y_1) = \frac{c(1)}{p-1} \int_0^1 k(y_1, y_2) \cdot f(1, y_2) dy_2 = 8 \cdot p \cdot (1-p)^2 \cdot \int_0^1 k(y_1, y_2) \cdot f(1, y_2) dy_2$$

and thus for any eigenvalue  $\lambda$  of  $\varphi \mapsto \int_0^1 k(y_1, y_2) \varphi(y_2) dy_2$  the value  $8 \cdot p \cdot (1-p)^2 \cdot \lambda$  is an eigenvalue of the operator of interest.  $\square$

## C.2. CALCULATION OF U-STATISTICS BASED ON CHATTERJEE'S RANK CORRELATION $\xi$

Code demonstrating how to calculate U-statistics based on Chatterjee's  $\xi$  efficiently. It is based on the idea that one can simply update an array of counts of all relevant subsequences. The first argument of `calc_xi` should be the vector  $(X_{\pi(1)}, \dots, X_{\pi(n)})$ .

```

1 function calc_xi_add0!(l::Array{Float64,4},m::Int64)
2     for j=reverse(2:m)
3         for k=1:m
4             l[1,k+1,j,1] += l[1,k,j-1,1]
5             for i=2:m
6                 l[i,k+1,j,1] += l[i,k,j-1,1] + l[i-1,k,j-1,2]
7             end
8         end
9     end
10    l[1,2,1,1] += 1
11 end
12
13 function calc_xi_add1!(l::Array{Float64,4},m::Int64)
14     for j=reverse(2:m)
15         for k=1:(m+1)
16             l[1,k,j,2] += l[1,k,j-1,2]
17             for i=2:m
18                 l[i,k,j,2] += l[i,k,j-1,2] + l[i-1,k,j-1,1]
19             end
20         end
21     end
22    l[1,1,1,2] += 1
23 end
24
25 function calc_xi_mat(s,m)
26     #jumps+1 x n_0+1 x length x last_bit
27     l = zeros(m,m+1,m,2)
28     for v in s
29         if v==0
30             calc_xi_add0!(l,m)
31         else
32             calc_xi_add1!(l,m)
33         end
34     end
35     # jumps+1 x n_0+1
36     dropdims(sum(l[:, :, m, :], dims=3), dims=3)

```

```

37 end
38
39 function calc_xi(s,m)
40     jumps = 1:(m-1)
41     n0 = (1:(m-1))'
42     n0n1 = n0 .* (m .- n0)
43     sum(calc_xi_mat(s,m)[2:end,2:(end-1)] .* jumps ./ n0n1) /
44         binomial(BigInt(length(s)),m)
45 end

```

C

### C.3. GENERATING RANDOM BITS WITH NON-UNIFORM PROBABILITY

Most programming languages offer the ability to uniformly sample unsigned integers and therefore blocks of uniformly distributed bits (i.e. block of i.i.d. Bernoulli random variables with probability  $p = 0.5$ ). There is usually also some way to sample bits with arbitrary  $p \in (0, 1)$ , but in e.g. the Julia programming language

```

1 import StatsBase
2 x=falses(2^20*8);
3 StatsBase.direct_sample!([true,false],StatsBase.AnalyticWeights([0.25,0.75]),x);

```

it is general purpose (`[true, false]` could be replaced by any subtype of `AbstractArray`) and therefore not as fast as a specialized version. For example on my AMD Ryzen 5 2600 it generates about 8 MiB/s of randomness which is a lot less than the approximately 7000 MiB/s of randomness the default Mersenne Twister pseudo random number generator (PRNG) produces. Indeed it is possible to generate bits with non-uniform probability (biased bits) more efficiently: As described in e.g. [Gry21, Proposition 3.1], if  $p \in (0, 1)$  has the expansion  $(b_i)_{i \in \mathbb{Z}}$  in base 2 where  $b_i \in \{0, 1\}$  and  $p = \sum_{i=1}^{\infty} \frac{b_i}{2^i}$ , then the following algorithm produces a biased bit which takes the value 1 with probability  $p$  and the value 0 with probability  $1 - p$ : Let  $(\tilde{b}_i)_{i \in \mathbb{N}}$  be a sequence of bits that are i.i.d. Bernoulli with  $\tilde{p} = 0.5$ . Denote by  $n := \min\{i \in \mathbb{N} : \tilde{b}_i = 1\}$  the first index such that  $\tilde{b}_n = 1$  and emit  $b_n$ . It can be easily seen that this is correct since the probability of looking at  $b_i$  is exactly  $2^{-i}$ . This is also quite efficient since the expected number of bits used to output one bit is two (cf. [Gry21]). For a practical implementation there are, however, relevant details not discussed in [Gry21]

1. How does one gain access to the binary expansion of  $p$ ? If it is stored as a floating point number, then the precision is already limited to finitely many non-zero binary digits  $b_i$ . Thus I advocate for only using  $b_1, \dots, b_{64}$  (for  $p$  represented as an IEEE 754-2008 binary64 point number, see Code 3 for an example). This decision has two main consequences: Firstly sampling with probabilities less than  $2^{-64} < 5.5 \cdot 10^{-20}$  is not possible, but this is unlikely to matter for most practical purposes. Secondly the absolute error in the probability with which the sampling is performed will be less than  $5.5 \cdot 10^{-20}$ .
2. What is the best strategy (in terms of performance) to manage the source PRNG? While it may, at first, seem like it would be best to use only as many bits as needed it turns out that there is a performance benefit to always using exactly 64 bits of randomness.<sup>1</sup>

<sup>1</sup>The first design choice implies that one can assume that all  $b_i$  for  $i > 64$  are zero. Thus the output of the algorithm will always be 0 if  $n > 64$  and it would therefore be pointless to use more bits from the source PRNG.

```

1 function getBinaryDigits(x::Float64)::Tuple{UInt64,UInt8}
2     @assert 0 < x && x < 1
3     mantissa::UInt64 = reinterpret(Unsigned,x) & Base.significand_mask(Float64)
4                       | 1 << Base.significand_bits(Float64)
5     adj = -exponent(x) - 1;
6     digits::UInt8 = Base.significand_bits(Float64)
7                   + 1 - trailing_zeros(mantissa) + adj;
8     (mantissa << (Base.exponent_bits(Float64)-adj), digits)
9 end

```

Code 3: Code demonstrating how to extract the first 64 bits in the base 2 expansion  $0.b_1b_2\dots b_{64}$  from binary64 floating point number in the Julia programming language.

Since PRNGs usually have quite a large state it is often not advisable to generate random numbers in an otherwise tight loop (i.e. a loop whose body is CPU cache and register friendly) since this can lead to a performance penalty. This can be avoided by filling a buffer with the needed random numbers before performing the actual calculations (in this case generation of the biased bits). To achieve optimal cache utilization the buffer size needs to be chosen carefully. The observed performance suggests that it is best to choose a buffer size that is approximately as large as the L1 cache available to each CPU core.

The following two subsections detail the relevant implementation details.

### C.3.1. SIMPLE IMPLEMENTATION USING 64 BITS OF RANDOMNESS

This subsection presents different variants of generating biased random bits. The variants can be obtained by either changing the constant at the beginning or eliminating the `@static` if's that don't match the desired variant and keeping only the bodies of those that do. The choice is between two methods of looking at the value of  $b_i$ :

1. `update_method=mask` Left shift followed by masking
2. `update_method=lookup` Using a lookup table.

Furthermore three variants of buffer management are considered:

1. `alloc_method=full` Immediately obtain all random numbers
2. `alloc_method=small_chunks` Obtain  $64 \cdot 64 = 4096$  bits at a time
3. `alloc_method=chunks` Obtain approximately `buf_size*64` bits (defaults to 16384 bits).

A few remarks on the details implementation are in order:

1. The `@static` if's are evaluated at compile-time. They therefore do not influence the benchmarks.
2. The (pseudo) random bit generation is performed in 64 bit chunks. Each chunk starts with all bits set to zero (line 49/80) and thus the update can be done branch free (cf. e.g. line 90). This is possible because it is known beforehand that the bit starts off with the value zero

Table C.1: Comparison of the measured performance (cf. [CR16]) of the different approaches discussed in this section when generating random bits with  $p = 0.25$ . The AMD machine had an AMD Ryzen 5 2600 cpu and the Intel machine had an Intel Core i7-4765T. Since these cpus are very different in terms of generation and tdp this is not meant to be a comparison of the performance on the different cpus. Instead the goal is to demonstrate, that the findings regarding the relative performance of the different approaches are not platform specific.

CPU type	Buffer allocation	Digit lookup	Peak MiB/s	Average MiB/s
AMD	chunks	lookup table	74.813	74.344
	chunks	mask	44.986	43.744
	full	lookup table	36.557	33.056
	small_chunks	lookup table	36.320	34.161
	full	mask	28.499	25.916
	small_chunks	mask	27.611	27.336
Intel	chunks	lookup table	50.296	47.923
	chunks	mask	44.011	42.298
	full	lookup table	31.574	30.125
	full	mask	29.015	28.020
	small_chunks	lookup table	26.296	23.898
	small_chunks	mask	24.394	22.928

C

which cannot be changed by the logical or. Thus the branch present in the implementation of core Julia<sup>2</sup> is not necessary.

- It may not seem optimal to write e.g. `chunks[i+1]` since this suggests an additional `add` instruction may be emitted by the compiler. This is however not the case since Julia uses 1-based indices. Unless the compiler has an opportunity for some kind of optimization it is actually `chunks[i]` that leads to an additional `add` instruction, cf. Code 4. Thus iterating with 0-based indices is preferable.
- `leading_zeros` will usually compile to a single instruction (e.g. `lzcnt` on x86 platforms).

According to the observed performance (cf. Table C.1) it is best to use a lookup table and it is important to choose the buffer size correctly. In practice using approximately the size of the L1 cache of a single core seems to work best, yielding approximately 75 MiB/s on my AMD Ryzen 5 2600 cpu.

```

1 import Random
2
3 const update_method = "lookup"
4 const alloc_method = "chunks"
5
6 function getDigitsArray(x::Float64)::Array{UInt64,1}
7     b = BitVector(undef, 64);
8     b.chunks[1] = getBinaryDigits(x)[1];
9     if b.chunks[1] == UInt64(0)
10         @warn "No non-zero digits"
11     end
12     vcat(map(x -> ifelse(x, 0x8000000000000000, UInt64(0)),
13         reverse(collect(b))), 0x0)
14 end
15
16 function bitrand_test(n,p::Float64,buf_size::Int64=-1)
17     @static if update_method == "lookup"
18         ba = getDigitsArray(p);
19     end
20     @static if update_method == "mask"
21         (mantissa::UInt64, digits::UInt8) = getBinaryDigits(p);

```

<sup>2</sup><https://github.com/JuliaLang/julia/blob/ae8452a9e0b973991c30f27beb2201db1b0ea0d3/base/bitarray.jl#L689>

C

```

22     end
23
24     mx = UInt32(ceil(n/64));
25
26     @static if alloc_method == "full"
27         buf_size = 64*(64*mx);
28         buf = Random.bitrand(buf_size);
29         chunks = buf.chunks;
30         k = 0;
31     end
32     @static if alloc_method == "small_chunks"
33         buf = BitArray(undef, 64*64)
34         chunks = buf.chunks;
35     end
36     @static if alloc_method == "chunks"
37         if buf_size < 0
38             buf_size = 1 << 14;
39         end
40
41         buf_size = min(buf_size, n);
42         buf_size = UInt32(ceil(buf_size/64)) * 64;
43
44         buf_chunks = Array{UInt64,1}(undef, buf_size)
45     end
46
47     ret::BitArray{1} = BitArray(undef, mx * 64);
48     ret_chunks = ret.chunks;
49
50     @static if alloc_method == "full" || alloc_method == "small_chunks"
51         for j::UInt64=1:mx
52             chunk::UInt64 = 0
53             @static if alloc_method == "small_chunks" Random.rand!(chunks) end
54             for i::UInt8=0:63
55                 @static if alloc_method == "full"
56                     @inbounds b::UInt64 = chunks[k+=1];
57                 end
58                 @static if alloc_method == "small_chunks"
59                     @inbounds b::UInt64 = chunks[i+1];
60                 end
61                 @static if alloc_method == "chunks"
62                     @inbounds b::UInt64 = buf_chunks[buf_i += 1];
63                 end
64                 cnt::UInt8 = leading_zeros(b);
65                 @static if update_method == "lookup"
66                     @inbounds chunk |= (ba[cnt+1] >> i);
67                 end
68                 @static if update_method == "mask"
69                     mantissa_bit::UInt64 = mantissa << cnt;
70                     condition::UInt64 = mantissa_bit & 0x8000000000000000;
71                     chunk |= (condition >> i);
72                 end
73             end
74             @inbounds ret_chunks[j] = chunk;
75         end
76     end
77     @static if alloc_method == "chunks"
78         j::UInt64 = 0;
79         while j < mx
80             Random.rand!(buf_chunks)
81             buf_i = 0
82             while buf_i < buf_size
83                 chunk::UInt64 = 0
84                 for i::UInt8=0:63
85                     @inbounds b::UInt64 = buf_chunks[buf_i += 1];
86                     cnt::UInt8 = leading_zeros(b);
87                     @static if update_method == "lookup"
88                         @inbounds chunk |= (ba[cnt+1] >> i);
89                     end
90                     @static if update_method == "mask"
91                         mantissa_bit::UInt64 = mantissa << cnt;
92                         condition::UInt64 = mantissa_bit & 0x8000000000000000;

```



```

93         chunk |= (condition >> i);
94     end
95 end
96 @inbounds ret_chunks[j += 1] = chunk
97 end
98 end
99 end
100 return ret;
101 end

```

C

```

julia> test(b::Array{Float64,1},i::Int64) = @inbounds b[i]
test (generic function with 1 method)

julia> test1(b::Array{Float64,1},i::Int64) = @inbounds b[i+1]
test1 (generic function with 1 method)

julia> @code_llvm test(zeros(100),10)

; @ REPL[2]:1 within `test'
; Function Attrs: uwtable
define double @julia_test_765(%jl_value_t* nonnull align 16 dereferenceable(40),
    i64) #0 {
top:
;   @ array.jl:809 within `getindex'
    %2 = add i64 %1, -1
    %3 = bitcast %jl_value_t* %0 to double**
    %4 = load double*, double** %3, align 8
    %5 = getelementptr inbounds double, double* %4, i64 %2
    %6 = load double, double* %5, align 8
;   L
    ret double %6
}

julia> @code_llvm test1(zeros(100),10)

; @ REPL[3]:1 within `test1'
; Function Attrs: uwtable
define double @julia_test1_772(%jl_value_t* nonnull align 16 dereferenceable(40),
    i64) #0 {
top:
;   @ array.jl:809 within `getindex'
    %2 = bitcast %jl_value_t* %0 to double**
    %3 = load double*, double** %2, align 8
    %4 = getelementptr inbounds double, double* %3, i64 %1
    %5 = load double, double* %4, align 8
;   L
    ret double %5
}

```

Code 4: A small example demonstrating that taking the  $(i+1)$ -th element of an array does not cause an extra instruction to be emitted by the Julia compiler.

### C.3.2. MORE EFFICIENT UTILISATION OF THE SOURCE RANDOM NUMBER GENERATOR

The expected number of bits obtained from 64 bits of source randomness using the algorithm of [Gry21, Proposition 3.1] is 32 bits. Thus it is quite wasteful to instead always use 64 bits of source randomness to generate only a single bit. Indeed, repeatedly applying this algorithm to the remaining bits will succeed most of the time and this is the idea behind the implementation presented in Code 5. Since this yields approximately 95 MiB/s of randomness on my AMD Ryzen 5 2600, this is the implementation I recommend.

Some further remarks on the implementation Code 5:

- The `& 63` in e.g. line 34 can be thought of as a hint to the compiler that the sign does not matter (in Julia shifting by a negative number has the semantic of instead shifting in the opposite direction) and that no wrap-around happens (since the shift will always be less than 64).
- If `shift > 63`, then the code in line 40 will not work as expected. But, since the buffer is anyways refilled due to `remaining ≤ 0`, this is not a correctness issue.

C

```

1  function bitrand(n,p::Float64,buf_size::Int64=-1)
2      # (mantissa::UInt64, digits::UInt8) = getBinaryDigits(p);
3
4      ba = getDigitsArray(p);
5
6      mx = UInt32(ceil(n/64));
7
8      if buf_size < 0
9          buf_size = 1 << 14;
10     end
11
12     buf_size = min(buf_size, n);
13     buf_size = max(128, buf_size)
14     buf_size = UInt32(ceil(buf_size/64)) * 64;
15
16     buf_chunks = Array{UInt64,1}(undef, buf_size)
17
18     r::UInt64 = Random.rand(UInt64);
19     remaining::Int8 = 64;
20
21     ret = BitArray{UInt64}(undef, n);
22     ret_chunks = ret.chunks;
23
24     j::UInt64 = 0;
25     while j < mx
26         Random.rand!(buf_chunks)
27         buf_i = 0
28         while buf_i+128 <= buf_size && j < mx
29             chunk::UInt64 = 0
30             for i::UInt8=0:63
31                 if r == 0 || remaining <= 0
32                     @inbounds r = buf_chunks[buf_i += 1];
33                     if remaining > 0
34                         r >>= remaining & 63;
35                     end
36                     remaining = 64;
37                 end
38                 cnt::Int8 = leading_zeros(r);
39                 shift::Int8 = cnt+1;
40                 r <<= shift & 63;
41                 remaining -= shift;
42                 @inbounds chunk |= (ba[cnt+1] >> (i & 63));
43             end
44             @inbounds ret_chunks[j += 1] = chunk
45         end
46     end
47
48     return ret;
49 end

```

Code 5: The suggested implementation of a variant of the algorithm of [Gry21, Proposition 3.1].

## C.4. EFFICIENT CALCULATION OF $D_n$ AND $\tau_n^*$

Similarly to the calculation of the U-statistic based on  $\xi_n$ , the calculation of  $D_n$  and  $\tau_n^*$  boils down to counting how often certain patterns appear as subsequences of  $\tilde{X} = (X_{\pi(1)}, \dots, X_{\pi(n)})$ . This can be done in time linear in  $n$  by using a simple principle which I will explain using an example: To count e.g. the number of times the pattern  $(0, 1, 0, 1)$  appears as a subsequence of  $\tilde{X}$  one can simply store the counts of all prefixes in an array and update them using a recursive relationship. For the example pattern the prefixes are  $(0), (0, 1), (0, 1, 0), (0, 1, 0, 1)$  and their counts can be stored in  $v \in \mathbb{Z}^4$ . The initial value of  $v$  is  $v = (0, 0, 0, 0)$  and it gets updated according to the following rules (for  $i = 1, \dots, n$ ):

- If  $\tilde{X}_i = 0$  then  $v_1 := v_1 + 1$  and  $v_3 := v_3 + v_2$
- If  $\tilde{X}_i = 1$  then  $v_2 := v_2 + v_1$  and  $v_4 := v_4 + v_3$

Code 6 demonstrates how to implement this idea to calculate  $\tau_n^*$  using Single Instruction Multiple Data (SIMD) parallelism by using the `shufflevector` intrinsic to arrange the vectors in a way optimal for the updates. Interestingly this manual vectorization is not necessary: Simply unrolling the inner loop as demonstrated in Code 7 allows Julia's LLVM backend<sup>3</sup> to automatically vectorize the code, yielding similar performance (albeit slightly different assembly instructions) to the manual vectorization. The code for calculating  $D_n$  can be found in Code 8.

```

1  import SIMD
2  function calc_tau_star_SIMD_(s::Main.BitArray)
3      cnt0 = 0.0
4      cnt1 = 0.0
5      m0 = (Main.SIMD).Vec{8, Main.Float64}(0.0)
6      m1 = (Main.SIMD).Vec{8, Main.Float64}(0.0)
7      idx0 = (Main.SIMD).Val{(:undef, :undef, 8, 1, 9, 12, 2, 10)}
8      idx1 = (Main.SIMD).Val{(:undef, :undef, 1, 0, 9, 11, 3, 4)}
9      @inbounds for val = s
10         if !val
11             v0 = (Main.SIMD).shufflevector(m0, m1, idx0)
12             v0 = (Main.Base).setindex(v0, cnt0, 1)
13             v0 = (Main.Base).setindex(v0, cnt1, 2)
14             m0 += v0
15             cnt0 += 1
16         else
17             v1 = (Main.SIMD).shufflevector(m0, m1, idx1)
18             v1 = (Main.Base).setindex(v1, cnt1, 1)
19             v1 = (Main.Base).setindex(v1, cnt0, 2)
20             m1 += v1
21             cnt1 += 1
22         end
23     end
24     return [m1[7], m1[6], m0[6], m0[7], m0[8], m1[8]]
25 end
26
27 function calc_tau_star(s::Main.BitArray)
28     sum(calc_tau_star_SIMD_(s) .* [-8, 16, -8, 16, -8, -8]) /
29     (n * (n-1) * (n-2) * (n-3))
30 end

```

Code 6: Single Instruction Multiple Data (SIMD) Julia code (cf. [Sch+19]) demonstrating how to calculate  $\tau_n^*$ .

<sup>3</sup>Specifically the SLP Vectorizer; cf. <https://llvm.org/docs/Vectorizers.html#the-slp-vectorizer>.

## C

```

1 function calc_tau_star_(s::Main.BitArray)
2     cnt0 = 0.0
3     cnt1 = 0.0
4     m = Main.zeros(16)
5     @inbounds for val = s
6         if !val
7             m[8] += m[11]
8             m[7] += m[3]
9             m[6] += m[13]
10            m[5] += m[10]
11            m[4] += m[2]
12            m[3] += m[9]
13            m[2] += cnt1
14            m[1] += cnt0
15            cnt0 += 1
16        else
17            m[16] += m[5]
18            m[15] += m[4]
19            m[14] += m[12]
20            m[13] += m[10]
21            m[12] += m[1]
22            m[11] += m[2]
23            m[10] += cnt0
24            m[9] += cnt1
25            cnt1 += 1
26        end
27    end
28    @inbounds return [m[15], m[14], m[6], m[7], m[8], m[16]]
29 end
30
31 function calc_tau_star(s::Main.BitArray)
32     sum(calc_tau_star_(s) .* [-8, 16, -8, 16, -8, -8]) /
33     (n * (n-1) * (n-2) * (n-3))
34 end

```

Code 7: Example code demonstrating how to calculate  $\tau_n^*$  in Julia.

```

1 import SIMD
2 function calc_D_SIMD_(s::Main.BitArray)
3     cnt0 = 0.0
4     cnt1 = 0.0
5     m0 = (Main.SIMD).Vec{8, Main.Float64}(0.0)
6     m1 = (Main.SIMD).Vec{8, Main.Float64}(0.0)
7     extra = (Main.SIMD).Vec{4, Main.Float64}(0.0)
8     idx0 = (Main.SIMD).Val{(:undef, :undef, 8, 0, 9, 1, 11, 12)}
9     idx1 = (Main.SIMD).Val{(:undef, :undef, 10, 12, 13, 2, 15, :undef)}
10    idx_extra = (Main.SIMD).Val{13, 4, 14, 3}
11    @inbounds for val = s
12        if !val
13            extra += (Main.SIMD).shufflevector(m1, m0, idx_extra)
14            v0 = (Main.SIMD).shufflevector(m1, m0, idx0)
15            v0 = (Main.Base).setindex(v0, cnt0, 1)
16            v0 = (Main.Base).setindex(v0, cnt1, 2)
17            m0 += v0
18            cnt0 += 1
19        else
20            v1 = (Main.SIMD).shufflevector(m1, m0, idx1)
21            v1 = (Main.Base).setindex(v1, cnt1, 1)
22            v1 = (Main.Base).setindex(v1, cnt0, 2)
23            v1 = (Main.Base).setindex(v1, extra[1], 8)
24            m1 += v1
25            cnt1 += 1
26        end
27    end
28    return [m1[6], extra[4], extra[2], m1[7], extra[3], m1[8]]
29 end
30
31 function calc_D(s::Main.BitArray)
32     sum(calc_D_SIMD_(s) .* [16, -8, -8, -8, 16, -8]) /
33     (4 * n * (n-1) * (n-2) * (n-3) * (n-4))
34 end

```

Code 8: Single Instruction Multiple Data (SIMD) Julia code (cf. [Sch+19]) demonstrating how to calculate  $D_n$ .