Carlos Renato Belo Azevedo

Anticipation in Multiple Criteria Decision-Making Under Uncertainty

Antecipação na Tomada de Decisão com Múltiplos Critérios sob Incerteza

Campinas
2014

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação

Carlos Renato Belo Azevedo

Anticipation in Multiple Criteria Decision-Making Under Uncertainty

Antecipação na Tomada de Decisão com Múltiplos Critérios sob Incerteza

Thesis presented to the School of Electrical and Computer Engineering of the University of Campinas in partial fulfillment of the requirements for the degree of Doctor in Electrical Engineering, in the area of Computer Engineering

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Elétrica, na área de Engenharia da Computação

Orientador: Prof. Dr. Fernando José Von Zuben

Este exemplar corresponde à versão final da tese defendida pelo aluno Carlos Renato Belo Azevedo, e orientada pelo Prof. Dr. Fernando José Von Zuben

Campinas
2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

Az25a

Azevedo, Carlos Renato Belo, 1984-
Anticipation in multiple criteria decision-making under uncertainty / Carlos Renato Belo Azevedo. – Campinas, SP : [s.n.], 2014.

Orientador: Fernando José Von Zuben.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Inteligência artificial. 2. Aprendizado de máquina. 3. Inferência bayesiana. 4. Otimização multiobjetivo. 5. Otimização de carteiras de investimento. I. Von Zuben, Fernando José,1968-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Antecipação na tomada de decisão com múltiplos critérios sob incerteza
Palavras-chave em inglês:
Artificial intelligence
Machine learning
Bayesian inference
Multi-objective optimization
Investment portfolio optimization
Área de concentração: Engenharia de Computação
Titulação: Doutor em Engenharia Elétrica
Banca examinadora:
Fernando José Von Zuben [Orientador]
Ricardo Hiroshi Caldeira Takahashi
Rodrigo Fernandes de Mello
Paulo Augusto Valente Ferreira
Romis Ribeiro de Faissol Attux
Data de defesa: 15-08-2014
Programa de Pós-Graduação: Engenharia Elétrica

COMISSÃO JULGADORA - TESE DE DOUTORADO

Candidato: Carlos Renato Belo Azevedo

Data da Defesa: 15 de agosto de 2014

Título da Tese: "Anticipation in Multiple Criteria Decision-Making Under Uncertainty"

Prof. Dr. Fernando José Von Zuben (Presidente): ____________________

Prof. Dr. Romis Ribeiro de Faissol Attux: ____________________

Prof. Dr. Ricardo Hiroshi Caldeira Takahashi: ____________________

Prof. Dr. Rodrigo Fernandes de Mello: ____________________

Prof. Dr. Paulo Augusto Valente Ferreira: ____________________

Abstract

The presence of uncertainty in future outcomes can lead to indecision in choice processes, especially when eliciting the relative importances of multiple decision criteria and of long-term vs. near-term performance. Some decisions, however, must be taken under incomplete information, what may result in precipitated actions with unforeseen consequences. When a solution must be selected under multiple conflicting views for operating in time-varying and noisy environments, implementing flexible provisional alternatives can be critical to circumvent the lack of complete information by keeping future options open. Anticipatory engineering can be then regarded as the strategy of designing flexible solutions that enable decision makers to respond robustly to unpredictable scenarios. This strategy can thus mitigate the risks of strong unintended commitments to uncertain alternatives, while increasing adaptability to future changes. In this thesis, the roles of anticipation and of flexibility on automating sequential multiple criteria decision-making processes under uncertainty are investigated. The dilemma of assigning relative importances to decision criteria and to immediate rewards under incomplete information is then handled by autonomously anticipating flexible decisions predicted to maximally preserve diversity of future choices. An online anticipatory learning methodology is then proposed for improving the range and quality of future trade-off solution sets. This goal is achieved by predicting maximal expected hypervolume sets, for which the anticipation capabilities of multi-objective metaheuristics are augmented with Bayesian tracking in both the objective and search spaces. The methodology has been applied for obtaining investment decisions that are shown to significantly improve the future hypervolume of sets of trade-off financial portfolios for out-of-sample stock data, when compared to a myopic strategy. Moreover, implementing flexible portfolio rebalancing decisions was confirmed as a significantly better strategy than to randomly choosing an investment decision from the evolved stochastic efficient frontier in all tested artificial and real-world markets. Finally, the results suggest that anticipating flexible choices has lead to portfolio compositions that are significantly correlated with the observed improvements in out-of-sample future expected hypervolume.

Key-words: Intelligent systems. Anticipatory engineering. Multiple criteria decision-making. Multi-objective optimization. Uncertainty handling. Stochastic optimization. Metaheuristics. Portfolio optimization.

Resumo

A presença de incerteza em resultados futuros pode levar a indecisões em processos de escolha, especialmente ao elicitar as importâncias relativas de múltiplos critérios de decisão e de desempenhos de curto vs. longo prazo. Algumas decisões, no entanto, devem ser tomadas sob informação incompleta, o que pode resultar em ações precipitadas com consequências imprevisíveis. Quando uma solução deve ser selecionada sob vários pontos de vista conflitantes para operar em ambientes ruidosos e variantes no tempo, implementar alternativas provisórias flexíveis pode ser fundamental para contornar a falta de informação completa, mantendo opções futuras em aberto. A engenharia antecipatória pode então ser considerada como a estratégia de conceber soluções flexíveis as quais permitem aos tomadores de decisão responder de forma robusta a cenários imprevisíveis. Essa estratégia pode, assim, mitigar os riscos de, sem intenção, se comprometer fortemente a alternativas incertas, ao mesmo tempo em que aumenta a adaptabilidade às mudanças futuras. Nesta tese, os papéis da antecipação e da flexibilidade na automação de processos de tomada de decisão sequencial com múltiplos critérios sob incerteza é investigado. O dilema de atribuir importâncias relativas aos critérios de decisão e a recompensas imediatas sob informação incompleta é então tratado pela antecipação autônoma de decisões flexíveis capazes de preservar ao máximo a diversidade de escolhas futuras. Uma metodologia de aprendizagem antecipatória on-line é então proposta para melhorar a variedade e qualidade dos conjuntos futuros de soluções de trade-off. Esse objetivo é alcançado por meio da previsão de conjuntos de máximo hipervolume esperado, para a qual as capacidades de antecipação de metaheurísticas multi-objetivo são incrementadas com rastreamento bayesiano em ambos os espaços de busca e dos objetivos. A metodologia foi aplicada para a obtenção de decisões de investimento, as quais levaram a melhoras significativas do hipervolume futuro de conjuntos de carteiras financeiras de trade-off avaliadas com dados de ações fora da amostra de treino, quando comparada a uma estratégia míope. Além disso, a tomada de decisões flexíveis para o rebalanceamento de carteiras foi confirmada como uma estratégia significativamente melhor do que a de escolher aleatoriamente uma decisão de investimento a partir da fronteira estocástica eficiente evoluída, em todos os mercados artificiais e reais testados. Finalmente, os resultados sugerem que a antecipação de opções flexíveis levou a composições de carteiras que se mostraram significativamente correlacionadas com as melhorias observadas no hipervolume futuro esperado, avaliado com dados fora das amostras de treino.

Palavras-chave: Sistemas inteligentes. Engenharia antecipatória. Tomada de decisão com múltiplos critérios. Otimização multiobjetivo. Tratamento de incerteza. Otimização estocástica. Metaheurísticas. Otimização de carteiras.

Summary

To my soulmate, best friend, and wife, Flávia Érika, with love: thank you for leading me to this moment and for adding meaning, joy, and purpose to my life.

Acknowledgments

A PhD thesis is not a book of memories and life experiences; yet it is, in a profound sense that is usually only fully comprehensible to none but its sole author. In order to celebrate the memories and experiences that I've registered in each and every word, symbol, and equation written throughout its hundreds of pages, I'd like to make sure the acknowledgments section is memorable enough to be perceived by its readers as the single most important aspect of being a researcher, graduate student, and human, for that matter. By doing so, perhaps I'll be also contributing for a better comprehension of the contributions of the thesis itself.

So, I would like to begin by thanking the person who is most responsible for this: my wonderful spouse, Flávia Érika, for everything she represents to me. I honestly can't imagine how I could have reached this far without her: from encouraging me to change to a computer science major back in 2004, to paying the professional costs of moving to Campinas in 2011. I hope those ended up being the most flexible choices maximizing our current joint happiness! I've learned and I'll always learn from her sense of dedication, fairness, responsibility, and integrity in everything she does. I'm so lucky to have you in my life!

I thank the world's best mother, Dione, for all her immense love and care, despite the geographic distance. I thank my father, Magal, for not giving up to transmit me his love, care, and values, on his own manner, despite the temporal distance. I thank both of them for showing me that love is transcendental and has no spatial or temporal boundaries.

I thank my sister Carol and my brothers, Louro and Rafa, for their love, prays, and encouragement. They are amongst the most altruist people I've ever met and I thank them for letting me be part of their lives.

I'm also convinced I wouldn't be as motivated to achieve this goal without the support and encouragement of the following key people.

My PhD advisor, Prof. Fernando Von Zuben, has always been a source of profound inspiration and a role model for my life and career. I can unreservedly say that he is committed to the formation of whole human beings, in the first place. Paraphrasing what he once said about the relationship between advisors and students, both parts must be capable of capturing the pain and joy of being the other part. My feeling after these extremely intense 3 and a half years is that we have successfully managed to achieve that goal together. As a result, I feel I'm a better human being than I was three years ago, for what I'll be forever grateful!

Moreover, Prof. Fernando must be credited for planting the seeds that motivated me to

investigate the role of *anticipation* in decision-making. I vividly remember that it was on his lecture about evolutionary biology during the Evolutionary Computation class, in March 2011, that I payed close attention to his observation that the emergence of *anticipatory behavior* on living beings was an interesting open question that could possibly be explained by evolutionary means. That moment prompted my unsuspected silent journey on investigating the concept of anticipation in biology.

Flexibility is a concept that was already rooted in my mind as an important one, since the first time I've read Branke and Mattfeld's paper [38] on the role of flexible ad-hoc heuristics to dynamic scheduling in 2010. Since then, I've intended to investigate the effects of ad-hoc heuristics of mine on inventory and routing control problems subject to stochastic demand.

In 2012, I abandoned that idea for a more general one that ultimately lead to the original PhD project that generated this thesis. As a result, the influence of flexibility on the experimental results that I had analyzed – although already highlighted in the thesis – was still not completely clear to me. That last until February 2014, when I was allowed to invest substantial effort to improve my understanding about the role of flexibility on the models and algorithms that I had already developed by that time, thanks to my FAPESP scholarship being extended for an additional 6 months period – for which I'm deeply grateful to the anonymous FAPESP reviewer for trusting in this work.

The truth is that Prof. Fernando never let the concept of flexibility go from my mind. He made sure to remind me of the importance of flexibility, whenever we met to discuss my thesis progress. In March 2014, I thus decided to invest the remaining five months of my PhD scholarship to go deeper into the relationship between hypervolume maximization and flexibility. The progress I've made on this topic during the final months of wrapping up the thesis text proved to be of immeasurable value and I thank Prof. Fernando for all the encouragement and wise guidance on this effort.

I thank my instructors at UNICAMP for their dedication and their excellent teaching skills: Profs. Fernando Von Zuben, Renato Lopes, Paulo Valente, Fernando Gomide. I also thank Profs. Joydeep Ghosh and Arindam Banerjee for their didacticism and competency during their brief but valuable time as visiting professors in 2012 at UNICAMP. All those lecturers may not be aware of this, but it was by attending their classes that many seeds that ultimately lead to this thesis were planted. I also thank all the FEEC/DCA staff for contributing to an incredibly supportive environment: Noêmia, Carmem, Adriana, Jerusa, and cia.

I thank the members that took part in the thesis committee evaluation: Profs. Fernando Von Zuben (FEEC/UNICAMP), Ricardo Takahashi (DM/UFMG), Rodrigo de Mello (ICMC/USP), Paulo Valente (FEEC/UNICAMP), and Romis Attux (FEEC/UNICAMP). The time and dedication each of them invested on not only appreciating this manuscript, but, mainly, on stimulating my mind throughout the thesis defense process have had a profound impact on my life and on the final version of the manuscript. I'm deeply thankful for their interest and trust on this work.

I'm grateful to Prof. Arindam Banerjee for being so kind to host me at his lab in the Univ. of Minnesota, Twin Cities campus, for a technical visit throughout the summer of 2013. I have greatly benefited from his experience and was inspired by his kindness and passion for pursuing

grand research goals.

I thank all my past teachers and instructors I had during my academic journey. I'd like to thank Prof. Múcio Queiroz as a representative of all of them. I was an undergrad teaching assistant of Prof. Múcio in 2005 when I was still lost in terms of my life calling and aspirations. I've always been impressed by his stories about how he was one of the first people in Brazil to get a Masters degree in computer science and how he co-founded the Department of Statistics and Informatics at the Catholic University of Pernambuco as well as the now prestigious Center for Informatics at the Federal University of Pernambuco. I remember that it was during his classes in 2004 when I first decided that I wanted to go through advanced graduate studies for the sake of giving back to society what he has always demonstrated to his students: the joy in the pursuit of knowledge.

I thank all my past advisors and research collaborators who taught me many valuable assets that were essential for my time at UNICAMP.

Prof. Tiago Ferreira was my first advisor in the scientific initiation undergrad program. Tiago is a visionary leader, a scientist by heart, and a great appreciator of Nature, deeply interested in comprehending how it works. He once said to me that quantum computation was "*such a beautiful thing*". I guess I've inherited his curiosity and enthusiasm. Thank you for being the first to show me what science really means!

Prof. Francisco Madeiro was my second scientific initiation advisor (and long-time tennis partner), but his enthusiasm for science is second to none! His working ethic and dedication for inspiring and helping young aspiring scientists to grow meant the world to me. I was very lucky to be able to learn from him when I was still giving my first steps into research. For instance, I remember well that it was him who taught me the basics of Shannon's entropy and Huffman coding, by estimulating my mind through lists of exercises from his draft book chapters on information theory. Moreover, I vividly remember his entropy-based proposals of *flexible* competitive learning clustering algorithms for which the "winners-take-all" rule for updating prototypes are relaxed by entropy terms. The influence of his lessons in my academic training has proven to be incommensurable. Thank you for contributing to my growth both as a person and as a researcher!

Prof. Scott Gordon trusted me while I was still an undergrad student and accepted my daring invitation in 2008 to co-write a joint research paper for the biggest conference at the time in evolutionary computation. The rest is history: we not only had one of the most memorable experiences of remotely going through the whole research preparation; paper writings; and the anxiety of the acceptance notification, but we also ended up finally meeting and having a great time at GECCO'09, in Montreal. My admiration for his trustiness, dedication, open-mindedness and didacticism will forever remain. Thank you for you friendship and continuous encouragement, Scott!

Prof. Claus Aranha also trusted me in 2009, when I was still a master student, and invited me to participate in a little part of his doctoral research on methods for financial portfolio optimization, after we met at GECCO'09. Needless to say that I've learned so much with him and from his research that I ended up seeing this area as a wonderful benchmark for my own research. Claus is also one of the most free-spirited and motivated people that I've ever met.

I'm inspired by his dedication and passion to teaching and researching and I wish him the best in his still young career.

Prof. Aluizio Araújo was my master degree advisor. As a researcher with grand aspirations, he was the one who first suggested me to work with multi-objective optimization in dynamic environments, a scaring and challenging research topic. Besides, during the months dedicated to the proper development of my masters project, I remember he encouraged me to investigate model reference-based methods for guiding diversity generation in populations of candidate solutions evolved with bio-inspired metaheuristics. Although I was not able to fully grasp the potential of the approach at the time, and despite the fact that my limited knowledge made me uncomfortable with the idea of accepting that challenge, by looking in retrospect, I can now see that the work delivered in this thesis was somehow influenced by my interaction with Prof. Aluizio. I'm impressed by his ability of advising and conducting me to achieve our research goals at a time where I was losing focus rapidly. Thank you very much for leading me into the right track, Prof. Aluizio.

The friends and colleagues that I made at Singularity University in 2010 are also unforgettable. That was one of the most profound lifetime experiences I've had to this day. My life was dramatically changed during that summer, for which I learned how to recognize (and separate) what I've always been looking for: like-minded people truly willing to make the world a better place.

I thank the relatively few but precious friends that I've made throughout my life. When I began this journey in 2011, I couldn't imagine that I would have the privilege of meeting so many more high-spirited and friendly people at UNICAMP, including my lab colleagues (and friends) at the Laboratory of Bioinformatics and Bio-inspired Computing (LBiC), UNICAMP, who have been kind enough to let me take part of so many fruitful discussions and activities. It is hard for me to not list all of my friends here, but I feel I must choose some over others as representatives. But to those not cited, you know who you are. Each of you know what I'm grateful for. A big thank you for sharing your knowledge, life experiences, and time with me.

I thank my friend and lab colleague André Gonçalves for convincing me, although he might not be aware of that, to use Kalman Filters for Bayesian learning instead of Gaussian Processes, as was originally specified in my PhD project. I definitely think in retrospect that this was the most flexible choice for this thesis. Besides being a top-notch researcher, André is one of the most gentle and insightful people that I known. Thank you for your friendship and intellectual guidance!

I thank my colleagues from APOGEEU, the graduate student association at FEEC, for trusting me and rapidly integrating me into so many joyful activities: Alan Godoy, Edgar, Eliezer, Rafael, Raul, and Thisien. I not only have been touched by their hard work for alleviating the struggles of the hundreds of students at FEEC/UNICAMP, but also learned and benefited from their leadership and dedication to the community.

I thank all the members of the student chapter of the IEEE CIS officially formed in 2012 at FEEC/UNICAMP. I'm happy that we truly did great work together and helped to inspire many other people on engaging in interdisciplinary research and to integrate otherwise isolated students into a conversation about the future of our areas of expertise.

I thank Profs. Armando da Rocha (FM/USP), Fabio Cozman (Poli/USP), Fernando Von Zuben (FEEC/UNICAMP), João Rosa (ICMC/USP), Paul Bourgine (École Polytechnique de Palaiseau), Renata Wassermann (IME/USP), Ricardo Gudwin (FEEC/UNICAMP), and Romis Attux (FEEC/UNICAMP) for their insightful and brilliant participation in the series of debates organized by our IEEE CIS student chapter on the practical and philosophical aspects regarding the future development of Artificial Intelligence.

I thank the FEEC/UNICAMP Profs. Fernando Von Zuben, Max Costa, Reginaldo Palazzo Jr., and Romis Attux for their passion and commitment to making the IEEE CIS student chapter organized debate on the intersections between artificial intelligence and information theory a memorable one.

I thank Prof. Reginaldo Palazzo Jr. for encouraging me to come to Campinas in 2009, although I ended up not attending the FEEC master degree program, staying safe in Recife. During the IEEE CIS organized debate, Prof. Reginaldo also encouraged us to focus on the geometric aspects of data analysis, instead of only reasoning over patterns of occurrences. His insight was key in motivating me to think about the subtle yet beautiful geometric interpretations of representing uncertainty in simplex spaces using Dirichlet distributions.

A special thank you to all my friends from the undergrad and master courses, who have inspired me in many different ways. The list is large, but I wish to pick my friend, Rafael Xavier, as a representative. "Dr. Ralph" was my classmate during the undergrad years and has also been through a tough journey for the pursuit of knowledge and for the sake of being professionally recognized as a scientist – we are all scientists since the childhood. Among many acts of generosity and comprehension, he presented me in 2009 with the famous book "The Evolution of Cooperation" by Robert Axelrod, a classic of the game theoretic literature on how competition and cooperation evolve over time, based on the study of elementary models of social and individual rewards arising from the interaction between two "rational" agents. This was one small but persistent seed that has somehow greatly influenced my views and my way of understanding the world. Thank you, Ralph, for your patience and for your generosity throughout all those years.

A special thank you to Dr. Rodrigo Pasti, who, besides being an admired scientist and great appreciator of Nature, was the first to greet us when we arrived in Campinas. I've never forgotten his willingness to help a young couple to find a place and to be integrated into the city. Thank you, Rodrigo!

Another special thank you to Conrado Miranda who, besides being a super-reliable and humorous friend, took the time to revise all the manuscript and contributed to substantially improve it, after a laborious reviewing work that took him weeks. If only more people were that dedicated and concerned with the quality and purpose of science...

I also thank Prof. Fabricio Olivetti who, besides being an inspiring creative and brilliant scientist, also did the same laborious reviewing work for the first version of this manuscript, as a member of my doctoral qualification exam board, and also helped me with the resulting submitted journal paper. I admire his altruism and willingness to make the world a better place. Thank you, Fabricio!

Besides Conrado and Prof. Fabricio, I'd like to thank Saullo Oliveira, Rodrigo Setti, and

Eliezer da Silva for accepting my challenge of improving not only the way scientists collaborate and work together, but, mainly, to help them to reunite with the very reasons of why they have chosen to pursue such career. It was certainly not only to publish papers and be personally rewarded by the resulting professional reputation that comes out of it; but to search for the truth and to engineer a better future for all humanity. I'm sure someday the world will hear about **Sci-Torrent** (http://scitorrent.github.io) and/or similar initiatives. Thank you for believing in this joint effort, my friends!

Last, but not least, I thank my good friends and tennis partners in Campinas with whom I enjoyed so many memorable experiences: Evandro Silva, Felipe Martinez, Guilherme, and Marcelo Ramos.

This thesis is also dedicated in memoriam of my good friend, lab colleague, brilliant scientist, and remarkable human being, Salomão Madeiro (2013). I shall never forget his enthusiasm for discovering new things; the tennis and soccer matches; the free Warcraft 3 strategy lessons; the jokes told while sharing a cup of coffee; the political discussions; and the mentally demanding yet enjoyable work done together. You'll never be forgotten; your ideas will live on and your dreams will be carried on by each of us. Thank you for every moment, my friend!

I acknowledge and thank the Brazilian funding agencies CNPq and FAPESP for the two doctoral scholarships that allowed me to develop my research under the best conditions. The research presented in this manuscript benefited from a CNPq PhD scholarship from March 2011 to September 2012, and from a FAPESP PhD scholarship from October 2012 until August 2014.

Finally, a big thank you for all dreamers all around the world (alive or dead) who have been working hard every day for a sustainable future wherein the human society can awake to its full potential and take action so that social balance, justice, and freedom can prevail over all forms of psychological impairment and intellectual slavery.

Campinas, September 23, 2014,
Carlos Renato Belo Azevedo

“You can have peace. Or you can have freedom. Don’t ever count on having both at once.”

Robert Anson Heinlein

List of Figures

List of Tables

List of Pseudocodes

List of Abbreviations

AMFC	Anticipated Maximal Flexible Choice
AS-MOO	Anticipatory Stochastic Multi-Objective Optimization
ASMS-EMOA	Anticipatory $\mathcal{S}$-Metric Selection Evolutionary Multi-Objective Algorithm
DD	Dirichlet Distribution
DJI	Dow Jones Index
DMOEA	Dynamic Multi-Objective Evolutionary Algorithm
EMFC	Estimated Maximal Flexible Choice
EMOO	Evolutionary Multi-Objective Optimization
FTL	Follow The Leader
FTSE	Financial Times Stock Exchange
GP	Gaussian Process
HSI	Hang Seng Index
Hypv	Hypervolume
KF	Kalman Filter
MAP	Maximum A Posteriori
MCDM	Multiple Criteria Decision-Making
MLE	Maximum Likelihood Estimation
MVP	Mean-Variance Problem
mHDM	maximal Hypv Decision Maker
MOEA	Multi-Objective Evolutionary Algorithm
MOO	Multi-Objective Optimization
MPT	Modern Portfolio Theory
NSGA-II	Fast Elitist Non-dominated Sorting Genetic Algorithm
OAL	Online Anticipatory Learning
PD	Pareto Dominance
PF	Pareto Frontier
PFR	Preferred Feasible Region
POCID	Prediction Of Change in Direction
PS	Pareto Set
RDM	Random Decision Maker
ROI	Return Over Investment

ROOT	Robust Optimization Over Time
RSMS-EMOA	Regularized $\mathcal{S}$-Metric Selection Evolutionary Multi-Objective Algorithm
SMOO	Stochastic Multi-Objective Optimization
SPF	Stochastic Pareto Frontier
SP	Stochastic Process
SR	Sharpe Ratio
TIP	Time Incomparability Probability
TL	Time-Linkage
TLF	Time-Linkage Free
TR	Turnover Rate
UP	Universal Portfolio

List of Symbols

$\mathbf{f}$	Vector-valued cost-adjusted random objective functions.
$\mathbf{g}$	Vector-valued original random objective functions.
$\mathbf{h}$	Vector-valued cost function.
$\mathbf{u}_t^{\max}$	Current decision vector of the maximal flexible choice.
$\mathbf{z}_t^{\max}$	Current objective vector of the maximal flexible choice.
$\mathbf{u}_t^{(i)}$	i-th ranked decision vector in terms of the first objective mean ($m_{z_{1,t}^{(i)}}$).
$\mathbf{z}_t^{(i)}$	i-th ranked random objective vector in terms of the first objective mean ($m_{z_{1,t}^{(i)}}$).
$\hat{\mathbf{u}}_t^{(i)}$	DD MAP estimation of $\mathbf{u}_t^{(i)}$.
$\hat{\mathbf{m}}_{\hat{\mathbf{u}}_t^{(i)}}$	Mean vector of the DD MAP estimation of $\mathbf{u}_t^{(i)}$.
$\hat{\mathbf{z}}_t^{(i)}$	Kalman filter estimation of $\mathbf{z}_t^{(i)}$.
$\hat{\mathbf{z}}_t \vert \hat{\mathbf{z}}_{t+1:t+H-1}$	Anticipatory distribution of $\hat{\mathbf{z}}_t$ given the predictive distributions $\hat{\mathbf{z}}_{t+1} \vert \hat{\mathbf{z}}_t, \cdots, \hat{\mathbf{z}}_{t+H-1} \vert \hat{\mathbf{z}}_t$, computed by using the OAL rule of Eq. (6.10).
$\tilde{\mathbf{Z}}_{1:t-1}$	Sum of historical objective vector KF squared residuals for a fixed portfolio.
$\mathcal{F}_t^N$	Set of N vector-valued objective evaluations over mean DD MAP vectors.
$\mathcal{S}$	Hypervolume function.
$\hat{\mathcal{S}}_{t+1}$	Sample average Hypv computed for future out-of-sample test data.
$\mathcal{Z}$	Feasible objective space.
$\pi(\mathcal{Z})$	set of all non-empty subsets of $\mathcal{Z}$.
Ω	Feasible search space.
$\Omega_t^\star \vert \mathbf{u}_{t-1}$	Current dynamic PS given $\mathbf{u}_{t-1}$ was taken.
$\mathcal{Z}_t^\star \vert \mathbf{u}_{t-1}$	Current dynamic SPF given $\mathbf{u}_{t-1}$ was taken.
$\hat{\mathcal{Z}}_t^{N\star} \vert \mathbf{u}_{t-1}$	Finite dynamic SPF estimation with N anticipatory distributions.
$\hat{\mathbf{U}}_t^{N\star} \vert \mathbf{u}_{t-1}$	Finite dynamic Pareto Set DD MAP estimation with N mean decision vectors.
$\mathbb{E}\left[\Delta_{\mathcal{S}}(\hat{\mathbf{z}}_t^{(i)})\right]$	Expected Hypv contribution of a given anticipatory distribution in the objective space.
$\preceq$	Weak-Pareto Dominance.
$\preccurlyeq$	Set-based weak-Pareto Dominance.
$\Vert$	Pareto incomparability.

$\mathcal{W}_t$	Current investor mean wealth.
$\boldsymbol{\mu}_t$	Current mean vector of asset returns.
$\boldsymbol{\Sigma}_t$	Current covariance matrix of asset returns.
$\mathcal{X}_t$	Set of current parameters of the joint assets return distribution.
η	Severity of change parameter.
τ	Periodicity of change parameter.
S^{d-1}	A $(d-1)$–Simplex subspace in $\mathcal{R}^d$.
N	Number of trade-off alternatives.
T	Number of decision stages (investment periods).
K	Window size of KF and of DD MAP estimation.
H	Anticipation horizon.
$\lambda_t^{(i)}$	Anticipation rate of the i-th ranked objective vector.
$\lambda_t^{(\mathcal{H})}$	Anticipation rate computed from the entropy of the incomparability probability between a current and a predicted objective distributions
$\lambda_t^{(\mathcal{K})}$	Anticipation rate computed from the normalized sum of historical Kalman Filter residuals

Chapter 1

Introductory Remarks

Those who cannot tell what they desire or expect, still sigh and struggle with indefinite thoughts and vast wishes.

– Ralph Waldo Emerson

Each choice is made in the context of whatever value system we have selected to govern our lives. In selecting that value system, we are, in a very real way, making the most important choice we will ever make.

– Benjamin Franklin

1.1 On the Nature of Change and the Emergence of Conflicts

The most self-evident and therefore fundamental property of Nature is that of *change*. Understanding and explaining how physical, chemical, and biological processes change is the ultimate goal of all natural sciences. As part of those processes, human behavior is at the same time constrained and driven by forces of natural change. The impetus of changing oneself towards an imagined and idealized future goal is nonetheless not exclusive of humans, but must be common to all sentient beings capable of *planning ahead* – whether when hunting a prey, climbing a tree, or landing safely at a target, regardless of circumstances.

Furthermore, if desires are considered as drivers of action towards change, then the inability of *anticipating* such desires (or *preferences*) of citizens, customers, users – or whatever group of individuals one is trying to cooperate with – may contribute to severe mismatches. Such discrepancies between what is (wrongly) expected and what is actually observed can thus generate frustration and, hence, lead to *conflicts* of interest. Failure to understand such processes of change – and to effectively cope with the conflicts that emerge out of it – can thus increase our perception of the world as *uncertain*.

In this context, it is not uncommon to hear about companies that spend significant time and effort planning, developing, and marketing products only to find that customer interests have already shifted toward new niches and opportunities. For example, in the emerging niche market of wearable computers, customers are reportedly concerned about the privacy of smart-glasses users and their driving safety[1]. Suddenly, preference for fancy and flyweight smart-glasses designs lost importance for less intrusive and distracting optical interfaces.

Fortunately, despite the frustration that comes from failures in planning and execution (and because of it), the ability of *learning from experience* provides sentient beings, organizations, and societies with the means for *adapting* such plans – as well as their execution – so better alignments between now refined future goals and reality can be achieved.

1.1.1 Three Grand Challenges Solved by Automation

In this thesis, we argue that three grand challenges for effectively adapting plans towards refined goals in scales wherein there is little central coordination between independent decision-makers, such as in certain human organizations, are

1. How to cope with multiple conflicting objectives – such as safety improvement, risk mitigation, quality control, and cost reduction;

2. How to best coordinate resources and direct effort to search for effective *new solutions* that can fill the needs which are *foreseen* to allow for improved *future* decisions; and

3. How to define what decisions would be *better* when assessing them under multiple views.

Clearly, the answers to any of the aforementioned challenges also shape the answers to the remaining ones. This thesis thus claims to investigate, for the first time, computational answers to all three challenges. In fact, to the best of our knowledge, the reported research is the first to jointly address such issues within an algorithmic and computational approach, and under a *machine learning* framework.

1.2 Research Goal: The Automation of Multiple Criteria Decision-Making

We summarize the logical chain followed until now in order to set the research problem addressed in this thesis: (a) processes change ⇒ (b) people change ⇒ (c) mismatches between imagined goals and reality occur ⇒ (d) conflicts emerge ⇒ (e) we thus perceive the world as uncertain.

It turns out artificial intelligence and machine learning provide the cornerstones for coping with such uncertainty regarding how to deal with the three aforementioned challenges. Nonetheless, we argue that most researches in (narrow) artificial intelligence have been *rushing* into a conservative, *myopic* agenda of rapid near-term incremental progress towards engineering inflexible and unadaptable algorithmic behavior and designs.

Furthermore, we argue that such narrow designs have been mostly aimed to *reducing* uncertainty, instead of *embracing* it. As a result, the research outcome of narrow artificial intelligence has been mostly composed of ad-hoc and problem-specific *decision-making* tools, which

[1]http://en.wikipedia.org/wiki/Smartglasses#Safety_considerations

are hindered by offline parameter tuning and unreasonable assumptions, and whose applicability cannot be easily extended to other problem domains. Although such inflexible machine learning algorithms can efficiently converge and identify “optimal” solutions to a given narrow context/objective, they usually cannot effectively cope with changes of the context/objective.

The resulting inflexible decision-making systems that take advantage of optimization technologies found in the narrow artificial intelligence literature have nonetheless mostly been designed for *a priori planning*, where one *fixed* solution is developed beforehand to operate in the same way, regardless of any eventual changes that may occur in the operational environment. Because those systems are built from the premise that the optimization problem, i.e., the objective-function to be optimized, is static, they do not possess the capability of reacting to changes and adapting old decisions to novel operational scenarios [211].

Therefore, this thesis proposes *automating Multiple Criteria Decision-Making* (MCDM) processes so that uncertainty can be *fully embraced*, instead of being myopically reduced. In order to achieve that, we developed methods for self-adapting the *confidence* which is internally represented by the proposed flexible machine learning algorithms in response to both (a) the observed discrepancies between expected and realized consequences of trade-off actions; and (b) the underlying uncertainty regarding predictions of future trade-off actions. See **chapter 5** and **chapter 6** for an in-depth discussion.

1.2.1 Research Agenda: Automating the Resolution of Conflicts

In 1772, Benjamin Franklin wrote a letter to a friend seeking advise from him on a complex matter. In that letter, he showed one of the early references to a *systematic decision-making* methodology based on evaluating pros and cons of all consequences resulting from a decision taken under uncertainty, including the assignment of weights of importances to each future consequence [143]. This letter is considered by the International Society on Multiple Criteria Decision Making as the “*earliest known reference relating to*” MCDM[2]. The full text of this letter is reproduced in this thesis **appendix**.

In a bold defense of a pragmatic agenda for human and social development, Roberto M. Unger indicated automation as a means to “*accelerate the production of the new*” [209]:

> “*A way to accelerate the production of the new is to turn the way people work together into a social embodiment of the imagination: their dealings with one another mimic the moves of experimental thought. To this end, the first requirement is that we save energy and time for whatever cannot yet be repeated. Whatever we can repeat we express in a formula and then embody in a machine. Thus, we shift the focus of energy and attention away from the already repeatable, toward the not yet repeatable.*”

It turns out, however, that it is not only narrow tasks that humanity has already learned how to repeat. In fact, several natural and biologically-inspired complex behavior have been modeled and experimented with in computational simulations for a myriad of general tasks, ranging from facial recognition using artificial neural networks; intrusion detection with artificial

[2]Source: http://mcdmsociety.org /facts.html.

immunological systems; or the automated search for a set of "best" solutions inspired by natural evolution – what is indeed performed and investigated in this thesis [65].

The inception of bio-inspired metaheuristics and hyper-heuristics can be also considered as attempts of modeling natural behavior for automating complex and otherwise computationally intractable problem-solving. Research on hyper-heuristics has been described in Burke et al. [45] as that comprising "*a set of approaches*" for "*automating the design of heuristic methods to solve hard computational search problems*" [45].

The research agenda posed in this thesis follows a similar path in which the goal is to automate a general decision-making process subject to *multiple conflicting views* about how to *generate and select* a sequence of candidate trade-off solutions for operating in changing and uncertain environments. Such sequential MCDM process could model in principle e.g. the same task automated by hyper-heuristics: that of generating and allocating computational resources for the most promising (and less risky) candidate heuristics to solve a given problem over time.

On the Unreliability of Preferences Elicitation Under Uncertainty

Automating MCDM processes is nonetheless an ambitious challenge and it would be a reckless move to claim that humanity is close to learning how to represent in a formula or algorithm the whole experience of rationality towards acting consciously about current and future choice trade-offs so to autonomously commit to a dynamic range of different preferences over time.

As shown by the 2002 economics Nobel laureate, Daniel Kahneman, human beings are subject to a myriad of *cognitive biases* [129] underlying their intuitive thought processes. In fact, the recent myriad of experimental research on social sciences has been challenging the assumptions of reductionists rational models of decision-making based on *value and risk*.

The idea that decision-makers in general *always know what they truly want* is obviously illusory when confronted to daily subjective and social experiences. One of the core reasons behind this indeterminacy of desires and goals are the often neglected interplay between primary needs, emotional states, and conscious thought about choice trade-offs.

For instance, in the book "Predictably Irrational" [7], research on how emotional states affect daily decisions that are modeled under simplistic assumptions in the decision theory literature was summarized by the Duke University professor of behavioral economics, Dan Ariely. In one of such studies, students from Duke University took part in a tickets lottery for the title match of the National College Athletic Association men's basketball league. When 100 students among winners and losers of the lottery were asked to price the tickets (for selling to non-owners or for buying from owners), a huge discrepancy (by a factor of 14) was observed: owners asked on average $2,400 dollars and buyers were willing to pay on average only $175. Ariely then had the following comment about such discrepancy [7]:

> "*From a rational perspective, both the ticket holders and the non-ticket holders should have thought of the game in exactly the same way. After all, the anticipated atmosphere at the game and the enjoyment one could expect from the experience should not depend on winning a lottery. Then how could a random lottery drawing have changed the students' view of the game – and the value of the tickets – so dramatically?*"

He later argued that cognitive biases related to the over-valuation (for ticket holders) and undervaluation (for non-ticket holders) of *ownership* could have been responsible for triggering different emotional states that affected the pricing of the tickets [7].

Recent works support the notion that human decision-makers are subject to a vast range of psychological effects that influence their answers and choices, contributing to a deviation from what would be considered as "rationality" by decision-theoretic accounts based on e.g. utility theory (see **chapter 2**, section 2.1.4 for a more in-depth discussion).

For instance, Nyhan and Reifler [162] showed that incentivizing American citizens participants to write about good life experiences that increased their self-esteem tended to increase the likelihood of the participants revealing, when asked about, previously known economical facts which contradict their prior political beliefs about the US economical agenda. The authors then concluded that "(...) *people may already know the correct answers but resist acknowledging them due to their threatening nature under normal circumstances*".

In addition, Kremer et al. [139] conducted a series of payed experiments which assessed the decisions taken by each of a total of 152 participants selected for managing a newsvendor supply under uncertain demand. The participants were asked to take decisions so to maximize profit, but were always offered the option of paying an additional fee to make the order decisions under full knowledge about the demand. The results showed that the decision-makers over-payed for eliminating the risks of the supply not matching the demand, what corroborated to the notion that the participants had hidden contextual preferences that did not match any utility-based risky choice model, such as risk-aversion or risk-neutrality. The authors concluded that their results "(...) *point to context-sensitive decision accounts such as the desire to minimize ex post inventory errors*" [139].

1.2.2 On Three Heuristic Meta-Preferences for Automating Sequential MCDM

In order to circumvent the necessity of eliciting preferences from psychologically and emotionally impaired decision-makers, this thesis proposes three *meta-preferences* for automating sequential MCDM processes, backed by the economics; the decision-theoretic; the neuroscience literature; and derived from the author's own intuition and subjective reflections. Those heuristic meta-preferences are:

1. Preference for flexibility;
2. Preference for long-term predictability; and
3. Preference for trajectories of minimal historical errors.

This thesis thus engineers those three heuristics into an autonomous decision maker (or agent) with continuous access to (a) a memory of multiple trade-off candidate choices; (b) complete historical information about discrepancies w.r.t. expected and observed choice outcomes; and (c) an internal predictive model. With those and other key components (see **chapter 6**), including a flexible and anticipatory mathematical model of behavior (see **chapter 5**), we design Bayesian predictive methods and optimization algorithms enabling resolution of conflicts[3], despite the *absence of any a priori preference specification* by human supervisors, except for the a priori *choice of the value system*, which comprises the specification of:

- Which set of decision criteria (or objective-functions) should be considered in the trade-off analysis[4] (e.g., the moments of multiple cost function distributions, see **chapter 3**);
- Meta-preferences automating the dynamic commitment to lower-level preferences over temporal rewards (e.g., long-term predictability, see **chapter 6**);
- Meta-preferences automating the dynamic commitment to lower-level preferences over error distributions (e.g., trajectories of minimal mean square historical prediction errors, see **chapter 6**); and
- Meta-preferences automating the dynamic commitment to lower-level preferences over choice menus (e.g. flexibility, see **chapter 3** and **chapter 5**).

Our proposal of automating MCDM implies the design of *flexible machine learning algorithms* and decision-making tools which are capable of operating under severe uncertainty, despite erroneous modeling assumptions about the changing behavior of the problem being solved. As a result, the flexible tools and methods proposed in this thesis are designed so as to be capable of *self-adapting* to the problem characteristics and of acting despite *lack of complete information* about *future consequences* of candidate actions. By doing so, we hope to improve the learning experience resulting from their application to challenging sequential[5] MCDM problems.

1.2.3 Research Applicability

Open-ended research on computational systems for *automating decision-making* has been rapidly unfolding within many application areas of machine learning and operation research. Perhaps the most popular examples are intelligent systems allowing for self-driving cars, e.g. [208], and certain types of industrial robots, e.g. [207]. Other examples of automated decision-making in practice include interactive recommendation and automated negotiation in distributed electronic commerce systems [144]; and the automatic revision of routes and schedules in vehicle routing subject to dynamic requests [183].

One of the most essential technologies at the core of such endeavor – of computing, via algorithms, *new solutions* emerging from mathematical formulas expressing what we have already learned how to repeat – is that of *optimization systems*, whose goal is to search for the best

[3]By conflict resolution we mean the ability of *autonomously* deciding, i.e., deciding without a *fixed and a priori* specification of preferences, which action to take among several options, each providing different trade-offs between a given set of conflicting criteria.

[4]A natural suggestion for future works is to investigate the ability of evolving meaningful objective-functions directly from data, perhaps using widespread surrogate models approaches such as Gaussian Processes that are commonly applied for expensive objective-function evaluations, c.f. [167].

[5]In sequential (or multi-period) decision-making, a sequence of decisions must be generated, one for each optimization environment (or decision period) over time.

possible decisions for the problem at hand. Optimization systems are useful when the computational requirements for identifying the best decisions exceed the amount of available human cognitive resources, as it is the case when the number of decision variables increases toward a combinatorial explosion of the number of feasible solutions; or when the rate of decisions to be taken over time is very large.

Such systems must cope with the fact that the decision outcomes are often not known in advance. Predicting the outcomes of decisions in environments subject to statistical fluctuations and time-varying conditions – foreseeing the reaction of customers, suppliers and competitors – can provide important information for decision-making and can be therefore key to reduce the mismatch between goals and reality.

Among the many industries relying on optimization for automating decision-making in a daily basis so as to satisfy a large set of conflicting criteria are airline and cruise companies, which must choose an optimal allocation for their fleet, crew, land personnel, and passengers in order to satisfy the demand and minimize time delays and fuel consumption [11,95]; and product portfolio managers, who must decide how to allocate their budget among many available product development projects [146] and how to schedule the development steps such that the estimated market success is maximal [220].

The list of applications is nearly endless, but the point is that the optimization criteria often resolve into finding a trade-off between minimizing operational costs and maximizing customer satisfaction, what calls for *Multi-Objective Optimization* (MOO) approaches. Presenting a finite set of optimal trade-off decisions found by an MOO algorithm to a Decision Maker (DM) allows for the design of Multiple Criteria Decision Making (MCDM) systems in which the DM must account for multiple conflicting views for evaluating candidate solutions – such as profitability, reliability, safety, responsiveness, etc.

In the MOO/MCDM approach, the DM does not interfere in the optimization process until he/she is required to identify the decision that most complies with his/her preferences among a set of trade-off decisions found by the MOO algorithm. The approach is therefore applicable when either (a) *very little is known* about the problem; or (b) when the current and future preferences of the DM over the optimization criteria are *not known* a priori. In this thesis, we propose advancing this approach towards an autonomous flexible and anticipatory MOO/MCDM framework, so that the need for eliciting preferences from a human decision-maker who has little subsides for assigning importances to conflicting criteria is alleviated.

1.2.4 Research Questions

The questions of *when* and *how* to automate a MCDM process have been nevertheless largely ignored in the literature. Despite what can be speculated about the reasons for the lack of research activity in this area, we argue that there is no fundamental bottleneck preventing such case to be addressed. In fact, we make the points that:

1. The MCDM process should be automated when the actual DM preferences over the conflicting criteria are missing or cannot be determined when required; and

2. A reasonable way of coping with preferences indeterminacy in a sequential MCDM setting

is to implement flexible provisional decisions that are expected to lead to improved sets of trade-off solutions in future decision stages.

Remark: With the proposed *online* MCDM automation strategy, the DM preference specification can be thus *indefinitely postponed* as he/she learns about choice consequences and refines his/her own set of preferences, thus progressively assuming control of the decision process.

One key enabler of such strategy is therefore the search for *flexible* provisional choices that are predicted to lead to high *quality* future admissible trade-off alternative sets. Before giving a precise definition of what is meant by the "quality" of future trade-off sets, we present the concept of *flexible anticipatory engineering*.

1.2.5 Approach: Engineering Flexible Anticipatory Systems

Anticipatory engineering [184] can be regarded as the strategy of designing solutions that enable DMs to respond to future changes. When *simulating* plausible futures, a DM can look after minimizing regret in the worst-case future scenario. In general, the worst-case scenario may never be realized and, thus, an opportunity loss also takes place because the resources invested in the solution could have been allocated to other purposes. Moreover, the strategy of preparing for an specific scenario does not fully address the underlying uncertainty about the future, especially when the DM does not know what are the trade-offs resulting from each possible alternative with regard to multiple conflicting performance measures.

Indecision in choice under uncertainty may in fact result in precipitated actions, yielding unforeseen consequences. However, when there is room for adapting the solutions in operation according to upcoming data, implementing *flexible* provisional alternatives can be critical to *keeping future options open*. Flexible decisions therefore not only mitigate the risks of early commitments to uncertain alternatives, but also may increase adaptability to future changes. In this thesis, the role of flexibility on mitigating preference indeterminacy in online MCDM under uncertainty is thoroughly investigated.

On the Intuitive Role of Flexibility in Decision-Making Under Uncertainty

In effect, we argue that a more *flexible* anticipatory strategy could search for *provisional* solutions, allowing for further adaptations by the time the uncertainty is reduced or revealed. In fact, Hart [108] considered flexibility as "*a law of response to uncertainty*". The basic reasoning is that, if a decision maker cannot guess the future states of Nature which will influence the success prospects of an action that must be taken immediately, flexibility becomes one of the most important aspects of the decision process.

Put in a more simple way, in an uncertain and changing world, the concept of *optimality* may be ill-defined (see e.g. section 3.2 in **chapter 3**). Hence, it might be futile to search for an "optimal" solution. In the strict mathematical sense, an optimal decision is one which is guaranteed to minimize a given cost function. For instance, the optimal decision within the set {*run away, stand still*} when a tree is falling towards an autonomous agent is certainly to run away, which minimizes the potential future hazards to his/her safety arising from the consequences of being hit. There is virtually no doubt regarding the payoffs associated with each available decision in this case: extremely negative in the worst scenario when standing still and almost always positive when running away – except on the scenarios in which the agent would dramatically stumble while executing his/her escaping plan.

It is worth emphasizing that, while the precise damages that would result from being hit are uncertain to the agent – e.g. he/she might be slightly hit in the shoulder and get only a minor cut, or might be hit directly in the head and die instantly –, standing still certainly leads to a worse off situation *by all accounts*, when compared to the safe scenario resulting from running away. Therefore, there is virtually no doubt that running away is worth the effort. It is then clear that for such kind of decisions under severe uncertainty – i.e., under a vast number of scenarios which cannot be all precisely evaluated by an agent –, on-the-fly (online) *approximate reasoning* is required. It turns out that this is one of the main research topics in this thesis.

Nevertheless, while it might be true that an agent does not always need to predict exactly what will happen to act optimally in uncertain environments, sometimes there are no such guarantees. Consider a second scenario in which an agent is lost in the wild and runs into a huge bear that stares right at him/her, and consider the same set of options: {*run away, stand still*}. The question is whether it is still clear what the "optimal" decision is among those two, in terms of minimizing the potential hazards of having the bear attacking the agent. The answer is no in this case. That is because bears can feel threatened by abrupt moves and turn into an attacking mode. Besides, one never knows precisely what are the real odds of a previously unseen bear doing so, or even what are the odds of escaping given that the agent decided to flee.

Hence, while it might be scaring to stand still in front of a wild bear, and while the agent cannot predict whether it will attack, standing still can be seen as *the most flexible decision* among those two. The reason is standing still allows the agent to reassess the hazards as the bear comes closer, *keeping the agent's future set of options open* by postponing the use of the most risky option until when absolutely necessary. That is to say, by standing still, the agent will always have the option of running at a future moment when the hostile intentions of the bear becomes completely clear. With proper planning and enough luck, the agent can escape.

While the agent could have run earlier and improved his/her chances of escaping, the problem with the strategy of running right away is that it is more probable that the bear will simply leave while standing still than it is when trying to escape. So, the payoff of running away is perceived to be more worthy than to stand still only later in the decision-making process, in the rare case of the bear attacking an inoffensive standing human. Of course, after running and turning the bear into attacking mode, suddenly stopping is not a viable option anymore, unless the agent wants to be seriously hurt. Clearly, running right away reduces the agent future *freedom of choice*, while leaving him/her more exposed to the risks of being attacked.

When Increasing Flexibility Might be Worthy

What then distinguishes the first scenario (the falling tree) from the second (the staring bear) so that in the first one an agent can rapidly identify the best action and instantly *react* to the environment, whereas in the second he/she might prefer to postpone the decision of running in favor of keeping his/her future set of options open?

We argue in this thesis that two answers for the aforementioned question are (1) the quality of the available predictive knowledge; and (2) the interaction between actions and the environment. It turns out that, for the falling tree problem, humans have evolved and learned the dynamical model which corresponds to the observed effect of gravity, allowing us to predict that no external force will likely deviate a falling tree from its inevitable trajectory of hitting the ground.

On the other hand, the dynamical model for accurately predicting what an animal as complex as a mother bear protecting her cubs might do in response to the threatening proximity of a human and the optimal plan of action are unfortunately much more complex to be learned and to be properly executed, respectively. In **chapter 7**, section 7.3.7, we investigate how using the *wrong* prediction models impacts the strategy of taking flexible decisions on future *diversity of choice*, while also accounting for the interaction between past actions and current actions.

Evidence Favoring Flexibility and Diversity Maximization

Axiomatic accounts of *preference for flexibility* [136,140] assume that DMs always enjoy enlarging their set of options. A related account of *freedom of choice* [169] regards essential options as those whose exclusion "*would reduce an agent's freedom*". Essential sets were interpreted as those containing the best options "*once future preferences are well determined*" [8], whereas *flexibility* in a two-period setting was defined as "*the size of the remaining second-period action set*" [149].

The concept of keeping track of the future consequences of action on multiple conflicting criteria has also been interpreted in terms of *freedom of choice* in Skulimowski [193], although in a non-axiomatic account. It has been defined as "*(...) the ability to choose an optimal solution from a given set of admissible alternatives with respect to a set of* [future] *selected optimality criteria that are specified explicitly*" [193].

Although flexibility (defined as the cardinality of future admissible option sets) was deemed as useless in deterministic environments or when no future information can be gained over time, Benjaafar et al. [25] developed formal relationships between flexibility, value, information, and risk, and presented theorems asserting that increasing flexibility does not conflict with value maximization and risk minimization. The work of Benjaafar et al. [25] thus indicates that maximizing flexibility may be a rare win-win situation in decision-making under uncertainty.

Kumar [142] generalized flexibility as "*diversity of choice*" and quantified it using entropic measures, where a set of μ options $\mathcal{A}^\mu$ would convey more freedom than $\mathcal{B}^\mu$ if the likelihoods of any $a \in \mathcal{A}^\mu$ being chosen are closer to $\frac{1}{\mu}$ than those of $b \in \mathcal{B}^\mu$.

Another striking evidence that diversity maximization can lead to robust search methods is given in Lehman and Stanley [145] in the paper entitled "Evolvability Is Inevitable: Increasing Evolvability without the Pressure to Adapt". The authors showed experimentally that by following a heuristic to maximize the diversity of candidate solutions in several simulated optimization problems, an observed increase in complexity of the candidate solutions – interpreted as "evolvability" by the authors – was observed, even without specifying any objective-functions to be explicitly optimized.

Finally, as this thesis was being wrapped up, the contemporary breakthrough work of

Wissner-Gross and Freer [221] experimentally showed how intelligent and even cooperative behavior can emerge in dynamical systems evolving to maximize the *causal entropic diversity* of future possible system macrostates. The causal entropic principle [221] thus states that intelligence not only emerges but is actually defined by the strategy of taking the actions predicted to maximize the conditional entropy of the distribution of future states of the world at which a decision maker can be. Thus, by maximizing the number of future achievable macrostates, an intelligent agent would enjoy more future freedom of choice, in a hedge against failure.

The intuitive yet remarkable connection between the proposed strategy of anticipating flexible decisions by maximizing the conditional future Hypervolume indicator in MCDM (see section 1.4) [10] and that of improving future freedom of choice as in Wissner-Gross and Freer [221] can be therefore regarded as one of the main contributions of this thesis.

In fact, we argue that taking actions predicted to yield maximal hypervolume future sets of trade-off options yields an even stronger notion of future freedom of choice in MCDM systems. That is because two future sets conveying the same entropy value, i.e., the same "*diversity of choice*" in Kumar's [142] account, can be distinguished by either comparing their proximity and/or coverage to/of the optimal trade-off surface (see **chapter 3**, section 3.1.1, for a more technical discussion). Therefore, our intuitive claim is that the conditional hypervolume maximization leads to an even stronger notion of future freedom of *essential* trade-off choices.

1.3 Motivating Scenario: Preference Indeterminacy

In *a posteriori* MCDM, a set of trade-off alternatives is usually searched for by a MOO algorithm. The DM does not interfere until he/she is asked to identify the most suitable alternative. Evolutionary MOO [66] (EMOO) may aid MCDM when preferences are undefined, by using the Pareto Dominance (PD, see [229]) for ranking alternatives and approximating the Pareto Frontier (PF) with a set of mutually Pareto-incomparable alternatives that can be presented to the DM. In the decision-making step, the classical approach is to aggregate the decision criteria into a utility function [40]. Preference elicitation is nevertheless prone to severe misspecification. For instance, the DM may not be able to rank every pair of alternatives [21], or the resulting rankings can prevent preferences to be represented with utility functions [216]. Still, the premise in MCDM is that the criteria importances can always be elicited after the DM examines a set of alternatives. This is however unrealistic when it is unclear how the trade-offs between the decision criteria *change over time* and when choice consequences are subject to *stochastic uncertainty.*

Therefore, when preferences cannot be reliably specified due to non-stationarity *and* noise, we propose *modeling* a DM whose goal becomes achieving future improved *flexible and predictable* sets of alternatives to choose from in later stages. In the following, we discuss motivating hypothetical scenarios. The bottom line is that automating MCDM systems makes sense when the DM is unable to determine the relevance of each criterion to his/her decision, in which case the use of *simulation* and *anticipation* can provide such systems with the required flexibility to *operate autonomously.* It is worth noting that this motivation is provided for didactic purposes. The goal is to aid the reader to better comprehend the research reported in the remaining chapters.

1.3.1 Handling Preferences Indeterminacy: Indifference vs. Flexibility

One of the most common pitfalls in sequential decision-making, however, is the temptation to compromise at early stages for immediate rewards: not only such rewards may have a multi-faceted nature, but they can also be embedded into a cloud of uncertainties, making it nearly impossible to determine which decisions will succeed in satisfying longer-term goals. What is worse, such goals are often subject to changes over time and may very well become undefined e.g. during periods of organizational restructuring.

This issue may arise for example when a decision for cutting operational expenses undermines the quality of service to a much greater degree than what might have been expected. This happens because the interplay between those two decision criteria is often subject to uncertainties. For instance, suppose a decision is made to reduce a supplier vehicle fleet and sales personnel in response to a steady reduction in the demand observed over the past few years. Contrary to the expectations, the demand might instead abruptly grow due to an unanticipated competitor bankruptcy event. As a result, it might then become impossible to satisfy all customers requests, what may cause the loss of many selling opportunities. This kind of disruptive scenario may be difficult to predict under imperfect information and, thus, the DM should safeguard the system against uninformed decisions, by carefully anticipating their outcomes according to an extensive set of plausible future scenarios, which is usually carried out by stochastic simulation.

It turns out that cost and quality are always conflicting goals in any system of interest. However, assigning precise importances to each criterion may be too complex to be determined beforehand (i.e., before an optimal decision could be pursued), often because the DM might yet not know in advance how customers will respond to e.g. the delivered quality or to the implemented pricing strategy, especially in the context of an unexplored market niche. Hence, having the ability to *learn from experience* may allow for correcting prediction models. Such corrections may in turn allow for improved simulations of the future consequences of diverse set of trade-off decisions, which can greatly aid the decision-making process.

Example: Flexible Anticipatory Career Choices

In order to get a better picture of how the assessment of alternative trade-off solutions could make the decision process more flexible, consider a DM who is hesitant between an engineering or a business career. The source of indecision may lie in the DM inability to assign importances to conflicting criteria such as average wage, spare time with family, and prospects of promotions. This is a byproduct of *indeterminacy*, which should be differentiated from *indifference*: the former implies that the DM does not have enough information to fully adhere to a goal in the present, whereas the latter simply means that, regardless of how much knowledge about the problem is available, two or more decision paths may end up being equally satisfying.

Shortage of future options may incur from biased decisions taken under little information.

When taking actions under undetermined (or partially determined) preferences, thus, care should be taken to not compromise by acting with indifference (e.g., randomly choosing an optimal trade-off decision), when such indeterminacy will be resolved as all relevant information is revealed over time. This may happen when an early decision favors more one goal over another, making it difficult to backtrack from certain choices. Therefore, instead of committing too early by enrolling in an engineering or in a business major, a DM following our methodology may *anticipate* that enrolling in e.g. an operations research major might *optimally preserve the future options* of pursuing either career after graduating, when he/she will be more experienced and in a better position to make a fully informed choice between those two career paths.

We thus make the case for an automated flexible MCDM strategy that can *postpone preference specification* under uncertainty, while allowing for a more diverse range of options in future decision stages, when the DM may have access to all relevant required knowledge to fully specify his/her preferences. The decision taken is then said to be flexible because it unleashes a larger set of future preferences; otherwise, if the DM compromises earlier for one path and later gravitates towards other, the cost to adapt can be too high.

1.3.2 Handling Uncertainty via Anticipation of Flexible Choices

Handling dynamic *and* stochastic MOO problems is challenging [102] because, while efficient alternatives can be identified in retrospect, their performance for upcoming data is unknown. To the best of our knowledge, despite existing EMOO approaches that can either cope with noisy (e.g. [121, 205]) *or* dynamic conditions (e.g. [135, 226]), research on designing MOO systems capable of *simultaneously* handling stochastic uncertainty *and* temporal changes is virtually nonexistent. As a consequence, there are no EMOO algorithms for *simultaneously* coping with noisy and time-varying objective functions. In a recent 2013 survey article on the modeling aspects of MOO problems subject to stochastic uncertainty in the objective functions, the so called Stochastic MOO Problems (SMOOPs), Gutjahr and Pichler [102] concluded that:

> *(...) the field of multistage SMOO (...) is still widely unexplored, even on the modeling level. A considerable amount of conceptual work, as well as the development of efficient solution algorithms and their test on applied problems, will still be needed to cope successfully with multistage SMOOPs.*

It is therefore evident that, despite its potential expressiveness for modeling complex, real-world problems, research on sequential SMOO has been scarce[6]. One interesting research challenge arising from SMOOPs – one that is set to be addressed in this thesis – is that it appears to be no silver bullet for integrating the 'what might be' part into a problem solver, i.e., there is no simple answer to how the perceived future uncertainty pertaining a decision should alter the evaluation of its worthiness. Therefore, in order to reduce this gap in the literature, we pose the following specific research goal:

[6]In fact, current research on *complex dynamical systems* modeling has been mostly credited for best handling such real-world problems. See **chapter 2**, section 2.3.1 for a more detailed discussion.

> **Research goal:** To deliver mathematical and algorithmic frameworks of flexible online anticipatory learning methods for *automatically* handling uncertainty in sequential SMOO.

The term *anticipatory learning* accounts for methods that can *handle uncertainty* based on predictive knowledge. By making best use of prediction, the hopes are to augment problem solving tools with enough flexibility to effectively cope with multiple conflicting, time-varying, and stochastic optimization criteria, so that adaptable sequences of trade-off decisions can be *anticipated* over time.

Such online anticipatory learning rules are thus intended to incorporate available predictive knowledge (e.g. in the form of predictive distributions) into the optimization tools so they can adapt to expectations regarding the future outcomes of candidate decisions available in the present, mediated by the uncertainty conveyed in the predictive model.

1.4 Research Vision: How It All Comes Together

The vision conveyed in this thesis regarding the research and development of automated decision-making systems relies on the rendering of optimization tools augmented with the following non-exhaustive list of capabilities:

- Conflict and preference indeterminacy resolution;
- Uncertainty/risk awareness;
- Adaptation to change;
- Anticipatory learning; among others.

All those skills are thought to compose the building blocks for rational decision-making and are often object of philosophical debate as well as of intensive research in psychology, biology, and neuroscience; they also constitute the central topics of inquiry in the doctorate research reported hereafter.

We thus propose the so called Anticipatory Stochastic Multi-Objective Optimization (AS-MOO) model [14] that can, in principle, be used to represent sequential SMOOPs with undefined preferences. When solving an AS-MOO problem, we are interested in obtaining a sequence of flexible trade-off decisions that are not inferior to any other decision and, at the same time, preserve the DM preferences indeterminacy over time regarding the stated optimization criteria. In order to improve the diversity of future set of options, we propose maximizing *set-based quality indicators* (see [227]) over time, among which we argue that the Hypervolume (Hypv, $\mathcal{S}$) [10] is consistent with our notion of *flexibility* (see **chapter 3**).

The property of foremost interest to this thesis is that the maximization of the $\mathcal{S}$-Metric measured over N mutually non-dominated solutions is guaranteed to yield a subset of N solutions in the optimal trade-off surface known as the *Pareto Frontier* (PF, see **chapter 3**) [10]. It is this guarantee that allows us to express the AS-MOO model as a recursive equation in which the optimal trade-off solutions at any decision period depend on the optimal solutions at later stages. Moreover, the resulting set of trade-off solutions is expected to provide good near-optimal approximation of the PF [41]. When solving real-world AS-MOO problems, however, high levels of uncertainty about the future decision stages can undermine the search for exact solutions.

It is arguable in statistical decision theory and, particularly, in Bayesian inference [172], that the quantified uncertainty can be used for updating predictions and other estimates based on accumulated historical evidence. The proposed AS-MOO models are therefore *approximately solved* by a hypervolume maximizer, the Anticipatory S-Metric Selection Evolutionary Multi-Objective Algorithm (ASMS-EMOA) [14].

We thus propose principled ways of augmenting the anticipatory capabilities of multi-objective solvers (metaheuristics) with Bayesian models to aid sequential MCDM in uncertain environments. The ASMS-EMOA proposal makes use of the Kalman Filter (KF) estimation [100] (see **chapter 2**, section 2.3.4) and of *Online Anticipatory Learning* (OAL, see **chapter 6**), for updating the performance vector (section 6.1) and the decision vector (section 6.2) of each candidate solution.

Each trade-off solution is ranked according to its expected anticipatory $\mathcal{S}$-Metric contribution, $\mathbb{E}\left[\Delta_{\mathcal{S}}\right]$ (see **chapter 6**). The pairwise covariance terms appearing in the computation of $\mathbb{E}\left[\Delta_{\mathcal{S}}\right]$ (see **chapter 6**, Theorem 6.3.1) merge trade-off information between multiple neighboring anticipatory distributions in the objective space, thus allowing for the *mutual interference* and *exchange of trade-off information* between parallel neighboring anticipatory paths.

Nevertheless, the AS-MOO model can hardly be exactly solved because, first and foremost, the underlying probability distributions are unknown and must be estimated from the available data, what requires the specification of models (such as the Gaussian assumption made in the KF estimation) that, while contributing to make the problem tractable, are often subject to information loss and inaccuracies[7]. Another reason for approximately solving the AS-MOO model is due to the fact that it is expressed as a recursive equation whose exact solution would require an exponential number of re-optimizations. That is why we argue that efficient approximated AS-MOO solution methods are the most suited for real-world practical applications, what ended up being the focus of this thesis.

1.4.1 On Anticipatory Multiple Criteria Decision-Making

The layer diagram shown in Fig. 1.1 provides a broader view of the proposed anticipatory MCDM architecture. The three heuristic meta-preferences described in section 1.2.2 are modeled by the proposed online anticipatory learning method. Uncertainty in the specification of the DM lower-level preferences for ranking and assigning importances to the underlying objective-functions prompts the definition of intermediate-level meta-preferences, such as preference for flexibility and preference for long-term predictability. By its turn, the uncertainty regarding those meta-preferences are set to be coordinated by higher-level hyper-preferences heuristics, such as Online Anticipatory Learning (OAL), which actuates directly over the uncertainty estimated from anticipatory Pareto order information among candidate trade-off decisions (see **chapter 6**).

[7]See **chapter 2**, section 2.1.3 for a more in-depth discussion.

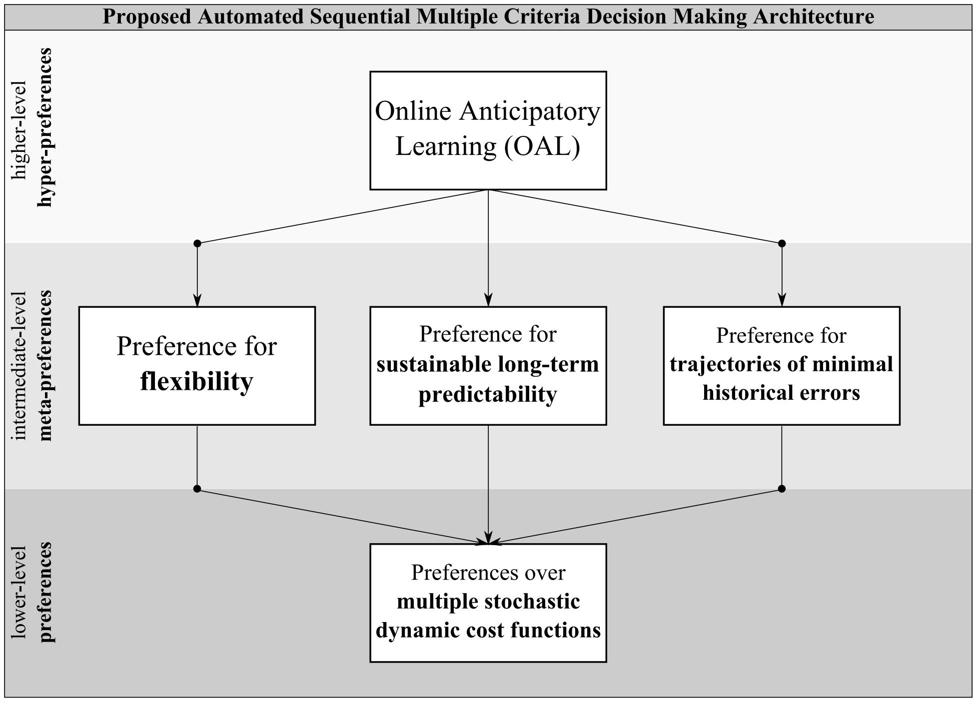


Figure 1.1: Diagram with the layered architecture proposed in this thesis.

In this architecture, the uncertainty of lower preference layers is propagated and handled by subsequent higher-level layers.

The closest related MCDM theoretical model that we found in the literature was that of Skulimowski [192], whose goal was:

> "*To use anticipated future consequences of a decision as a source of additional preference information in multivariate decision problems*."

That framework has been further developed and culminated in the recently proposal of the so-called "anticipatory networks" [194], in which "*(...) non-cooperative future decision makers influence the same decision unit"*.

It should be noted that our anticipatory proposals for automating sequential MCDM processes were developed without any influence of those past works and that we only became aware of their existence when the thesis was being wrapped up. Besides, although those early models have identified the key role of *prediction of future consequences of actions* in MCDM processes, they were not necessarily intended to *automate* the MCDM process.

Nevertheless, we acknowledge that Skulimowski [192] provided impressively vanguard and advanced models for anticipatory MCDM at the early 1980's, when the computational and engineering resources to design and experiment with algorithmic solutions were unfortunately most likely prohibitive at the time.

Therefore, we believe that the flexible anticipatory MCDM models as well as the algorithmic and computational contributions provided in this thesis also represent a first step on making Skulimowski's [192–194] theoretical contributions implementable.

1.4.2 Research Roadmap and Main Contributions

The thesis main contributions are thus as follows (see Fig. 1.2):

Contribution 1 An open-ended research problem: how can a DM behave over time when preferences cannot be reliably elicited?

Contribution 2 An answer to that question motivated in decision theory: the DM may exhibit preference (a) for flexibility; (b) for long-term predictability; and (c) for decisions whose estimated outcomes possess minimal historical error (see **chapter 6**);

Contribution 3 An online decision-making model: select a (flexible) Pareto-efficient alternative predicted to maximize future Hypervolume;

Contribution 4 A flexible anticipatory MCDM methodology: solve an online sequence of Anticipatory Stochastic MOO (AS-MOO) problems over time;

- An Anticipatory $\mathcal{S}$-Metric Selection EMOO algorithm (ASMS-EMOA) for solving AS-MOO problems;
- *Online Anticipatory Learning* rules incorporating Bayesian tracking into ASMS-EMOA for self-adjusting time preferences based on the estimated future uncertainty and on historical prediction errors; and

Contribution 5 An experimental validation of the proposed online anticipatory MCDM methodology in financial portfolio selection.

1.4.3 Potential Real-World Applications

There are several real-world applications that can benefit from AS-MOO, including:

- Air traffic tactical planning – to obtain a set of non-inferior daily flight plans minimizing the probability of congestion within flight sectors while minimizing the total expected delay to takeoff [51];
- Vendor-managed inventory replenishment – to obtain a set of non-inferior daily replenishment and vehicle routing plans minimizing transportation and inventory costs while minimizing the probability of product shortage [132];
- Portfolio selection – to obtain a set of non-inferior daily wealth allocations maximizing the expected return over investment while minimizing the expected risk [2]; etc.

All those problems rely on stochastic simulation, in which scenarios are sampled from a model describing uncertain quantities such as time to arrival, demand, and return. Besides, while one can search in retrospect for decisions that would be optimal for historical, observed data, there are no guarantees that such decisions will remain optimal when adopted in future environments. Hence, the role of anticipation is crucial for those problems, especially when

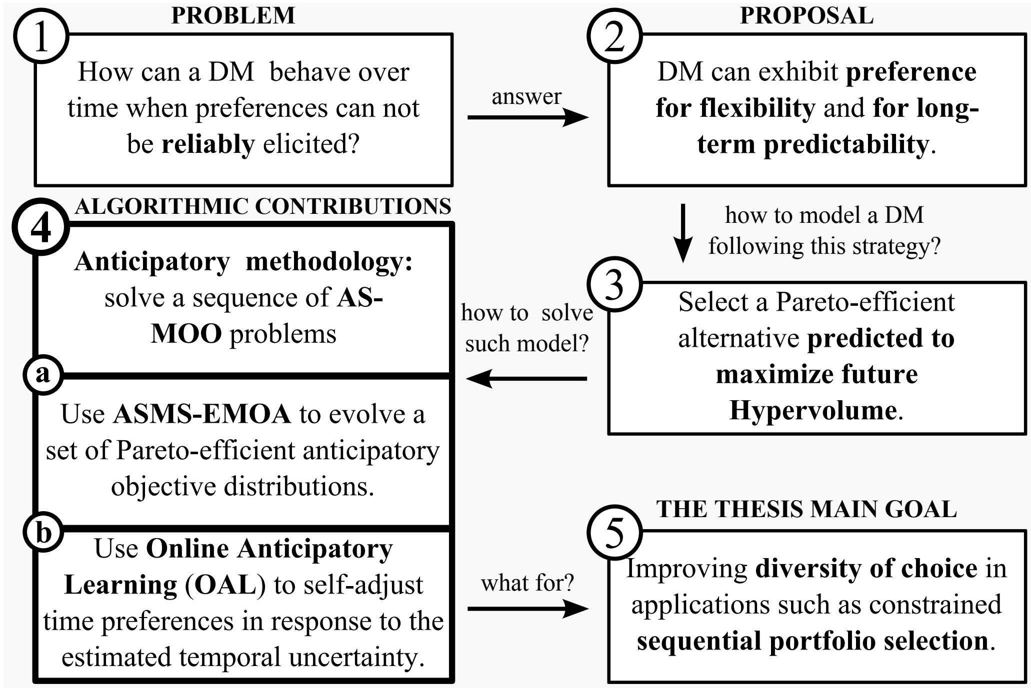


Figure 1.2: Research roadmap showing the goal and contributions of the thesis.

there are temporal dependencies (time-linkage) between the decisions taken and the upcoming future optimization environments.

It is natural to assess our methodology and algorithmic contributions in portfolio selection because, besides its real-world interest, financial markets invite the handling of strong sources of uncertainty. Active asset allocation implies that the implemented portfolio should be monitored and rebalanced to reflect market changes. Markowitz's modern portfolio theory [151] provides only a backward-looking (myopic) strategy, in the sense that the analysis suggests outdated portfolios upon market changes. The tracking of changing portfolios has been recently described as a promising future research avenue in Kolm et al. [134], who suggested that all "*trade-offs between risk-adjusted returns and trading costs*" should be considered. As put by the 2003 economic Nobel laureate, Robert Engle, "(...) *the risk of a portfolio depends not on what the correlations were in the past, but on what they will be in the future*" [80] and, thus, "(...) *we must 'anticipate correlations' if we want to have optimal risk management, portfolio selection, and hedging*" [80].

Our experiments are thus designed to determine the degree to which anticipating predicted flexible trade-off alternatives in ASMS-EMOA can improve the quality of the obtained sets of candidate non-inferior decisions in out-of-sample environments, as measured by the Hypervolume indicator. We investigate the performance of ASMS-EMOA and of its myopic counterpart (SMS-EMOA [32]) on artificial scenarios of portfolio selection, for which we have designed a dynamic instance generator, controlling for periodicity and for severity of disruptive change, representing different regime switching scenarios, that, in financial markets, may result from periods of crisis or abrupt changes in governmental policies. In addition, we present results for real-world data from three stock indexes: London's FTSE-100, Dow Jones Index, and Hong Kong's HSI.

1.4.4 Applications for Optimization within Simplex Spaces

In this thesis, it is assumed that the feasible search space is defined over the $(d-1)$-simplex $S^{d-1} \subset \mathcal{R}^d$, where d is the number of decision variables to be optimized (see **chapter 2**, section 2.3.5, Eq. (2.34)). We have thus designed specific anticipatory methods for handling this setting, motivated by our choice of assessing the proposed anticipatory methods in the portfolio optimization domain (see [14, 15] and **chapter 7**), although optimizing over the simplex has many other practical applications in machine learning, bioinformatics, social network analysis, and computational chemistry.

For instance, Motzkin and Straus [157] highlighted the remarkable connection between the problem of maximizing the Lagrangian of a graph G over the simplex and the clique number of G. It turns out the the maximum clique problem in graphs is of interest to many real-world problems. Tang et al. [203] and Bulò and Pelillo [176] generalized the Motzkin-Straus theorem for hypergraphs, and the latter proved that the problem of *hypergraph clustering* – defined as "*(...) the process of extracting maximally coherent groups from a set of objects using high-order (rather than pairwise) similarities*" [176] – can be posed as one of finding the equilibria of a cooperative game expressed as an optimization of a polynomial function over the simplex.

Other applications that can benefit from finding maximal cliques by maximizing the Lagrangian in hypergraphs include the computational analysis of high-throughput gene expression and transcriptomic data [77], and the prediction of biologically active compounds for drug discovery and chemical genomics [49].

Bunea et al. [43] have studied the problem of the optimal (minimal loss) convex aggregation of (Gaussian) statistical models for regression, in which the aggregation function Q is expressed in terms of the mapping $Q : \mathcal{X} \times S^{d-1} \mapsto \mathcal{R}$, where $\mathcal{X}$ is the sample space of the individual statistical models. Because the true probability distributions governing the data are unknown, they modeled the problem as a stochastic optimization program.

Mørup [156] has provided an exact, continuous optimization relaxation in the simplex for handling the combinatorial explosion of hard assignments in clustering problems, in which the vertices of the simplex are shown to lead to stable hard assignment solutions. Examples of a simplex-based clustering algorithm are provided for community detection in complex networks.

Finally, Kotropoulos and Moschou [138] addressed the problem of reducing the number of classes in speech classification by applying clustering techniques over simplex subspaces, for which feature vectors describing probability assignments from samples to classes are available. The application considered the probabilistic re-assignment of neutral speech features into more informative classes describing emotional states.

Before addressing the technical aspects of how anticipation is implemented in our models and problem-solving tools to handle uncertainty, we pave the way for a better understanding of the science of anticipation, from its biological basis to how it has been studied theoretically. In fact, this concept has influenced the scientific thought in many different areas, but has yet to be elevated to a discipline on its own.

1.5 Biological Basis of Anticipatory Systems

One of the first serious accounts for the emergence of anticipation in biological systems is due to Rosen [174], who defined an anticipatory system as:

> *"(...) a system containing a predictive model of itself and/or of its environment that allows it to change state at an instant in accord with the model's predictions pertaining to a later instant."*

At this point, it is worth making a distinction between prediction and anticipation. The former is merely a representation of a future event or outcome, whereas the latter is a future-oriented decision, based on prediction. Butz et al. [47] described anticipation as:

> *(...) a process or behavior that does not only depends on past and present but also on predictions, expectations, or beliefs about the future.*

Among the many examples of biological anticipation found in Rosen [174], we highlight the case of negatively phototropic organisms, i.e., organisms that are attracted to darkness. Rosen argues that while darkness in itself is physiologically neutral, it is *correlated* with non-neutral environmental features such as the absence of sighted predators. Effectively, such organisms would act in accordance with a built-in model that has somehow learned the correlation between moving towards darkness and higher chances of survival [174]. In this thesis, we argue for this particular viewpoint of anticipation as being triggered by learned correlations.

1.5.1 Signaling Theory

Meaningful correlations triggering anticipatory actions permeate the field of *signaling theory*. For instance, Thomas and Stoddart [206] made the case that some tree species react to the shortening of daylight in autumn by changing color and shedding their leafs in *anticipation* to the high physiological costs of maintaining them during winter, what otherwise could supersede the benefits from photosynthesis. Hamilton [105] suggested that "*autumn coloration serves as an honest signal of defensive commitment against autumn colonizing insect pests*".

Signaling theory plays an important role for explaining the co-evolutionary dynamics between different species. Signals would evolve to influence the behavior of the receiver to benefit the signaler. While those signals can be honest or dishonest[8], they may in effect *anticipate* a counteraction to mitigate or nullify the potential future actions of the receiver as predicted by the signaler. Models of mate selection are heavily influenced by this theory. For example, experiments with the mating dynamics of the jumping spider *Phidippus clarus* species in Sivalinghem et al. [191] showed that signaling rates "*significantly predicted mating success in all copulations*". Males of other species such as *Habronattus dossenus* that made strong use of courtship signals were also found to suffer lower pre-mating cannibalism rates.

In an application of signaling theory to explain the dynamics of the job-market, Spence [197] defined hiring as an "*investment under uncertainty*" in which applicants signal their skills to

[8]The adaptive dynamics of the proportions of individuals within species relying on honest and dishonest strategies are commonly predicted by evolutionary invasion analyzes.

the employer by showcasing their credentials. The informational value of such signals to the contractor emerges when he/she analyzes the benefits of hiring the applicants for which they predict positive correlation between the credentials shown and superior work performance.

1.5.2 Hormonal Regulation and Sensorimotor Control

In vertebrate species, anticipatory hormonal regulations can be triggered by the autonomic nervous system in response to stressful stimuli followed by the perception of endangerment. For instance, subconscious processes triggered by acute stress may release hormones such as epinephrine (adrenaline) into the bloodstream causing a diverse range of physiological responses in different organs, including increased heart and respiratory rates, increased blood vessels diameter, and muscle contraction.

Because such anticipatory responses prepare the body to respond to a potential future (harmful) event, not only they are thought to be crucial to triggering fight-or-flight-or-freeze [48] behavior, but also they play an essential role in sensorimotor control. For instance, muscle contractions can compensate eventual disturbances in the body's balance, in anticipation of the potential harms of falling to the ground. Several experiments regarding the motor system dynamics have supported the existence of an *internal model* accountable for estimating the relative location of body members, what is essential when sensorial feedback is not available, as it is the case e.g. when walking in a dark room [222].

1.5.3 Anticipation as a Built in Feature of the Brain

It seems to us that the concept of a self-referential conscious brain fits nicely into Rosen's definition of anticipatory systems. For instance, the acclaimed evolutionary biologist Richard Dawkins presented an interesting explanatory narrative of how the capacity of simulating the future may have evolved to enable animals to benefit from *vicarious experience*[9] [64]:

> *"Natural selection built in the capacity to simulate the world as it is because this was necessary in order to perceive the world. You cannot see that two-dimensional patterns of lines on two retinas amount to a single solid cube unless you simulate, in your brain, a model of the cube. Having built in the capacity to simulate models of things as they are, natural selection found that it was but a short step to simulate things as they are not quite yet-to simulate the future. This turned out to have valuable consequences, for it enabled animals to benefit from 'experience', not trial-and-error experience in their own past (...), but vicarious experience in the safe interior of the skull. And once natural selection had built brains capable of simulating slight departures from reality into the imagined future, a further capacity automatically flowered. Now it was but another short step to the wilder reaches of imagination revealed in dreams and in art, an escape from mundane reality that has no obvious limits."*

[9] We understand *"vicarious experiences"* as those gained after generating and processing the feelings arising from the imagined (or perceived) consequences of potential actions upon oneself (or someone else's self).

One of the most prominent evidences that (primates) brains indeed encode mechanisms supporting simulation and vicarious experiences was the discovery of the so-called *mirror neurons* in the premotor cortex [171]. They are regarded as a distinctive class of neurons because they exhibit activity both when a motor task is executed and when the subject has access to a visual feedback of another subject performing the same motor task.

Barbey et al. [20] suggested that counter-factual thinking about the past (e.g. "Had I studied harder, I would have passed the exam.") for improving future performance or simulating future behavior (e.g. "What would I do if X happens in the future?") depend on structures in the prefrontal cortex. From functional magnetic resonance imaging (fMRI) experimental analyzes, Van Hoeck et al. [212] concluded that counter-factual and standard episodic thinking "*(...) share a common brain network, involving a core memory network (hippocampal area, temporal lobes, midline, and lateral parietal lobes) and prefrontal areas that might be related to mentalizing (medial prefrontal cortex) and performance monitoring (right prefrontal cortex)*".

1.5.4 Game Theoretic Analyzes of Anticipation

It turns out that anticipation accordingly to feelings and expectations about the future can lead to a form of *personal equilibria* in a game theoretic framework. Kószegi [137] conceived a model of sequential decision-making wherein the utility function of a DM is elicited not only by physical outcomes but also from their own expectations over future outcomes. He then argued that those two payoff components – future expectations and present actions – can interact and jointly lead to personal equilibrium, subject to the requirement that the current action conditioned on past expectations must attain a certain degree of consistency.

Kószegi's [137] model thus predicts equilibrium points resulting from the positive feedback between expectations and behavior, what resembles Merton's notion of *self-fulfilling prophecies* [154]. This viewpoint may explain to a certain extent human behavior and emotions arising from the degree to which the DM is consistent over time. For instance, the willingness of a pessimistic DM to keep doing a riskless, low rewarding action can be higher if he/she had been wrongly believing that all other available options would also yield low rewards. Because of this prior belief, he/she may also not be willing to assume the cost of acquiring more information about alternative options – whose rewards could actually have been greatly improved since when the DM formulated his/her beliefs. Hence, because of the time consistency constraint, such DM will end up materializing his/her own prior expectations of acquiring low rewards.

1.6 Meaning and Purpose of Anticipation in the Thesis

Much has been said throughout this chapter about the role of anticipation in explaining natural phenomena or biological/sociological behavior, as well as in engineering flexible and robust systems. Nonetheless, the meaning of "anticipation" used in this thesis is still non-orthodox with regard to its general understanding. That is because we are following the broad notion conveyed in Rosen's definition [174] that *any* decision or action mediated by prediction *is* anticipation. The daily common sense understanding, on the other hand, is more restrictive and almost always only refers to decisions or actions that are brought forward *ahead of time* (in advance), i.e., actions or decisions that are taken at an *earlier* time than they normally would without further consideration.

Rosen [174] made no such distinction about the *timing* of specific decisions when defining anticipation, nor do we do that in this thesis. It follows from this broader perspective that specific decisions or actions that are *postponed* (i.e., that are set to be taken at a later time than they normally would without further consideration) are also anticipations, since such postponements are justified by predictive knowledge. It turns out that the *sequence* of decisions/actions taken mediated by prediction will look much different from other sequences produced under different assumptions and expectations about the future (see section 3.3.1 of **chapter 3**).

The *purpose* of anticipation in this thesis should also be elucidated and distinguished from the common sense. We claim that there are at least three reasons that makes anticipation in decision-making absolutely necessary:

1. *Value maximization* – anticipation as a means of *capitalizing on* and taking advantage of foreseen opportunities. Consider for instance the negotiation of a contract renewal. The DM can either decide to approach the other party ahead of time (e.g. several months or years before the contract expiration date) or even at a later time (i.e., after the contract expiration) *because* he/she foresees a favorable environment to bargain better conditions;

2. *Risk mitigation* – anticipation as a means of avoiding unnecessary risks and mitigating inevitable ones. The DM can decide to take immediate action to completely eliminate or reduce the likelihood and impact of a foreseen undesired outcome, including investing time and effort on building contingency plans so he/she is prepared to mitigate the outcome shall it occur. Sometimes, however, the DM can realize that the undesirable outcome is acceptable to make no changes to his/her plan or strategy and just hope that it does not materialize. Consider for instance the payment of e.g. municipal taxes. The DM may decide to pay his/her taxes in advance *because* he/she foresees to take advantage of promotional discounts and also to avoid the risk of being penalized with the high interests rates and fees for forgetting to do so within the imposed deadline;

3. *Uncertainty handling* – anticipation as a means of handling the uncertainty arising from lack of information (indeterminacy). The DM can *postpone* a decision *because* he/she foresees to be in a better position to take an informed decision after investing in (often costly) information acquisition, therefore reducing his/her uncertainty and the possibility of taking unnecessary risks and/or underperforming decisions. Such DM thus may search for robust intermediate actions that can account for the current uncertainty levels without compromising into the unknown. This is the scenario in which anticipation will be essential in this thesis for sequential MCDM under uncertainty.

Another concept that should be highlighted in our understanding of anticipatory systems is the presence of *internal models*. This is an important aspect to autonomous behavior because it is what allows such systems to construct their own beliefs about the environment and about the consequences of their actions. Such models thus enable the system to process their own uncertainty and to modify their own behavior accordingly. For instance, we argue that the capability of relying on internal models to predict and to estimate the underlying uncertainty about the dynamics of the optimal trade-off solutions in the search space and their corresponding evaluations in the objective space is what distinguishes our proposed anticipatory problem-solving tools from existing optimization tools passively relying on *external* prediction models (e.g. [89]).

It should be made clear that we are not suggesting to know how anticipatory systems should be engineered. Actually, we do not even know whether the anticipatory systems perspective possess what it takes to completely explain and handle complex dynamic systems. We admit, though, that this particular understanding of anticipation and of anticipatory systems has strongly influenced our models and problem solving tools for sequential MCDM under uncertainty and, thus, anticipation plays a very definite role in our proposals presented in section 5.2.

1.7 Thesis Organization

The remaining contents of the thesis are organized as follow:

Chapter 2 An overall introduction to sources of uncertainty in decision-making, as well as of models for representing it, is presented, including Shannon's entropy, which is one of the main concepts permeating the algorithmic contributions of **chapter 6**. Besides, the Bayesian tracking methods utilized in this thesis for estimating and predicting time-varying distributions in both the objective and the search space are discussed. Those tools are the Kalman Filter and a recursive maximum a posteriori for tracking Dirichlet distributions in simplex spaces.

Chapter 3 The technical background for MOO and sequential MCDM is presented, including the Hypervolume indicator and its not surprising – yet incomprehensibly neglected – axiomatic relation with the concepts of preference for flexibility and diversity of choice appearing in the economics and decision-theoretic literature. The chapter thus provides the foundations for the thesis main algorithmic and modeling contributions.

Chapter 4 The application scenario of financial portfolio optimization is described and the classical Markowitz [151] mean-variance problem is detailed. In addition, a preemptive multi-objective strategy for improving the stability of portfolios in changing (albeit deterministic) environments is suggested. In other words, regularization of the hypervolume indicator is utilized as a preemptive strategy to account for unpredictable changes, in the absence of a predictive model.

Chapter 5 The thesis contributions on the modeling of flexibile and anticipatory behavior are described, including an anticipatory stochastic MOO model, and an anticipatory flexible methodology for improving future diversity of choice.

Chapter 6 The thesis main algorithmic contributions are described, including an anticipatory MOO hypervolume-based maximization algorithm, an anticipatory learning strategy for self-adjusting meta-preferences in terms of the entropy of the probability of Pareto dominance between two temporal performance distributions, and in terms of historical estimation errors, and the computation of the expected anticipatory hypervolume contribution of a candidate trade-off solution in the objective space, under a joint Gaussian assumption.

Chapter 7 The application of the proposed anticipatory methodology for portfolio selection subject to expected return maximization and expected risk minimization is presented. The experimental design and the results achieved are discussed in details and confronted with the thesis main goals, considering controlled experimental scenarios with nonstationary stochastic processes and real-world financial data.

Chapter 8 Conclusions, limitations, and suggestions of future research avenues supported by our anticipatory MOO approach are presented.

1.8 Publications Related to the Thesis

As of September 9, 2014, the list of research papers published by Carlos R. B. Azevedo during his time as a doctoral student at the Laboratory of Bioinformatics and Bio-inspired Computing (LBiC), UNICAMP, are listed in Table 1.1. Three papers contain material that appears in the thesis and/or that is directly derived from the results of the thesis. Other papers have been published on related research initiatives, but were not directly derived from the thesis contributions. Other journal submissions are under preparation for publishing additional results that are not contained in the thesis, but that were generated while the thesis was being wrapped up.

Table 1.1: List of papers published by Carlos R. B. Azevedo during the doctoral project.

Paper reference	Summary	In thesis?	Chapters
C. R. B. Azevedo and F. J. Von Zuben. Learning to Anticipate Flexible Choices in Multiple Criteria Decision-Making Under Uncertainty. IEEE Trans. On Cybernetics, (under review), 2014.	Presents the main results from the application of the *time-linkage* AS-MOO methodology to portfolio selection.	**yes**	1–3, 5–8
C. R. B. Azevedo and F. J. Von Zuben. Anticipatory stochastic multi-objective optimization for uncertainty handling in portfolio selection. In Evolutionary Computation (CEC), 2013 IEEE Congress on, pages 157–164, June 2013	Presents results from the application of the *time-linkage free* AS-MOO methodology to portfolio selection.	**yes**	1–3, 5, 6
C. R. B. Azevedo and F. J. Von Zuben. Regularized hypervolume selection for robust portfolio optimization in dynamic environments. In Evolutionary Computation (CEC), 2013 IEEE Congress on, pages 2146–2153, June 2013.	Presents results from the application of a preemptive multi-objective metaheuristic to portfolio selection.	**yes**	2–4
C. Aranha, **C. R. B. Azevedo**, and H. Iba. Money in trees: How memes, trees, and isolation can optimize financial portfolios. Inf. Sci., 182(1):184–198, 2012.	Presents results from the application of population-based Nature-inspired metaheuristics for dynamic portfolio selection.	no	4
C. Aranha, **C. R. B. Azevedo**, V. S. Gordon, H. Iba. Topological and Motion Strategies for Cellular Memetic Tree-Based Portfolio Optimization, X Brazilian Congress on Computational Intelligence, November 2011.	Presents results from the application of Nature-inspired metaheuristics for dynamic portfolio selection.	no	4
C. R. B. Azevedo and A. F. R. Araújo. Nondominance landscapes for guiding diversity generation in dynamic environments. In VIII Best MSc Dissertation/PhD Thesis Contest in Artificial Intelligence (CTDIA). Proceedings of The Brazilian Conference on Intelligent System, 2012.	Presents results from the application of preemptive diversity-based metaheuristics to dynamic multi-objective optimization benchmark problems.	no	3
A. R. Gonçalves, R. Veroneze, S. S. Madeiro, **C. R. B. Azevedo**, F. J. Von Zuben. The Influence of Supervised Clustering for RBFNN Centers Definition: A Comparative Study. In: International Conference on Artificial Neural Networks (ICANN), 2012, Lausana. Lecture Notes in Computer Science: ICANN 2012 - Part II. Berlin: Springer, 2012. v. 7553. p. 148–155.	Presents results from the study of the influence of using label information for data clustering algorithms in pattern classification benchmark problems.	no	n/a

Chapter 2

Representing, Measuring, and Handling Uncertainty in Decision-Making

Not only does God play dice, but sometimes he throws the dice where we can't see them.

– Stephen Hawking

When one admits that nothing is certain one must, I think, also add that some things are more nearly certain than others.

– Bertrand Russell

This chapter discusses statistical tools for modeling and processing stochastic uncertainty in decision-making, including Bayesian methods for estimation and prediction of time-varying uncertain quantities in discrete and continuous spaces. The algorithmic innovations proposed in this thesis have benefited from the techniques addressed in the following.

2.1 Sources of Uncertainty

A reliable decision-making system requires a means to *understand* its surrounding environment as well as the consequences of its potential decisions. While, conceptually, understanding is not the same as modeling, it is convenient to conceive rough descriptions of the object of interest, known as *models*, to aid understanding. The act of modeling conveys reduction techniques, grounded on the Scientific Method, that amounts to (1) discarding what is irrelevant; and (2) representing to some extent what is relevant to explain the behavior of an object of interest. Models can thus be useful for reducing the gap in a system's knowledge by exploring the simulated behavior of whatever we are trying to gain a better understanding.

Such descriptions are of course incomplete by construction and a certain amount of *uncertainty* on the system's understanding about what is being modeled is inevitable. The resulting granularity loss undermines the system's knowledge about its environment, but it is not the only source of uncertainty the system has to deal with. In fact, Parsons [165] discussed several taxonomies about types of uncertainty proposed in the literature and agreed with Elkan's [78] conclusion, stated in 1994:

> *"A complete and consistent analysis of all the many varieties of uncertainty involved in human thinking and revealed in human language is a philosophical goal that we should not expect to achieve soon."*

After 20 years, while it still appears to be no evidence that such philosophical achievement is near, there are a few common sense[1] notions of what uncertainty means, including:

- Lack of precision arising from imperfect measurement;
- Lack of reliability arising from the passage of time;
- Lack of confidence arising when trying to abstract complex phenomena; or
- Lack of definiteness arising from gaps in the system's knowledge.

The four aforementioned notions are indirectly addressed in our proposal for sequential (multi-period) Multi-Criteria Decision-Making (MCDM) under uncertainty (see **chapter 5** and **chapter 6**). Some of them need to be handled during the design of sequential SMOO problem-solvers, while others require a better modeling of the optimization problem, as will be discussed in the following.

2.1.1 Uncertainty Arising from Imperfect Measurement

One fundamental source of uncertainty in the real-world comes from measurements. Although through mathematics we have an expressive alphabet such as the set of real numbers to represent any measurable quantity up to arbitrary precision, our physical measuring tools and mechanisms are still limited. Werner Heisenberg's famous *uncertainty principle* [111] indeed showed that there are, at the quantum scale, physical limits to the precision one can measure a system of two conjugate variables, such as the position and momentum of a subatomic particle.

Taylor [204] goes as far as to speak of "*inevitable uncertainty*" and gives a dramatic example of how a carpenter could never be satisfied with his measurement of the height of a doorway before installing a door. First and foremost, his measuring tape may be insufficiently or poorly graduated (either because the marks are too far away or due to failures of the printing device). Moreover, the carpenter may have to approximate the measurement if the top of the doorway is not aligned with one of the marks. Not only that, but he/she may also find later that the measured height differs at different locations, and even that the heights vary with time, depending on the temperature and humidity conditions at the room. Granted, the carpenter frustration illustrates several aspects of uncertainty arising from measurement, even at one of the most controllable and deterministic environments one can think of. Experimental evidences also suggest us that there might be external sources of interference in our measurements that we might not even be *aware* of.

[1] Here we apply the criterion of what can be found in mainstream English language dictionaries.

Imperfect Measurement in Financial Markets

The lack of awareness is a topic at the heart of the controversy regarding the efficiency of financial markets. The efficient market hypothesis [180, 181] suggests that, unless an investor has access to inside information that other investors are not aware of (e.g. privileged access to quarterly earning reports, or intended acquisitions, etc.), he/she cannot tell with certainty if assets prices are too low, too high, or just right. This implies that, on average, "*(...) it is not possible to outperform the market consistently on a risk-adjusted basis after accounting for transaction costs by using available information.*" [83]. This impossibility comes from the belief that every measure required for taking investment decisions are completely reflected in asset prices [83], which, according to Fox [87], has lead many key characters of the global economy to believe that markets were rational:

> *"Financial markets knew best. They moved capital from those who had it to those who needed it. They spread risk. They gathered and dispersed information. They regulated global economic affairs with a swiftness and decisiveness that governments couldn't match. (...) Financial markets possessed a wisdom that individuals, companies, and governments did not."*

Alan Greenspan, the president of the U.S. Federal Reserve at the time, when testimoning at the House of Representatives in the midst of the financial crisis of 2008, blamed the misuse of the available data for the chaotic scenario that had been then unfolding [99]:

> "*The whole intellectual edifice collapsed in the summer of last year, however, because [of] the data inputted into the risk management models (...)*"

and admitted that there was no such thing as a rational market: "*it was the failure to properly price such risky assets that precipitated the crisis.*" [99]. Following Greenspan's logic, it seems that the decisions taken under uncertain and imperfect data can have a huge impact on the real-world, implying dangerous feedback loops that may be difficult to mitigate with confidence.

Modeling Measurement Uncertainty

Among the many mathematical frameworks used to represent uncertainty, including possibility theory [224] (from which fuzzy theory can be derived) and Dempster-Shafer theory of evidence [72, 186], arguably probability theory is the most widespread. We shall not discuss the merits of using probabilistic models for designing intelligent systems nor we will enter in the details of the other frameworks, but we refer to Cheeseman's essay [54], which generated many responses from prominent researchers working on uncertainty models, including Judea Pearl and Shafer himself. What is interesting to note, however, are the connections all alternative frameworks possess to probability theory, suggesting that it may be difficult to not rely on likelihoods and observed frequencies of occurrence of events for quantifying uncertainty.

Dubois and Prade [76], for instance, devised a possibilistic version of Bayes' rule, used for reversing the direction of reasoning in the same way as its probabilistic counterpart. Zadeh [224] has more clearly exposed the connections between possibility and probability using the following example (also found in Parsons [165]): consider the statement "*Hans ate x eggs for breakfast*". One can model the uncertainty regarding the value of x by attributing a set of possibilistic outcomes based on the ease with which Hans can afford eating x eggs, say, $\{1, \cdots, 8\}$, and then, one could assign a probability distribution over those possible outcomes. Although one might conclude that it would be, in principle, possible for Hans to eat 8 eggs (thus assigning a value of 0.2), in a probabilistic sense, the subjective attribution of the likelihood for him to have such an outrageous breakfast would be much closer to zero.

When the determination of the value of a quantity x is not merely speculative but comes from the results of measurements, one simple probabilistic model to represent such uncertainty is that of an additive noise:

$$\hat{x} = x + \epsilon, \tag{2.1}$$

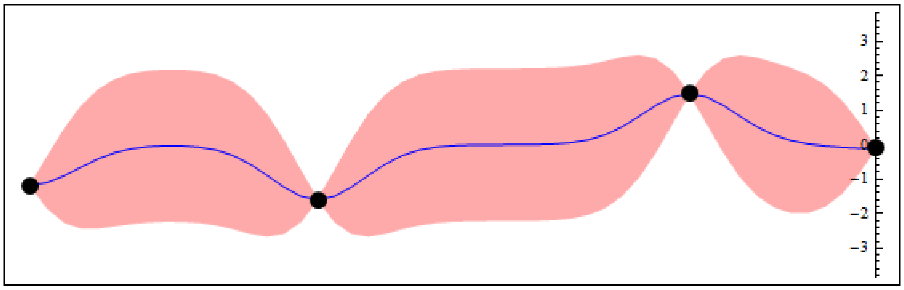
in which $\hat{x}$ is an estimation of the "true" value, x is the measured value and ϵ is an error factor accounting for variations in the measurement. A commonly used model assumes the error to be a normally distributed random variable, i.e., $\epsilon \sim \mathcal{N}(m_\epsilon, \sigma_\epsilon^2)$. This is a drastic reduction because it is usually impossible to determine the ground error distributions. Nonetheless, Robert [172] points out that the choice of the normal distribution in this case is usually justified because it allows for sound "*statistical analysis, which remains valid even if the distribution of ϵ is not exactly normal*"[2]. Another reason [172] is due to the Central Limit Theorem, which suggests that the additional external influence of many small factors of similar magnitude would converge asymptotically to a normal distribution. The normal distribution assumption will also be used in the uncertainty models devised throughout this thesis.

2.1.2 Uncertainty Arising from the Passage of Time

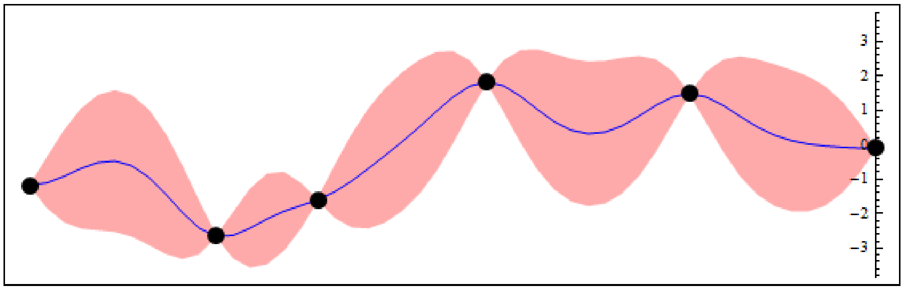
Variations and errors arising from non-stationarity are another common source of uncertainty. For instance, when collecting observations at a constant sample rate – say, the daily closing prices of a stock –, the volume of trades in any given day can drastically vary and, thus, valuable daily trend information will not be reflected within the collected data. In fact, given a collection of samples taken between equally spaced intervals, there are infinitely many continuous functions of time that would output exactly those same values (see Fig. 2.1). This means that all measures of return and risk in the stock market will be taken over incomplete information, since one cannot know if what happened between any two given sampling intervals carries relevant information. As a consequence, an estimative of risk taken from historical data will inevitably suffer from lack of reliability after some time and, thus, such estimatives must be periodically recalculated upon arrival of new data points.

The field of Stochastic Processes (SP) (or random processes) has evolved in statistics to deal with such scenarios. Formally, a SP is a collection of random variables $X(T) = \{X_t, t \in T\}$ (T is a set of indices that can be either discrete or continuous). If T is continuous, then, sampling from a given SP model is equivalent to sampling from a subspace of continuous functions. Such sample is known as an *realization* of the SP. In other words, each realization $\{x_t, t \in T\}$

[2] Apparently, Robert [172] was referring to an equivalence between (a) least-square estimators of the slope of linear regression models under only a simple assumption about the expected value of the independent variables; and (b) maximum likelihood estimators of the slope under a normality assumption.



(a) A GP model with 4 samples.



(b) A GP model with 6 samples.

Figure 2.1: Gaussian Process models estimated over (a) four and (b) six samples. The horizontal axes depict time. The blue lines are the mean processes; whereas the red shadowed areas depict 95% confidence intervals computed from the variance processes.

sampled from the SP may represent one *hypothesis* for fitting a given finite set of observations. For instance, a Gaussian Process (GP) [170] is a SP for which each $X_t \sim \mathcal{N}(m_t, \sigma_t^2)$. Figure 2.1 illustrates how the uncertainty is modeled in GP models: (a) with an realization of only four samples, the total variance of the GP is higher; whereas (b) with two additional samples, the uncertainty about the behavior of the temporal measurement process is reduced.

Roughly speaking, assuming T to be a set of positive integers, a strict stationary process is one whose joint probability distribution of a subset $X(t_1 : t_l)$ of $X(T)$ is the same as of $X(t_1 + h : t_l + h)$, for all l, h. A weaker notion of stationarity relaxes the joint distribution assumption and, instead, states that the mean and variance of the process is what remain constant upon a shift in time. Conversely, (weak) non-stationary processes are those for which the joint distribution (or the parameters) changes with a shift in time. Time-varying quantities such as asset prices can be modeled as non-stationary processes. The most simple model is the so called *random walk*:

$$P_t = P_{t-1} + \epsilon_t, \tag{2.2}$$

which describes how the distribution of prices, P_t, depends on the previous price plus a time-varying error distribution, $\epsilon_t \sim \mathcal{N}(m_{\epsilon_t}, \sigma_{\epsilon_t}^2)$, modeling volatility. No matter what model one takes to represent a non-stationary process, the data analysis involved requires clever techniques to keep track of the varying statistics associated with the collected realizations. In section 2.3, two such techniques will be described: (1) the Kalman Filter [100], for normally distributed data following linear temporal dynamics (such as in the random walk model); and (2) an approach for estimating time-varying probabilities modeled as Dirichlet distributions. Both of them are integrated in the problem-solving tools devised in this thesis.

Modeling Changes with the Sliding Window Approach

Intelligent systems for real-time monitoring and diagnosis can be of great practical interest in a complex dynamic environment. A system capable of continuously tracking time-varying uncertain statistics can pro-actively take exploratory actions for investigating potential problems, which can be achieved by using active learning techniques [148]. In addition, such systems can recommend promising trade-off actions or decisions for mitigating eventual bottlenecks and problems that might be detected.

In many practical scenarios, tasks such as the monitoring, analysis, and prediction of time-varying phenomena rely on the detection of changes in order to suggest adjustments so the predictive models can remain consistent with the environment evolving dynamics. In the following, the problem of change detection in realizations from stochastic processes is roughly described (we follow the notation of Dasu [62]). Let $\{x_1, x_2, \cdots\}$ be a data stream describing some nonstationary process. For instance, x_t can be taken as vectors in $\mathbb{R}^n$ representing the observed consumption rates of a product inventory. The underlying probability distributions can be estimated using *sliding windows*, denoted as $W_{t_1,t_l} = (x_{t_1}, ..., x_{t_l})$. Then, for two empirical distributions, say, $F_{t_1:t_l}$ and $F_{t_1+h:t_l+h}$, estimated from two adjacent windows (realizations) $W_{t_1:t_l}$ and $W_{t_1+h:t_l+h}$, the distance $d_t(F_{t_1:t_l}, F_{t_1+h:t_l+h})$ can be computed and used to test if a statistically significant change has occurred. The null hypothesis in this case is [62]

$$\mathcal{H}_0 : F_{t_1+h:t_l+h} = F_{t_1:t_l}. \tag{2.3}$$

Note that, depending on the choice of h (particularly, if $h < l$), some samples in W_{t_1,t_l} will also appear in W_{t_1+h,t_l+h}. The distance function d_t can be taken as any statistics that represents the discrepancy between two empirical distributions, such as the Kullback-Leibler divergence [62].

Besides change detection, some classes of change patterns can be diagnosed by analyzing data streams. For instance, coagulation and dissolution regions in the state space can be detected from the analysis of temporal and spatial profiles of the velocities with which the empirical distributions modeling the phenomena are changing. This can be done from the estimation of the gradient vector of the adjacent realizations for each spatial coordinate [3]. Specifically, non-parametric Gaussian Kernel-based estimators are useful for this task because they are able to represent multimodality.

By integrating the density of spatial velocity, one can estimate a global change rate that can be useful to measure the evolution degree of a stochastic process. With this approach, it is also possible to find a set of minimal evolving projections that can be valuable for multivariate trend analysis [3].

2.1.3 Uncertainty Arising from Model Misspecification

In the absence of complete information, assumptions about the phenomenon of interest need to be made (i.e., reductions in the form of models of behavior). Perhaps one of the most striking example of model misspecification are claims about economic progress using Gross Domestic

Product (GDP) statistics, which measures only output but cannot distinguish between the variables that are more significant to long term economic stability, let alone the fact that "progress" is a too much subjective concept to be modeled by GDP growth. In other words, there is very low confidence about a nation's economic progress when taking only the perspective of GDP growth.

When it comes to statistical data-driven modeling approaches, two classes of probabilistic models are possible: parametric and nonparametric ones. The latter ones make usage of minimal assumptions and are therefore more flexible. The downside is that such models are more prone to error when the sample size is small. Their main strength is that they can be asymptotically optimal for estimating the ground probability distribution when the sample size goes to infinity [172]. The parametric models, on the other hand, are more pragmatic in the sense that the functional form of the probability density is first assumed, $f(x|\theta)$. Then, statistical techniques such as Maximum Likelihood Estimation (MLE) are employed to determine the value of the parameter θ that best fits the available data. The main advantage is that this may work well even for small sample sizes, provided that the choice of f was a reasonable one. The downside is of course that the choice of a fixed f is problematic for modeling ill-behaved phenomena.

Model Misspecification and Robustness in Financial Data

The model choice and other assumptions thus lead to a drastic and pragmatic reduction that carries itself a lot of uncertainty. Returning to the financial market domain, Hutchison and Pennachi [123] argued that the perfectly competitive market assumption prevents accurate measures of interest rates risk. One evidence of this, they say, are banks that "*exercise market power in setting retail deposit interest rates*".

Even Henry Markowitz, the father of Modern Portfolio Theory, admitted at his Nobel Prize speech that "*Perhaps some other measure of portfolio risk will serve (...) Semi-variance seems more plausible than variance as a measure of risk.*" In fact, despite its intuitive appeal, the practical application of the Markowitz portfolio approach is not reliable and has been reported to behave poorly in real-world out-of-sample data, e.g. [82, 98]. This is mainly due to the prevalent presence of outliers in stock data which makes the tails of the estimated returns distribution fatter than those of the Gaussians assumed in Markowitz portfolio theory [88]. Hence, approaches to robustly estimate the returns distribution parameters have been investigated with the purpose of mitigating such model risk, e.g. [83, 98].

Robustness was defined by Huber and Ronchetti [120] as "*insensitivity to small deviations from* [distributional] *assumptions*", when estimating population parameters from finite samples. This insensitivity is usually traded off by the *efficiency* of the estimator. Statistically, maximum likelihood estimators, such as the sample mean and variance of the Gaussian, carry optimal asymptotic efficiency when the data follow the assumed distribution. However, the presence of outliers in the data may lead to large estimation errors.

Robust statistics may not have optimal efficiency for the assumed model, but they can usually tolerate larger proportions of outliers in the sample before diverging from the true population parameters. For instance, the sample median can tolerate levels up to 50% of contamination (outliers) before breaking down, whereas the sample mean may tolerate none.

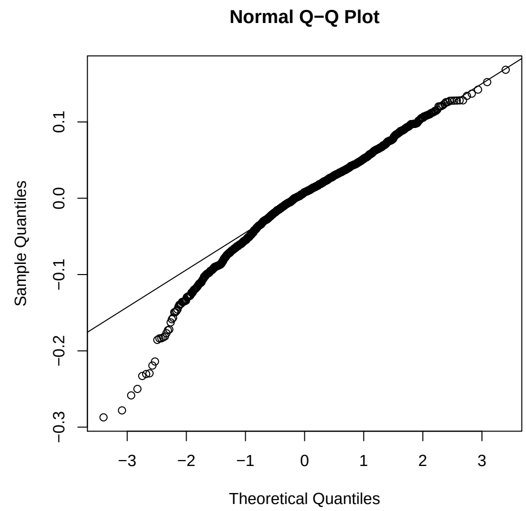


Figure 2.2: Q-Q normal plot for FTSE 100 25 day horizon returns data for the last 6 years.

Venables and Ripley [214] pointed out that the goal of robust statistics is to maintain high (although suboptimal) efficiency in the neighborhood of the assumed distributions.

The motivation for using robust statistics to measure risk and return is to improve the estimation efficiency when the returns data deviates from the Gaussian distribution. In fact, real-world financial data hardly fit the normality assumption. Figure 2.2 shows the normal Q-Q plot for the 25 lagged FTSE 100 index fund daily data from November 2006 to November 2012. It can be observed that the positive returns (profits) portion of the data presents a good fit to normality, while the probability of extreme negative returns (losses) is underestimated by the normality assumption. Remarkably, these are the same observations reported for NASDAQ and S&P500 data by Frahm and Jaekel [88], even considering different time frame and lags. Such violation of the elliptical assumption causes the sample standard deviation to underestimate downside risk and, thus, one can expect the estimation error to impair the performance of Markowitz's portfolios in practice.

In the following, we present a general framework for robustly estimating the covariance matrix under a normality assumption proposed by Alqallaf et al. [4]. This technique may allow for a more robust model by smoothening the influence of outliers and is thus a means to mitigate uncertainty due to model misspecification. Azevedo and Von Zuben [15] have investigated the benefits of this estimation technique for a preemptive multi-objective metaheuristic for evolving a set of robust mutually trade-off investment portfolios for real-world financial data, with promising results when compared with the use of MLE [15] (see also **chapter 4**).

A Robust Pairwise Covariance Matrix Estimation

In this section, we describe a pairwise robust estimation technique for covariance matrices [4]. The main advantages of pairwise approaches are: (i) their quadratic complexity on the dimensionality of the sample vectors (N^2) when compared to the exponential complexity (2^N) of other methods; (ii) the possibility of using any robust univariate statistic for the first and second moments of the distribution; and (*ii*) the resulting matrix is guaranteed to be positive definite, what is not true for other pairwise approaches in the literature.

Despite providing a computationally efficient method, pairwise approaches had been somewhat ignored in the literature because the estimate does not guarantee the resulting matrix to be positive definite. This scenario changed after Maronna and Zamar's [153] work, who proposed an intuitive spectral decomposition approach for restoring the positive definiteness property. The pairwise proposal of Alqallaf et al. [4] is presented in the following.

Given a $T \times N$ dataset matrix $\mathbf{X}$, the $N \times N$ robust covariance matrix estimator is given as follows [4]:

- Compute the univariate robust statistics for the central tendency and dispersion for each column:

$$\hat{m}_j = Q_j^{(2)}(\mathbf{x}_j) \tag{2.4}$$

$$\hat{s}_j = b \cdot \left(Q_j^{(3)}(\mathbf{x}_j) - Q_j^{(1)}(\mathbf{x}_j) \right), \tag{2.5}$$

  with $b = 0.7413$ chosen to provide an unbiased estimate of the standard deviation when the columns follow a Gaussian distribution. Note that Eq. (2.4) corresponds to the computation of the median (breakdown point of 50%), while Eq. (2.5) corresponds to the interquartile range, defined as the difference between the upper ($Q^{(3)}$) and lower ($Q^{(1)}$) quartiles (breakdown point of 25%).

- Compute the quadrant correlation estimate:

$$\hat{c}_{lk} = \frac{1}{T_{lk,0}} \sum_{t=1}^{T} \psi\left(x_{tl} - \hat{m}_l\right) \psi\left(x_{tk} - \hat{m}_k\right), \tag{2.6}$$

  in which

$$\psi(y) = \begin{cases} 1, & y \geq 0 \\ -1, & y < 0 \end{cases} \tag{2.7}$$

  and $T_{lk,0}$ is the number of rows with non-zero entries for both $(x_{tl} - m_l)$ and $(x_{tk} - m_k)$. Note that $\hat{c}_{lk}$ is effectively a robust estimate of the correlation coefficient using rank statistics, which has an intrinsic bias. For x_{tl} and x_{tk} jointly Gaussian, the bias is removed by computing (see [120]):

$$\hat{\rho}_{lk} = \sin\left(\frac{\pi}{2}\hat{c}_{lk}\right), l \neq k; \hat{\rho}_{lk} = 1, l = k. \tag{2.8}$$

  The initial robust covariance matrix estimate is $\hat{\mathbf{\Sigma}}_0 = \{\hat{s}_l \hat{s}_k \hat{\rho}_{lk}\}$, although not necessarily positive definite.

- Compute the spectral decomposition, $\hat{\mathbf{\Sigma}}_0 = \mathbf{U}\mathbf{\Lambda}\mathbf{U}'$, in which the columns of $\mathbf{U}_{N\times N}$ form the orthonormal basis of eigenvectors and $\mathbf{\Lambda}$ is the diagonal matrix whose entries are the corresponding eigenvalues, which are not necessarily positive. By (a) row-wisely projecting the original dataset $\mathbf{X}$ onto the eigenvector basis, i.e., $\tilde{\mathbf{x}}_t = \mathbf{U} \cdot \mathbf{x}_t$ ($\mathbf{x}_t$ is the t-th row of $\mathbf{X}$); (b) computing the univariate robust dispersion statistics $\tilde{s}_j =$

$0.7414 \cdot \left(Q_j^{(3)}(\tilde{\mathbf{x}}_j) - Q_j^{(1)}(\tilde{\mathbf{x}}_j)\right)$; and (c) noting that, for $\hat{\mathbf{\Sigma}}_0$, the entries of $\mathbf{\Lambda}$ are the variances of the projected data on the eigenvector basis, one can simply replace all λ_j by $\tilde{s}_j^2$ to form $\tilde{\mathbf{\Lambda}} = \left\{\tilde{s}_{jj}^2\right\}$.

- The final step consists in applying a series of column permutations over $\tilde{\mathbf{\Lambda}}$ so that the new eigenvalues are ordered from the largest to the smallest one in the main diagonal. The final unbiased robust estimative for the covariance matrix, with the positive definiteness property restored is $\hat{\mathbf{\Sigma}} = \mathbf{U}\tilde{\mathbf{\Lambda}}\mathbf{U}'$.

2.1.4 Uncertainty Arising from Indeterminacy

Heisenberg mostly used the term *indeterminacy* when referring to what was later translated as uncertainty in his original paper [111]. Indeterminacy (or indefiniteness) may thus arise from disjunctive statements of the form *the spin of the electron x is positive or negative*. In the context of MCDM – the main topic of this thesis –, a Decision Maker (DM) may have undefined preferences among multiple conflicting decision criteria due to lack of knowledge about all potential consequences of a given decision.

Consider for instance a DM who must decide between two alternative trade-off daily flights schedules, one possessing a lower estimated congestion probability but with higher estimated takeoff delays. How can the DM state with confidence that the lower takeoff delay schedule is preferable to the one possessing lower congestion probability? What would be worse: dealing with angry passengers because of takeoff delays or because of landing delays? One could argue that takeoff delays would be easier to handle because passengers can comfortably await at the boarding room, while others could argue that most would not even notice short landing delays while entertained with movies and other inflight treats. Moreover, would the trade-offs between takeoff and landing delays be subject to seasonal or regional effects? For instance, letting customers awaiting at a very crowded boarding rooms during a Christmas eve could have far worse consequences than usual. No matter what the DM may think about those subjective questions, the uncertainty conveyed in such decision can be so high that he/she may simply abstain from it, leaving preferences specification undetermined.

The classical decision theory requires the DM preferences to be specified by a total order relation such that it is possible to rank every pair of decisions. This is not the case when the DM cannot determine which among two decisions is preferable. If the DM cannot rank every pair of decisions, then his/her preference relation can be modeled as a partial order. When eliciting preferences by requesting the DM to rank pairs of decisions (e.g. [21, 131]), however, one must be aware that the resulting DM rankings can lead to an order relation $\preceq$ that violates the von Neumann-Morgenstern axioms of rationality [216], as defined in classical utility theory. For instance, transitivity is absolutely necessary for preventing inconsistencies such as $d_1 \preceq d_2$, $d_2 \preceq d_3$ but $d_1 \npreceq d_3$, not to mention cycles of the type $d_1 \preceq d_2 \preceq d_3 \preceq d_1$.

One of the main contributions of this thesis is a strategy for handling the DM preferences indeterminacy by automating the MCDM process. Preferences specification can be thus postponed through the identification of a sequence of flexible provisional trade-off decisions most likely to lead to a larger range of options in future decision periods (see **chapters 5** and **6**).

2.2 Measuring and Reasoning Under Probabilistic Uncertainty

Information Theory (IT) is a theoretical framework with profound implications on automated model building, including data representation (compression), which became widespread after the work of Claude E. Shannon [188]. The most widespread concept contributed by IT to the scientific community was the concept of Shannon's *entropy*, which quantifies the average information contained within a random variable. In this section, we explore the fundamental properties of IT that ultimately contributed to the design of optimization-based multiple criteria decision-making systems possessing *uncertainty awareness* (see **chapter 6**).

2.2.1 Information Theory as a Framework to Quantify Probabilistic Uncertainty

It was then but a short step to use Shannon's entropy to quantify the *uncertainty* contained within a random variable. Mathematically, the Shannon's entropy for a discrete random variable X, $\mathcal{H}(X)$, is defined in terms of a probability distribution $\mathbf{p} = (p_1, \cdots, p_n)^\intercal$ over the n possible values in the domain of X:

$$\mathcal{H}(X) = \mathcal{H}(p_1, \cdots, p_n) := -\sum_{i=1}^{n} p_i \log p_i, \tag{2.9}$$

with the logarithms usually taken to the base 2. When defined in this way, the assumption is that the basic unit of information is the bit (derived from "**bi**nary digi**t**"), which can be used to represent and distinguish between two orthogonal[3] states of a physical system of interest.

In communication systems, X is known as the *source*, and the n possible values taken by X, i.e., $x_1, \cdots, x_n$, are known as *symbols*. Hence, Shannon proposed quantifying the information revealed by the occurrence of the symbol x_k with probability p_k in a data stream produced by the source X as

$$I(p_k) := \log \frac{1}{p_k} = -\log p_k. \tag{2.10}$$

The choice of the logarithm of the probability distribution by Shannon in the definition of $I(p_k)$ can be justified by its algebraic properties. For instance, the amount of information generated by the toss of a fair coin (or, equivalently, the measurement of any system yielding two orthogonal states with equal probability) is one bit. Hence, it makes sense to think that the repeated, independent toss of N such coins will yield N bits of information (i.e., N bits will be required to represent the output of the whole sequence of N tosses). The complete additive property of the logarithm (i.e., $\log(ab) = \log(a) + \log(b)$ and $\log(1) = 0$, for all $a, b \in \mathcal{R}$), was thus crucial in the development of the theory. Therefore, after inspecting Eq. (2.9), it can be concluded that Shannon's entropy for a source X can be interpreted as the expected value of information

[3]Here, orthogonal states should be understood as those which do not share any commonalities with each other, i.e., no (linear) operation over any subset of orthogonal states can be used to perfectly reconstruct other orthogonal state which was not originally present into the combination.

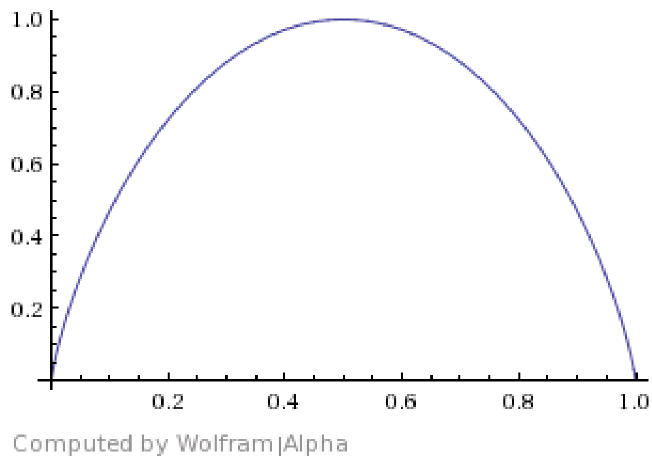


Figure 2.3: The Shannon entropy function (vertical axis) for a binary source (p_1 is represented in the horizontal axis).

contained within a (discrete) source X with probability distribution $\mathbf{p}$:

$$\mathcal{H}(X) := \mathbb{E}\left[I(\mathbf{p})\right]. \tag{2.11}$$

Therefore, for a binary discrete source X (i.e., a Bernoulli random variable) yielding values x_1 and x_2 with probabilities p_1 and $p_2 = (1 - p_1)$, $\mathcal{H}(X) = -(p_1 \log p_1 + p_2 \log p_2)$ represents the average number of bits of information generated for all possible symbols of the source X. From Fig. 2.3, it can be noted that the entropy (uncertainty) reaches its maximum value when the symbols are equiprobable (i.e., $p_1 = 0.5$), in which case the number of bits required to represent a long sequence of N symbols produced by X is N. Conversely, for a source whose probabilities are null except for one symbol x_j for which $p_j = 1$, then $\mathcal{H}(X) = 0$, i.e., the uncertainty is completely absent, what means that no bits are required to store an already expected measurement result.

That is to say, a constant sequence of N symbols emitted by a deterministic discrete source X does not need to be registered once the receiver *knows* about the determinism of X. Therefore, besides the commonsense understanding of Shannon's entropy as a measure of information, uncertainty, and randomness (which are attributes of the source X), this line of reasoning supports two other interpretations, when one allows the empirical probability distribution $\hat{\mathbf{p}}$ to be updated online as more symbols are collected, and when one makes usage of $\hat{\mathbf{p}}$ to update the entropy estimation of X and to measure the change in entropy upon receiving a new sequence of symbols: entropy as a measure of *surprise*, an attribute of the receiver, and *novelty*, an attribute of the data stream.

2.2.2 Information Theory as a Framework to Explain Autonomous Behavior

The problem of explaining self-motivation towards new discoveries is ill-defined. It can be formulated in terms of an endless list of multi-disciplinary approaches and it has been mostly studied within clinical and social psychology research circles. However, taking an information theoretic perspective, Schmidhuber [182] provided an elegant theoretical principle explaining the emergence of creativity, "fun", and motivation. The principle is formulated in terms of information theory; algorithmic complexity; machine learning; and pro-active learning. Schmidhuber [182] postulated that self-motivation (i.e., autonomous behavior) is as intense as the availability of sensorial input exhibiting: (1) novelty; (2) non-determinism (i.e., not fully predictable sequences of symbols); and (3) enough regularity.

In other words, an intelligent agent would lose interest (self-motivation) for events (a) that are unpredictable (or fully predictable); or (b) that do not exhibit enough regularity and, hence, are computationally difficult to be represented by the agent's internal models. On the other hand, an agent would be self-motivated towards exploring input sequences that are surprising (or novel), but regular enough to be comprehended. The main contribution of Schmidhuber [182] to artificial intelligence was then:

> "(...) *theoretically optimal (but not necessarily practical) ways of implementing the basic computational principles on exploratory, intrinsically motivated agents or robots, encouraging them to provoke event sequences exhibiting previously unknown, but learnable algorithmic regularities.*"

Schmidhuber's learning model [182] is then posed in terms of optimal predictors and compressors of data streams. Implicit in such model, is a trade-off between data compressibility – i.e., the agent's ability to consistently summarize and recall events and patterns – and surprise – i.e., how unlike is the observed event given the agent's current internal model of the world surrounding it. Interpreting this implicit trade-off, we can conclude that an autonomous agent in Schmidhuber's [182] perspective would not be motivated towards discovering facts or patterns already known. On the other hand, an agent must be able to near-optimally compress (or represent) the new knowledge according to its internal models. Otherwise, the data stream would be regarded as too complex to be worth analyzing, given the agent's current learning capabilities.

Therefore, it becomes clear that the problem of defining and modeling an autonomous agent exhibiting self-motivation towards new discoveries cannot be put in absolute terms, but only in the context of the internal learning and predictive models an agent has access to, what corroborates Rosen's [174] *anticipatory systems* framework for explaining and designing complex, intelligent behavior.

Very recently, Kaelbling and Lozano-Pérez [128] developed a framework for robot task planning under uncertainty about both the robot current and future states, prompting exploratory actions that resulted in enhanced performance. The contributions provided in **chapter 6** follows a similar path by integrating the uncertainty about the current and future states of a multiple criteria decision-making process into an flexible anticipatory machine learning framework.

2.2.3 Information-theoretic Measures of Autonomy and Synergy

Because there is no clear causal links and arrows in a complex system in which every node influences one another, one way to evaluate the response of a decision-making system to changes in its surround environment may be given by an intuitive notion of autonomy. Bertschinger et al. [27] proposed measuring the conditional mutual information between consecutive system states and the whole history of the environment, which provides the input stream.

When the interaction model between the decision-making system and the environment is known, Bertschinger et al. [27] proposal allowed the further derivation of a causal autonomy measure, which resolves the ambiguity of whether it is the decision-making system which is driving the environment state transitions, or if it is the environmental input stream that is driving the system trajectory. From the experimental results regarding a study of such autonomy measure in the so-called "game of life", philosophical issues about the attribution of control were raised [27].

The result of the complex interactions between intelligent agents (or decision-makers) and the environment leads to the intuition of synergy, i.e., the notion that the macro effects of such interactions carry more information than that conveyed when analyzing the information generated by the agent and the environment separately. In this regard, a more general measure of synergistic mutual information between random variables was proposed in Griffith et al. [101].

With this approach, it is possible to quantify the contribution of each node among a complex network of random variable in predicting the behavior of another random variable [101]. In practice, this can be viewed as a measure of how much information the network conveys about the interactions between two or more individual nodes. Thus, the level of coupling between the elements of such network can be studied. Quantifying how much such synergistic information influences the network evolution through time is nevertheless another interesting open research question.

2.2.4 On the Implicit Relation Between Shannon and Boltzmann-Gibbs Entropies

When the symbols emitted by the source represent orthogonal states of a physical system, the two concepts of entropy in information theory (after Shannon) and in thermodynamics (after Boltzmann and Gibbs [34]) can be seen as closely related. For instance, suppose that the orthogonal states of a classical bit (denoted as "zero", $|0\rangle$, and "one", $|1\rangle$) can be represented in a vector space as the following orthogonal (in the linear algebra sense of an inner product space) vectors:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.12}$$

Moreover, consider that the joint state of two bits (i.e., $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$) can be represented in matrix form by the tensor product between the original bits, yielding the following orthogonal vectors:

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.13}$$

Then, the XOR operator is defined such that $XOR(|00\rangle) = |0\rangle$, $XOR(|01\rangle) = |1\rangle$, $XOR(|10\rangle) = |1\rangle$, and $XOR(|11\rangle) = |0\rangle$. Note that there is no unitary operator (roughly speaking, a linear transformation represented by an invertible matrix) which can be used to represent the mapping performed by the XOR operator. In addition, note that, given only the output of the XOR operator, it is impossible to determine *with certainty* what the input states were, what result in *information loss.* In other words, the phenomenon of information loss is strongly linked to the increase of uncertainty (that can be measured by Shannon's entropy) in the identification of inverse mappings. In the particular case of the XOR operator, it is easy to see that the probability of the input states being $|00\rangle$ given that the observed output was $|0\rangle$ is $p_{|00\rangle||0\rangle} = 1/2$. So is $p_{|11\rangle||0\rangle} = 1/2$. Hence, the Shannon's entropy of the distribution of possible inverse mappings associated with the XOR operator must be maximal.

Therefore, one of the main reasons why linear operators naturally appear in definitions of orthogonality is to preserve the existence of inverse mappings. It turns out that the German-American physicist Rolf Landauer discovered in the 1960's that

> "*(...) any logically irreversible manipulation of information, such as the erasure of a bit or the merging of two computation paths, must be accompanied by a corresponding entropy increase in non-information bearing degrees of freedom of the information processing apparatus or its environment*" [26].

In other words, the loss of information in a physical information processing system must increase the Boltzmann-Gibbs entropy of the system – i.e., the number of possible macrostates it can be at – in the form of heat dissipation. This is one of the main reasons why Quantum Mechanics (QM), when accounting for the probabilistic uncertainty of the trajectories of a physical system of interest, allows only unitary, reversible operators supporting linear physical theories.

Hence, we can state that Shannon and Boltzmann-Gibbs entropies are intrinsically related under the QM perspective and can be thus considered as normative measures of uncertainty of physical dynamical systems when such uncertainty is modeled by probabilistic means[4]. In **chapter 6**, we make intensive use of entropy functions to enable metaheuristics to reason under uncertainty about present and future probabilistic outcomes, thus supporting the automation of complex multiple criteria decision-making processes.

2.2.5 Decision-Making Under Uncertainty

In decision-making under (probabilistic) uncertainty, a Decision Maker (DM) must choose an action u among a set $\mathcal{U}$ whose consequence depends on an uncertain state $x \in \mathcal{X}$. Regardless of whether the state probability distribution $p(x)$ can be estimated a priori, the exact next state assumed as a consequence of the action taken, u, over x is unknown. The *utility function*, which maps actions and states onto real values representing how desirable are the consequences of one action u over a state x for the DM, may be represented as $J(u, x)$.

[4]In the design of quantum information systems, the von Neumann entropy is commonly considered. It can be seen as the quantum generalization of the classical Gibbs entropy. It turns out that the Shannon entropy is also a special case and can be recovered from the von Neumann entropy, if its parameter (the density matrix) is written in terms of its eigenvectors, which are orthonormal base (classical) states [24].

An *optimal* action $u_B^\star$ can be determined by *maximizing the expected utility*:

$$u_B^\star = \arg\max_{u\in\mathcal{U}} \int_{\mathcal{X}} J(u,x)p(x)\mathrm{d}x. \tag{2.14}$$

The obvious drawback about solving Eq. 2.14 is that the ground state distribution $p(x)$ is unknown and must be estimated from past realizations. However, except when the process generating the realizations is *ergodic*[5], this strategy can lead to severe model misspecification. One way to cope with this difficulty is by making use of the *sliding window* estimation approach (see section 2.1.2) so that the state distribution can be updated over time without the influence of past uncorrelated realizations. Besides, one can use statistical tracking techniques such as the Kalman Filter (KF, see section 2.3.4) to obtain $\hat{p}(x)$. In **chapter 5**, our contributions to anticipatory multiple criteria decision-making utilizes a maximum expected utility approach, with sliding window techniques (see **chapter 6** and **chapter 7**), and KF estimation, although the proposed utility function does not carry any information regarding the DM preferences and desirabilities.

An alternative to the expected utility approach is a worst-case approach known as the *minimax* method. The minimax rule does not require the specification of $p(x)$ when selecting the action $u_M^\star$ minimizing the maximal possible loss [164], but requires the estimation of $L(u,x) = \sup_{u\in\mathcal{U}}\{J(u,x)\} - J(u,x)$, i.e., the maximal difference in utility between the optimal action and a candidate action for state x. The minimax rule is expressed as follows:

$$\begin{aligned} u_M^\star &= \arg\min_{u\in\mathcal{U}} \max_{x\in\mathcal{X}} L(u,x), \\ L(u,x) &= \sup_{u\in\mathcal{U}}\{J(u,x)\} - J(u,x). \end{aligned} \tag{2.15}$$

Hence, the minimax decision is that leading to the minimal maximal loss over all possible states. The maximum expected utility and the minimax approach generally do not lead to the same choice, but there are specific conditions over $p(x)$ for $u_B^\star = u_M^\star$ [164].

2.3 Learning and Prediction Under Uncertainty

A real-valued time series X can be represented as a mapping $X : \mathcal{T} \times \mathcal{E} \to \mathcal{R}$ such that, for each fixed t, $X(t,e)$ is a random variable in a probability space, where e represent elementary events and $\mathcal{T} = 1,...,T$ is a set of indices that enumerate the resulting real-valued measures sequentially. The notation X_t describes the probabilistic state of a system of interest at time t. The problem of time series prediction can be thus consider observing a system of interest until time t. The goal is to predict the value of X_{t+H}, where H is known as the prediction horizon. The prediction of X_{t+H} is denoted as $\hat{X}_{t+H}$, and can be represented as a function of the historical data samples collected at times $0, 1, \cdots, t-1$.

[5]Ergodic processes can be intuitively viewed as those whose moments (e.g. mean, variance, skewness, kurtosis, etc.) of the ground distribution can be inferred with arbitrary accuracy from sufficiently large samples of the process. Non-ergodic processes, therefore, can be viewed as those whose changing patterns of the ground distribution impair the confidence with which one can infer the process moments, despite the samples availability.

2.3.1 On Dynamical Systems Modeling and The Takens Theorem

When representing dynamical systems with d interacting parts, it is reasonable to describe the system by d independent real-valued variables: $\{X^1, X^2, \cdots, X^d\}$. For instance, one can represent the joint temporal evolution of a complex financial market by describing the trajectory over time of a state vector in $\mathcal{R}^d$, $\mathbf{s}_t$. Each of the d components of $\mathbf{s}_t$ represent all different relevant interacting factors that completely describe the system. For instance, one can encode in $\mathbf{s}_t$ several economical output indicators, as well as relevant business related variables pertaining to N different investment assets in a stock market, although such description will always be incomplete for systems as *complex* as financial markets.

Current research on *complex dynamical systems* modeling has been mostly credited for best representing large-scale real-world problems wherein a large number of agents interact with each other, producing complex self-organizing and emergent macro effects. Despite its theoretical appeal, one practical difficulty arising from such approach is that efficient methods for inference and approximate reasoning within the complex systems paradigm are virtually non-existent.

Although the characterization of what is meant by the term "complex" varies across disciplines (e.g., game theory, dynamical systems, complex network theory, among others), we briefly describe the notion of complexity in statistical mechanics. In a very recent work co-authored by the 1969 physics Nobel laureate, Murray Gell-Mann, Hanel et al. [107] implicitly defined complex systems as those that can be best modeled as non-Markovian and non-ergodic stochastic processes.

Hanel et al. [107] also proved, for the first time, the existence of the principle of maximum entropy for complex systems. The principle provides mathematical, and, hence, algorithmic procedures for statistically reducing a priori knowledge bias when modeling a process by fitting probability distributions, under a Bayesian inference framework. In addition, they derived a class of entropy measures that satisfies Shannon-Khinchin axioms, which prescribe the intuitive notions that should be satisfied by any measure of information, e.g., separability, complete additiveness, among others.

The mathematical procedure requires finding a multiplicative factoring of the relative entropy function. Nevertheless, Hanel et al. [107] provide the first *exact* derivation of the maximum entropy principle for complex systems with long-term memory. The application of such principle may lead to effective Bayesian inference in complex systems, what may result in a paradigm shift for machine learning.

The Takens Theorem

Let $M \subseteq \mathcal{R}^d$ represent the manifold containing all possible trajectories of a complex dynamical system, at any time frame. Let also X^i be any constituent part of M. The Takens Theorem [202] establishes the necessary conditions upon which it is possible to construct a "shadow" manifold, $M_{X^i} \in \mathcal{R}^d$, that preserves not only the topology of the original manifold M, but also its Lyapunov exponents[6] [200]. Such shadow manifold can be formed by carefully choosing the

[6] It is possible to estimate the exponents of the so-called Lyapunov functions describing discrepancies between infinitesimally close trajectories within the phase state of a dynamical system. The analysis of the maximal Lyapunov exponent can reveal information about the degree of predictability and chaos within such systems.

minimum dimensionality of time lags (d) from the original time series so to encode any value in the trajectory of X^i as the mapping $g : \{X^i_{t-h_d}, \cdots, X^i_{t-h_1}\} \rightarrow X^i_t$.

Moreover, Takens [202] proves that there is a one-to-one map between M_{X^i} and M, what means that all information contained within the dynamical system manifold M can be recovered by carefully processing the historical measurements taken from any of its projections onto an individual constituents, say, X^i. Despite being a remarkable and intuitive result, Takens' assumptions require the system to be stationary and free from quantization errors [195].

Application of The Takens Theorem for Causality Detection

Nevertheless, in a recent application, Sugihara et al. [200] made use of the following transitive argument to design a test of causality between two or more time series, denoted as convergent cross mapping, which takes advantage of the Takens Theorem. Let M_{X^i} and M_{X^j} be two shadow manifolds of minimum dimensionality w.r.t. a dynamical system evolving in M. Since there are one-to-one maps leading both M_{X^i} and M_{X^j} to M, then there must also be an one-to-one map leading directly from M_{X^i} to M_{X^j}.

After measuring the similarity between the trajectories within M_{X^i} and M_{X^j} by finding corresponding nearest neighbors states between both shadow manifolds, Sugihara et al. [200] argued that it is possible to infer whether the time series X^i and X^j describe concepts belonging to the original dynamical system. The inference relies on comparing the convergence rates (in terms of prediction error) upon increasing the available time series length (L) – i.e., the number of available historical data samples – when cross-mapping from a shadow manifold to reconstruct the other time series. In other words, with the increase of L, the correlation coefficient between the ground and the reconstructed values of the cross-mapping $\hat{X}^i_t|M_{X^j}$ should converge above a statistically significant level faster than that of the cross-mapping $\hat{X}^j_t|M_{X^i}$.

Although the non-stationarity assumption of the joint asset return time series used in the experiments described in **chapter 6** does not match the assumptions of the Takens Theorem, we investigated (a) the effects of increasing the number of historical observations (window size factor) on the prediction accuracy of the proposed Bayesian estimation tools; and (b) the impact of that factor on the performance of the resulting anticipatory flexible decision-making system for financial portfolio selection.

2.3.2 The Bayesian Learning Approach

In many problems, one has to infer unknown states from noisy and time-varying observations, such as the position and velocity of objects in tracking systems, the volumetric flow rate and temperature in industrial processes, or the interest rates in a financial market. More generally, the moments of the distributions describing stochastic processes must be estimated. In such scenarios, Bayesian models are commonly used because they allow for mathematical optimal ways of combining observations to prior knowledge in order to reduce the uncertainty about a given system of interest, by updating the posterior distribution using the Bayes Theorem [172].

The 2011 Turing Award recipient Judea Pearl [166] has made the compelling point that while probabilistic knowledge of a system of N uncertain quantities is defined in terms of joint distributions in the traditional frequentist approach, the Bayesian approach for handling uncertainty relies instead on the learning of conditional relations between two or more such variables. Thus, in the frequentist point of view, conditional distributions are merely an artifact without much importance relating joint distributions, whereas, in the Bayesian view, they are the principal means by which uncertain knowledge is encoded. As put by Parsons [165], the focus on conditionals implies that Bayes Theorem becomes "*a normative rule for updating probabilities in line with new information*".

Conditional probabilities such as $Pr(Y|X)$ convey the notion that *if I am given X, then I can guess the value of Y*. The Bayes Theorem provides a means to go the other way around as to *infer* the probable values of X starting with Y. If X and Y are discrete events such that $Pr(X) \neq 0$, then

$$Pr(Y|X) = \frac{Pr(X|Y)Pr(Y)}{Pr(X)}, \tag{2.16}$$

in which $Pr(X) = Pr(X|Y)Pr(Y) + Pr(X|\bar{Y})Pr(\bar{Y})$ and $\bar{Y}$ denotes the complement of Y. The continuous counterpart (for continuous random variables) of the Theorem is:

$$g(y|x) = \frac{f(x|y)g(y)}{\int f(x|y)g(y)dy}. \tag{2.17}$$

Interpreting Eq. (2.17), $g(y|x)$ is proportional to the likelihood $f(x|y)$ of y given that we gathered the evidence x weighted by how likely we believe such value of y – whose probability we want to *update* – to show up, i.e. $g(y)$. As pointed out by Robert [172], Bayes and Laplace quickly figured out that the Bayes Theorem would be of main importance for parametric estimation. Particularly, the uncertainty on the parameters θ of a given probability density function f could be modeled through a distribution π over the parameter space, known as the prior distribution.

Given the sample distribution $f(x|\theta)$ and the prior distribution $\pi(\theta)$, there are four elements at the core of learning with Bayes rule that can be inferred [172]:

- The joint distribution of x and θ: $h(x, \theta) = f(x|\theta)\pi(\theta)d\theta$;

- The marginal distribution of x: $f(x) = \int h(x, \theta)d\theta = \int f(x|\theta)\pi(\theta)d\theta$;

- The posterior distribution of θ given evidence x: by using Bayes Theorem,

$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta)d\theta}; \tag{2.18}$$

- The predictive distribution of a given value of interest y, considering $y \sim g(y|\theta, x)$:

$$\pi(y|x) = \int g(y|\theta, x)\pi(\theta)d\theta. \tag{2.19}$$

2.3.3 Maximum a Posteriori (MAP) Estimation

The posterior distribution derived using Bayes Theorem has another important role in the estimation of parameters of probabilistic models. It can be used in a point estimation framework to infer the most likely value of the parameter θ given a piece of evidence x. This approach is known as the Maximum a Posteriori (MAP) estimation of θ, denoted as $\delta^*(\theta)$, and is given as the mode of the posterior distribution:

$$\begin{aligned}\delta^*(\theta) &= \arg\max_\theta \frac{f(x|\theta)\pi(\theta)}{\int f(x|\phi)\pi(\phi)d\phi} \\ &= \arg\max_\theta f(x|\theta)\pi(\theta).\end{aligned}$$

This point estimator possess some interesting theoretical properties, one being that it is asymptotically optimal when the sample size goes to infinity (i.e., when the estimation is repeated for a long series of x's), being equivalent to the MLE in such case. It is also a robust estimator for small sample sizes in the sense that it can be interpreted as a regularized version of the MLE [172].

Example: MAP estimation for a Gaussian mean

The MAP estimation for the mean vector of a multivariate Gaussian $\mathcal{N}(\boldsymbol{\theta}, \boldsymbol{\Sigma})$ with conjugate distribution (the prior distribution) $\mathcal{N}(\mu, \mathbf{A})$ and known covariance matrix given an observation $\mathbf{x}$ yields

$$\delta^*(\mathbf{x}) = \left(\Sigma^{-1} + \mathbf{A}^{-1}\right)^{-1}\left(\boldsymbol{\Sigma}^{-1}\mathbf{x} + \mathbf{A}^{-1}\mu\right). \tag{2.20}$$

Note that $\delta^*(\mathbf{x})$ is expressed as a convex combination between the evidence $\mathbf{x}$ and the prior mean μ, where the weights are proportional to the inverse of the covariance matrices [172].

In the following, a Bayesian network model for estimating noisy, time-varying quantities assuming Gaussian priors and linear, Markovian dynamics is described.

2.3.4 The Kalman Filter

In several real-world applications, a solution must be deployed in a time-varying, noisy and unseen environment for which it has not been optimized for. In this thesis, we make use of Bayesian networks for computing the posterior state distribution of such quantities, so to allow the search for solutions for which better performance on the targeted future environments is predicted. For stochastic MCDM, the idea is to achieve this goal by estimating the trajectory of random objective vectors over time.

The Kalman Filter [100] is the most widespread Bayesian network with countless successful applications, providing closed-form expressions for updating the state posterior distributions if the uncertainty can be modeled as a multivariate Gaussian and if the system dynamics are linear. As a result, exact and computational inexpensive one-step ahead inference of the state predictive distribution is available. The KF computes in closed-form the best possible estimation (in the mean-square sense) for the distribution of the hidden state vector. The main assumption for the

KF is that of a linear Gauss-Markov state-space model, given as the following joint distribution factored representation of hidden states and measurements:

$$p(\mathbf{X}, \mathbf{Y}) = p(\mathbf{x}_1)p(\mathbf{y}_1|\mathbf{x}_1) \prod_{t=2}^{H} p(\mathbf{x}_t|\mathbf{x}_{t-1})p(\mathbf{y}_t|\mathbf{x}_t) \text{ (factored joint distribution)}, \tag{2.21}$$

$$\mathbf{x}_0 \sim \mathcal{N}\left(\mathbf{m}_{\mathbf{x}_0}, \mathbf{C}_0\right) \text{ (initial state distribution)}, \tag{2.22}$$

$$\mathbf{x}_t|\mathbf{x}_{t-1} \sim \mathcal{N}\left(\mathbf{A}\mathbf{m}_{\mathbf{x}_{t-1}}, \mathbf{C}_{t|t-1}\right) \text{ (predictive distribution)}, \tag{2.23}$$

$$\mathbf{y}_t|\mathbf{x}_t \sim \mathcal{N}\left(\mathbf{M}\mathbf{m}_{\mathbf{x}_t}, \mathbf{R}_{t|t-1}\right) \text{ (measurement distribution)}, \tag{2.24}$$

in which $\mathbf{X} = \{\mathbf{x}_1, \cdots, \mathbf{x}_H\} \subset \mathcal{R}^n$ is the set of unknown states, and $\mathbf{Y} = \{\mathbf{y}_1, \cdots, \mathbf{y}_H\} \subset \mathcal{R}^p$ is the set of available measurements (which may have lower dimensionality, $p < n$). The matrices $\mathbf{A}$ and $\mathbf{M}$ encode the state linear dynamics and the measurement function, respectively. The uncertainty is encoded by one covariance matrix for the unknown states, $\mathbf{C}_{t|t-1}$, and by another covariance for the observations, $\mathbf{R}_{t|t-1}$.

The equations for the KF come from the recursive application of the Bayes Theorem to the conditional densities $p(\mathbf{x}_t|\mathbf{y}_t)$ and $p(\mathbf{x}_{t+1}|\mathbf{x}_t, \mathbf{y}_t)$, corresponding to the updating of the posterior given a measurement at time t, and the prediction of the next state using the known dynamical model, $\mathbf{A}$. Both densities are Gaussian and can be analytically determined assuming the stochastic dynamical model in (2.21)–(2.24).

The KF is then capable of tracking state vectors evolving according to linear dynamics under Gaussian uncertainty. There are two steps required for recursively applying the KF for estimating a Gaussian random vector evolving on time: (a) a prediction step; and (b) a measurement correction step.

KF Prediction Step

This step makes use of the available dynamical model describing the trajectory of the hidden state vector over time to compute an estimation for the distribution of the hidden state vector in the current time step, given past estimations, as represented in Eq. (2.23). The dynamical model is assumed to be[7]:

$$\mathbf{x}_t = \mathbf{A}_t\mathbf{x}_{t-1} + \mathbf{q}_t, \quad \mathbf{q}_t \sim \mathcal{N}\left(\mathbf{0}, \mathbf{Q}_t\right), \tag{2.25}$$

where $\mathbf{q}_t$ is a zero mean noise process with covariance $\mathbf{Q}_t$, which can be estimated from the available historical data. The closed-form KF estimation is possible due to the mathematical tractability resulting from the linear algebra over Gaussian random vectors.

The computation of the predictive estimation of $\hat{\mathbf{x}}_t|\mathbf{x}_{t-1}$ from the prior distribution of $\hat{\mathbf{x}}_{t-1}$ is done over the parameters:

$$\mathbf{m}_{\hat{\mathbf{x}}_t|\mathbf{x}_{t-1}} = \mathbf{A}_t\mathbf{m}_{\mathbf{x}_{t-1}} \tag{2.26}$$

$$\mathbf{C}_{\hat{\mathbf{x}}_t|\mathbf{x}_{t-1}} = \mathbf{A}_t\mathbf{C}_{t-1}\mathbf{A}_t^\intercal + \mathbf{Q}_t, \tag{2.27}$$

where $\mathbf{Q}_t$ is the covariance of the additive noise process affecting the dynamics of $\mathbf{x}_t$. The predictive distribution $\hat{\mathbf{x}}_t|\mathbf{x}_{t-1}$ thus expresses the KF *prior belief* of what the true hidden state distribution $\mathbf{x}_t$ might be.

[7]We assume no control input in our tracking application, i.e., the objective vectors are assumed to evolve independent of external actions.

KF Measurement Correction Step

This step updates the estimated $\hat{\mathbf{x}}|\mathbf{x}_{t-1}$ distribution obtained in the prediction step upon the collection of a new sample $\mathbf{y}_t$ resulting from measuring the hidden state $\mathbf{x}_t$. The measurement distribution is denoted as $\mathbf{y}_t|\mathbf{x}_t$ and results from the following expression:

$$\mathbf{y}_t = \mathbf{M}_t\mathbf{x}_t + \mathbf{v}_t, \quad \text{with} \quad \mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_t), \tag{2.28}$$

in which $\mathbf{M}_t$ is a *measurement function model* mapping the true state $\mathbf{x}_t$ into a measurement state $\mathbf{y}_t$ and $\mathbf{v}_t$ is a measurement noise process that is assumed as a zero mean Gaussian additive noise process, with a noise covariance $\mathbf{R}_t$ that can also be estimated from available historical data. The KF estimation for the current hidden state distribution given the measurement $\mathbf{y}_t$ is then represented as $\mathbf{x}_t|\mathbf{y}_t$ and is also computed in closed-form over the parameters.

The closed-form equations arising from the KF measurement correction step make use of feedback from the *residual* between the observed measurement $\mathbf{y}_t$ in Eq. (2.28) and the measurement expected from the estimation $\hat{\mathbf{x}}_t|\mathbf{x}_{t-1}$ computed in the prediction step, i.e.,

$$\tilde{\mathbf{y}}_t = \mathbf{y}_t - \mathbf{M}_t\hat{\mathbf{x}}_t|\mathbf{x}_{t-1}, \tag{2.29}$$

where the residual $\tilde{\mathbf{y}}_t$ is also known as an *innovation* process containing novel statistical information that was not already known from the series of historical measurements $\mathbf{y}_0, \cdots \mathbf{y}_{t-1}$. Similarly, the residual measurement covariance matrix is expressed as

$$\mathbf{S}_t = \mathbf{M}_t\mathbf{C}_{\hat{\mathbf{x}}_t|\mathbf{x}_{t-1}}\mathbf{M}_t^{\intercal} + \mathbf{R}_t. \tag{2.30}$$

The so-called *Kalman gain* – represented as the matrix $\mathbf{K}_t$ – adjusts the importance of the residual relative to the prior estimate $\hat{\mathbf{x}}_t|\mathbf{x}_{t-1}$ in the correction step of the true state distribution. Intuitively, when the pairwise covariances in the measurement noise covariance matrix $\mathbf{R}_t$ are high, less importance should be assigned to the observations and, therefore, the Kalman gain should be small. Conversely, when the estimated prior covariance $\mathbf{C}_{\hat{\mathbf{x}}_t|\mathbf{x}_{t-1}}$ is high, the state vector is expected to convey high variability and, hence, the Kalman gain should be high, to more strongly account for the observed measurements. In fact, the Kalman gain is the minimum mean-square error factor for correcting the prior estimation $\hat{\mathbf{x}}_t|\mathbf{x}_{t-1}$ while accounting for the residuals resulting from the observed measurements $\mathbf{y}_t$ and is expressed as

$$\mathbf{K}_t = \mathbf{C}_{\hat{\mathbf{x}}_t|\mathbf{x}_{t-1}}\mathbf{M}_T^{\intercal}\mathbf{S}_t^{-1}. \tag{2.31}$$

The correction step is then completed by updating the parameters of the KF prior state estimation in light of the residual:

$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_t|\mathbf{x}_{t-1} + \mathbf{K}_t\tilde{\mathbf{y}}_t, \tag{2.32}$$

$$\mathbf{C}_{\hat{\mathbf{x}}_t} = (\mathbf{I} - \mathbf{K}_t\mathbf{M}_t)\mathbf{C}_{\hat{\mathbf{x}}_t|\mathbf{x}_{t-1}}. \tag{2.33}$$

Prediction in Metaheuristics Using KF

The first attempt to integrate KF estimation into metaheuristics was due to Stroud [199], who designed an extended KF to support an active learning heuristic for further reducing the uncertainty of the system by re-evaluating specific solutions. Particularly, an extended KF was designed to support heuristic decisions on "*(...) when to generate a new individual, when to re-evaluate an existing individual, and which one to re-evaluate*".

Rossi et al. [175] applied the KF for tracking the optimum solution (for single-objective, unimodal problems) and proposed several ad-hoc heuristics to incorporate the predictive knowledge in order to guide the evolutionary process toward promising regions. Three heuristic classes were assessed: (*i*) incorporating the predictive knowledge into the metaheuristic variation operators (*ii*) generating new solutions around the estimated optimum; and (*iii*) biasing the search by increasing the fitness evaluation inversely proportional to the distance of the current solution to the predictive optimum, discounted by the amount of uncertainty in the prediction.

The latest work found in the literature applying KF estimation in metaheuristics had just been published by the time this thesis was being wrapped up and is due to Muruganantham et al. [159], who proposed tracking each candidate solution in the decision space of a Multi-Objective Optimization (MOO) solver in time-varying – yet deterministic – environments. The general idea somewhat resembles our proposal for tracking mutually non-dominated decision vectors presented in **chapter 6**, although, in our case, the estimation is done in the $(d-1)$-simplex and, thus, the linear Gaussian-Markov assumption does not hold, what leads to the need for a different Bayesian estimation model based on the Dirichlet distribution.

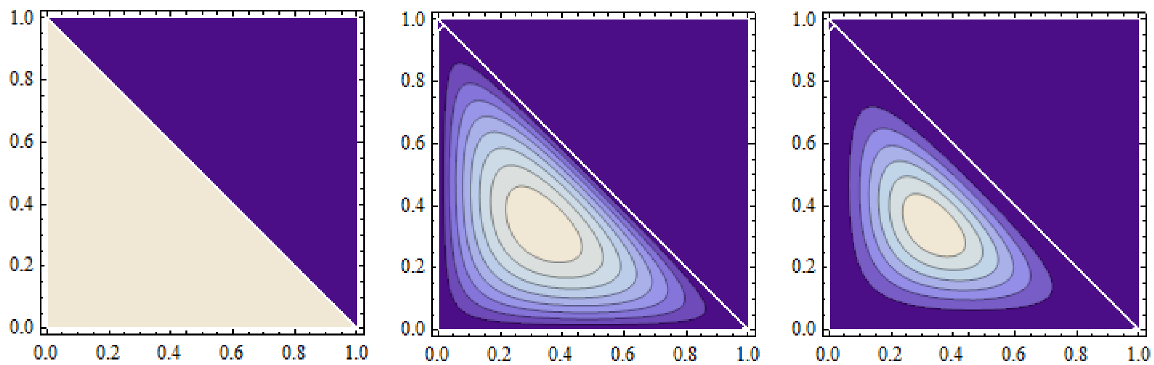
2.3.5 Dirichlet Priors for Bayesian Estimation in Simplex Spaces

Sometimes, as it is the case in this thesis (see **chapter 6**), it will be useful to estimate uncertain quantities in structured spaces other then $\mathcal{R}^n$. For instance, the following expression describes a vector $\mathbf{w}$ belonging to a *simplex space*:

$$\mathbf{w} = (w_1 \ w_2 \ \cdots \ w_d)^\intercal \in S^{d-1} = \left\{ \mathbf{w} \in \mathcal{R}^d : \sum_{k=1}^{d} w_k = 1, \ \forall \ j \ w_j \geq 0 \right\}. \tag{2.34}$$

Note that we are interested on representing $\mathbf{w}$ as a random vector of *proportions* lying in the $(d-1)$-simplex space, S^{d-1}. It is worth noting that simplex spaces possess elementary properties that may simplify the resulting analytical expressions when performing Bayesian inference over proportions. The first property regards a basic notion of *complementarity*. It describes the fact that each individual variable (proportion) w_i can be completely described as the complement of the sum over the remaining variables to the unity, i.e., $w_i = 1 - \left(\sum_{k \neq i} w_k\right)$. This is the reason why simplex spaces are denoted as S^{d-1} and why we only need to account for $d-1$ variables when depicting S^{d-1} (see Fig. 2.4). In other words, each and every variable in a simplex space complements one another; i.e., any subset of $d-1$ variables from the simplex space yields complete information about the value of the remaining variable.

The second intuitive property is given as a geometrical interpretation of the maximum entropy discrete distribution represented as a point in S^{d-1}, when viewing the simplex as a natural space for representing all possible discrete probability distributions over d orthogonal states. Take $d = 3$ as an example, for which S^2 can be depicted in an extended form as an equilateral triangle. Then, the orthocenter, circumcenter, center of mass, and even the incenter



(a) $\mathcal{D}(1,\ 1,\ 1)$ (b) $\mathcal{D}(2,\ 2,\ 2)$ (c) $\mathcal{D}(3,\ 3,\ 3)$

Figure 2.4: Contour plots for the Dirichlet distribution in the 2-simplex.

of S^2 all perfectly coincide at the point $\mathbf{w}_C = (1/3\ 1/3\ 1/3)^\intercal$, which results from the intersection of the three lines of symmetry of S^2 [81][8]. It turns out that $\mathbf{w}_C$ thus also represents the discrete distribution of maximal Shannon's entropy (see section 2.2.1).

The Dirichlet Distribution

One of the most important questions when designing Bayesian inference methods for simplex spaces is: what distribution should we use to represent our uncertainty regarding $\mathbf{w} \in S^{d-1}$? We could in principle try fitting an unbounded distribution such as the multivariate Gaussian to represent the data collected in S^{d-1}. But such parametric model would lead to large estimation errors nearby the boundaries of S^{d-1}, since the Gaussian would always attribute positive density to unfeasible samples, i.e., points that lie outside S^{d-1}.

A more plausible choice could be to fit distributions in the surface of a unit hypersphere, such as the von Mises-Fisher distribution [18]. But while such distribution conveys some structure, i.e., the elements in S^{d-1} are also unit vectors and can be thus represented in the unit $d-1$ hypersphere surface, there are still elements outside of S^{d-1} that could be attributed positive probability density.

One sound choice that is just right for modeling uncertainty in the S^{d-1} simplex (i.e., that indeed takes S^{d-1} support) is the Dirichlet Distribution (DD) [160], which is a multivariate version of the Beta distribution. Because any multinomial probability distribution can be represented as a point in S^{d-1}, the DD is also known as a distribution over distributions. Thus, the DD can also be used to represent the underlying uncertainty over true probability distributions, what conveys a notion of *second-order uncertainty*.

This is one of the main reasons why the DD is becoming popular in several Bayesian machine learning applications, e.g. text mining and topic models [19]. Another reason is that the DD is part of the exponential family [160], which means that the conjugate prior used for estimating the posterior distribution using Bayes Theorem is also from the exponential family. This very

[8]This property only holds for equilateral triangles. For other triangles, Euler [81] proved that those notable points (except the incenter) are always different, yet collinear, lying in the so-called Euler's line.

much simplifies the calculations, otherwise, one would need to rely on numerical methods to estimate the posterior.

In order to motivate its usage, consider a simple coin flipping experiment. The goal is to represent our uncertainty about the probability that a coin flip will yield heads, based on previously observed flippings. We may be very well dealing with a fair coin and therefore we may want to assign $P(\text{heads}) = P(\text{tails}) = \frac{1}{2}$. On the other hand, we have no idea from where this coin came from or how it was fabricated. So we do not want to rule out the possibility that the coin is biased. Here is when the DD usage becomes very handy: we can assign positive probability density to all the possible heads odds. Not only that, but also, by using the DD[9], we can model the decrement of our uncertainty about the heads odds as we observe more and more flippings, making the probability distribution P to progressively peak at the true odds value.

Let $\mathbf{w}$ denote a vector of proportions so that w_k represents the proportion of the k-th item. For example, in financial portfolio applications, w_k can be the proportion of wealth invested in the k-th asset. The DD can be written as:

$$\mathbf{w} \sim \mathcal{D}(\boldsymbol{\alpha}) = \frac{\Gamma\left(\sum_k \alpha_k\right)}{\prod_k \Gamma(\alpha_k)} \prod_k w_k^{\alpha_k - 1}, \tag{2.35}$$

where $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ is the Gamma function[10] and the parameter $\boldsymbol{\alpha} = (\alpha_1 \ \cdots \ \alpha_N)^\intercal$ determines the shape of the distribution (with $\alpha_k > 0$ for all k). Note that if the value of all but one of the N components of $\mathbf{w}$ are given, then the value of the remaining one is also known because $\sum_k w_k = 1$. Thus, the Dirichlet density will lie in a $d-1$ hyperplane (the $(d-1)$-simplex) in d-dimensional space.

The DD can be re-parametrized by the mean vector $\mathbf{m} = \left(\frac{\alpha_1}{\alpha_C} \cdots \frac{\alpha_N}{\alpha_C}\right)^\intercal$, where $\alpha_C = \sum_k \alpha_k$ is known as the *concentration* parameter, which controls the dispersion of the distribution around $\mathbf{m}$: the higher α_C is, the more concentrated the density will be around the mean vector in the $(d-1)$-simplex (see Fig. 2.4).

The concentration parameter has an interesting interpretation in terms of learning (and therefore uncertainty reduction): if each α_k is interpreted as *pseudo-counts*[11], then the more we accumulate evidence, the more certain we will be about the location of $\mathbf{w}$ in the $(d-1)$-simplex.

Remark: When $\boldsymbol{\alpha} = \mathbf{1}$, the DD describes a uniform distribution in the $(d-1)$-simplex, which is the *maximum entropy distribution* for random vectors whose moments are not constrained [163], i.e., whose only constraint is that it is contained in the distribution support.

The moments of the DD are:

$$\mathbb{E}[w_k] = \frac{\alpha_k}{\alpha_C}, \tag{2.36}$$

[9] In the case of a simple coin experiment, the DD actually reduces to a Beta distribution.

[10] $\Gamma(x)$ is an extension of the factorial function in the real line, defined as $\Gamma(x) = (x-1)\Gamma(x-1)$ and $\Gamma(1) = 1$. For a positive integer n, $\Gamma(n) = (n-1)!$.

[11] Pseudo-counts are used to adjust models based on observed frequencies of occurrences. Rare (but not impossible) events that were not yet observed are usually inaccurately attributed zero probability, but this can be artificially corrected by adding small quantities to such frequencies so the model can account for such events.

$$\mathrm{Var}(w_k) = \frac{\alpha_k(\alpha_C - \alpha_k)}{\alpha_C^2(\alpha_C + 1)}. \tag{2.37}$$

The covariances are always non-positive for any two proportions in the DD, since whenever w_k is incremented, some or all other proportions must be necessarily decremented so as to satisfy $\sum_k w_k = 1$. Thus,

$$\mathrm{Cov}(w_i, w_j) = \frac{-\alpha_i \alpha_j}{\alpha_C^2(\alpha_C + 1)}. \tag{2.38}$$

It also is worth noting that the marginal distribution of w_k is a Beta distribution, i.e.,

$$w_k \sim Beta(\alpha_k, \alpha_C - \alpha_k). \tag{2.39}$$

Remark: James and Moisimann [124] showed that the Dirichlet distribution is characterized by *neutrality*, which is a strong notion of statistical independence on spaces whose coordinates must add up to a certain sum. That is to say, if $\mathbf{w} \sim \mathcal{D}(\boldsymbol{\alpha})$, then the vector $\mathbf{w}$ is completely neutral because $w_k \sim Beta(\alpha_k, \alpha_C - \alpha_k)$ is provenly independent of

$$\mathbf{w}^{(-k)} = (\frac{w_1}{1 - w_k} \ \cdots \ \frac{w_{k-1}}{1 - w_k} \ \frac{w_{k+1}}{1 - w_k} \ \cdots \ \frac{w_d}{1 - w_k})^{\intercal} \sim \mathcal{D}(\boldsymbol{\alpha}^{(-k)}), \tag{2.40}$$

for all $1 \leq k \leq d$, what implies that any permutation of $\mathbf{w}$ is neutral [124]. We can thus see that the neutrality property of the DD is thus intuitively linked with the complementarity property of S^{d-1} (see section 2.3.5): although every element in $\mathbf{w}$ perfectly complement the sum of the others, removing an element w_k from $\mathbf{w}$ while equally redistributing its mass across the remaining $d-1$ elements yields a new Dirichlet distribution which is completely independent of w_k.

The Dirichlet Distribution as the Conjugate Prior of a Multinomial

It is also well-known that the DD and the multinomial distribution are the conjugate priors of each other [160]. Let $\mathbf{x}$ be a vector of N observed frequencies of occurrence (counts) following a multinomial distribution with parameter $\mathbf{p}$ (here, $\mathbf{p}$ is a proportion vector in the $(d-1)$-simplex specifying the probabilities of observing an item k). Its probability mass function is expressed as:

$$\mathbf{x} \sim Mult(\mathbf{x}|\mathbf{p}) = \frac{(\sum_k x_k)!}{\prod_k x_k!} \prod_k \theta_k^{x_k}. \tag{2.41}$$

If $\mathbf{p}$ follows a DD (i.e., $\mathbf{p} \sim \mathcal{D}(\mathbf{p}|\boldsymbol{\alpha})$), then the posterior distribution of $\mathbf{p}$ given that we observed the counts $\mathbf{x}$ also follows a Dirichlet:

$$\begin{aligned}
p(\mathbf{p}|\mathbf{x}) &= p(\mathbf{x}|\mathbf{p})p(\mathbf{p}) \\
&= Mult(\mathbf{x}|\mathbf{p})\mathcal{D}(\mathbf{p}|\boldsymbol{\alpha}) \\
&\propto \prod_k \theta_k^{x_k} \prod_k \theta_k^{\alpha_k - 1} \\
&\propto \prod_k \theta_k^{\alpha_k - 1 + x_k} \\
&= \mathcal{D}(\boldsymbol{\alpha} + \mathbf{x}).
\end{aligned}$$

This suggests an intuitive strategy for updating our beliefs about proportions evolving over time: start with a Dirichlet prior $\mathbf{p} \sim \mathcal{D}(\boldsymbol{\alpha})$ and just add the observed counts to the *history* of observed frequencies $\boldsymbol{\alpha}$, thus increasing the concentration parameter (α_C) and, as a result, decreasing our uncertainty.

Tracking Changing Probabilities Over Time with Dirichlet Priors

Bertuccelli and How [30] devised a Bayesian dynamic model to track the transition probabilities of a Markov-Chain model evolving in time. By using one Dirichlet prior for each row of the transition matrix, they derived an online, recursive mean-variance Maximum A Posteriori (MAP) approach to estimate the transition matrix as counts on the number of state transitions were collected.

Classical MAP estimators may however fail to respond in a timely manner to changes in the transition matrix because, since the estimators are based on the whole history of observed counts, "*a large number of new transitions will be required for the change detection.*" [30]. The authors have then proposed adding pseudo-noise by heuristically scaling the variance of the Dirichlet distributions, which may help to reduce the dependency of the model on old observations. The recursive MAP equation for updating the posterior mean is as follows [30]:

$$\bar{p_k}(t+1|t+1) = \bar{p_k}(t+1|t) + \Sigma_{kk}(t+1|t)\frac{\delta_{k,k'} - \bar{p_k}(t+1|t)}{\bar{p_k}(t+1|t)(1-\bar{p_k}(t+1|t))}, \tag{2.42}$$

where $\delta_{k,k'}$ is an indicator function for transitions observed from a state k to any new state k', and $\Sigma_{kk} = Var[p_k]$. Moreover, in the absence of a known dynamical model for how the transition probabilities evolve over time, they simply assumed the following prediction steps:

$$\bar{p_k}(t+1|t) = \bar{p_k}(t|t), \text{ and} \tag{2.43}$$

$$\Sigma_{kk}(t+1|t) = \Sigma_{kk}(t|t). \tag{2.44}$$

In this thesis, we make use of the Dirichlet mean MAP recursive estimation to update our beliefs about how trade-off investment proportions evolve over time in financial portfolio applications. One of our contributions is to devise a dynamical model for the mean prediction step of Eq. 2.43), considering a sliding window update approach (see section 2.1.2 and **chapter 6**).

2.4 Summary of the Contributions

This chapter's contributions to the thesis are as follows:

1. Several notions of what it is meant by *uncertainty* throughout the thesis are presented and motivated, as well as practical approaches to model and handle it. Although this concept can be easily mistaken for *risk*, we are sympathetic with the points made by Hybbard [119], that we summarize in the following way: risk = (uncertainty, undesirable outcomes);

2. A robust covariance matrix estimation technique for mitigating model misspecification under a Gaussian assumption was presented in section 2.1.3. This technique is investigated in our preemptive multi-objective metaheuristic proposal presented in **chapter 4**. The algorithm has been designed to search for stable investment portfolios over time, i.e., portfolios for which the expected return and risk estimates change as little as possible between investment environments;

3. Finally, two Bayesian models that are used and adapted in the anticipatory learning tools proposed in **chapter 6** were presented: (1) the Kalman Filter, which is used to keep track of evolving Gaussian distributions over time under Markovian, linear dynamics; and (2) an approach to track the mean of the Dirichlet priors used to represent uncertain, time-varying proportions.

In the next chapter, we provide an overview of what we mean by multi-objective optimization under uncertainty, the central topic of this thesis.

Chapter 3

Multi-Objective Optimization Under Uncertainty

Do not quench your inspiration and your imagination; do not become the slave of your model.

– **Vincent Van Gogh**

The test of a first-rate intelligence is the ability to hold two opposed ideas in the mind at the same time, and still retain the ability to function.

– **Francis Scott Fitzgerald, The Crack-Up (1936)**

This chapter summarizes the main concepts required for a better understanding of the doctorate research scope. In addition, a literature review is presented comprising the historical roots of Multi-Objective Optimization (MOO) under uncertainty. Recent research on Evolutionary MOO (EMOO) in uncertain environments is also covered. From this point forward, we assume minimization of the objective functions, unless indicated otherwise.

3.1 Foundations of Multi-Objective Optimization

The simultaneous optimization of two or more objective functions requires a trade-off analysis over the candidate decisions. This is so because the notion of *optimality* in an MOO problem is not straightforward as in mono-objective optimization. In the latter, a global optimum can be defined as a point $\mathbf{u}^* \in \Omega$ in the set of feasible solutions (search space), $\Omega \subset \mathcal{R}^d$ (d is the number of decision variables), such that $f(\mathbf{u}^*) \leq f(\mathbf{u})$, $\forall\, \mathbf{u} \in \Omega$; whereas, in MOO, we are given a vector of m objective functions, $\mathbf{f}(\mathbf{u}) = [f_1(\mathbf{u}), \cdots, f_m(\mathbf{u})]$, which provides m different evaluations of a candidate decision, $\mathbf{u} \in \Omega$. All $f_1, \cdots f_m$ must be *simultaneously* minimized. If one wants to replace f for $\mathbf{f}$, then the standard order relation $\leq$, which ranks decisions according to their *objective value*, must also be replaced for one that can rank *objective vectors* (a.k.a. performance vectors). An early formulation of an MOO problem appeared in Kuhn and Tucker [141].

While $\leq$ leads to a total order in the real line, it turns out that there is no total order for ranking elements of $\mathcal{Z} \subseteq \mathcal{R}^m$, for $m \geq 2$, unless one makes use of a real-valued (ordered) aggregation function (e.g. weighted sum, min, max, etc.) that specifies the importance of each objective function. This requires a complete *a priori* preference expression by the decision maker (DM)[1], but if this were the case, then we could transform the MOO problem back into a mono-objective one which minimizes the aggregation function.

True MOO problems, thus, deal with optimization in the absence of *a priori* preferences, in which the Pareto Dominance (PD), $\preceq$, is the standard order relation used for ranking objective vectors. The PD is defined as $\mathbf{u} \preceq \mathbf{y} \iff u_i \leq y_i,\ \forall\ i$ and $\exists\ j$ such that $u_j < y_j$. The PD was shown by Voorneveld [217] to be "*(...) the unique nontrivial partial order on the set of finite-dimensional real vectors satisfying* [at the same time] *a number of intuitive properties*", including the axiom of independence of duplicated states, i.e., if $\dim(\mathbf{u}) = \dim(\mathbf{y}) \geq 2$ and there are coordinates i, j for which $i \neq j,\ u_i = u_j$, and $y_i = y_j$, then $\mathbf{u} \preceq \mathbf{y} \iff u_{-i} \preceq y_{-i}$ (the subscript $-i$ notation denotes the removal of the i-th element).

This axiom implies that the PD depends only on objectives with different evaluations, what conveys a notion of ignorance w.r.t. the probability distributions of each variable: if the variables having identical values did matter, the implicit considerations of the probability of occurrence of those values would be required to modify the ranking, but this is not the case for the PD. That is to say, a DM relying on the PD is ignorant w.r.t. the underlying uncertainty regarding the objective vectors. One of the tasks of MOO *under uncertainty*, therefore, is to design a dominance relation possessing *uncertainty awareness*. Before showing how this can be achieved, we present a basic model of MOO for static and deterministic environments:

$$\min\ \mathbf{f}(\mathbf{u}) = (f_1(\mathbf{u}) \cdots f_m(\mathbf{u}))^\intercal\,,\ \text{s.t. } \mathbf{u} \in \Omega, \tag{3.1}$$

in which $\mathbf{f} : \mathcal{R}^d \mapsto \mathcal{R}^m$ and $f_i : \mathcal{R}^d \mapsto \mathcal{R},\ i = 1, \cdots, m$. When solving (3.1), there are no guarantees of finding a unique decision that simultaneously minimizes all objective functions, as this would require no conflict between the objective functions, in the whole domain. Thus, the min operator (or max, when maximizing) is understood as: obtain the set of all decisions for which improving the value of one objective function worsens the value of at least one other objective.

We therefore require an optimality concept that is compliant with PD for revealing the underlying trade-off among the objectives. The Pareto optimality property does just that: it reveals all feasible *non-inferior* (a.k.a. efficient, Pareto optimal, or non-dominated) decisions.

Definition 3.1 (Noninferior decision). *A decision* $\mathbf{u}^* \in \Omega$ *with an associate objective vector* $\mathbf{z}^* \in \mathcal{F}(\Omega)$ *is non-inferior iff* $\mathbf{f}(\mathbf{u}') \not\preceq \mathbf{f}(\mathbf{u}^\star), \forall\ \mathbf{u}' \in \Omega$.

The *set of non-inferior decisions*, $\Omega^\star \subset \Omega$, is also known as the *Pareto Set* (PS), whereas its image in the objective space, $\mathcal{F}^\star = \mathcal{F}(\Omega^\star) \subset \mathcal{F}(\Omega)$, is known as the *efficient frontier*, or the *Pareto Frontier* (PF), which reveals the geometry of the underlying trade-off among all m objective functions.

The idea behind solving MOO problems in Multi-Criteria Decision-Making (MCDM) is to find a diverse set of efficient solutions providing a good representation of the existing trade-off

[1]The DM preferences are encoded into the dominance relation. Even when using an ordered aggregation function (e.g. min, max, etc.) there is an implicit indifference assumption about the importance of each objective.

between the conflicting optimization criteria, so that the decision-making process can be done *a posteriori*. One important advantage of the *a posteriori* approach is that the DM can acquire a better knowledge about the nature of the problem, which is not always available before the visualization of the obtained approximation of the efficient frontier solutions. The DM can thus solve an MOO problem to obtain a finite approximation of the PS, $\mathcal{U}^\star = \{\mathbf{u}_1^\star, \cdots, \mathbf{u}_N^\star\} \subset \Omega^\star$, and then analyze the corresponding objective vectors in $\mathcal{F}(\mathcal{U}^\star) = \{\mathbf{z}_1^\star, \cdots, \mathbf{z}_N^\star\}$, in which $\mathbf{z}_k^\star = \mathbf{f}(\mathbf{u}_k^\star)$, $k = 1, \cdots, N$, for selecting the non-inferior decision that better represents his/her preferences, not directly incorporated into the existing objective functions.

Depending on the geometric properties of the objective functions (e.g. on whether they are convex, quasi-convex or non-convex) and of the feasible region in the search space, the shape of the PF in the objective space can pose significant challenges to classical MOO problem solvers. For instance, several MOO solvers struggle to handle concave or discontinuous PFs. Such solvers often require several runs to approximate the non-inferior solutions set by e.g. optimizing at each run a single objective function, while iteratively parameterizing the remaining optimization criteria as constraints over the objective space, which are difficult to tune unless one has complete information about the feasible region in the objective space.

Therefore, depending on the extent to which each objective function presents multi-modality (i.e., induce several local optima), those classical methods may converge to local PFs. It turns out that most classical MOO solvers do not possess the ability of escaping from such local PFs and, thus, there is often no guarantee that they will manage to converge to the true, global PF, except under very special circumstances, which very rarely are directly found in real-world problems. In general, the MOO solver is required to be stochastic to have a chance of escaping from local optima.

3.1.1 The Hypervolume ($\mathcal{S}$-Metric) Indicator

Let $\mathcal{Z} = \mathcal{F}(\Omega) \subset \mathcal{R}^m$ be the feasible alternatives set in the objective space (with m objective functions) and $\pi(\mathcal{Z})$ the set of all non-empty subsets of $\mathcal{Z}$. Indicators of the form $I : \pi(\mathcal{Z}) \mapsto \mathcal{R}$ can be used to measure the quality of sets [227]. The Hypervolume (Hypv or $\mathcal{S}$-Metric) [10] is a sound theoretical indicator that captures in a single real value the proximity to the true PF, the coverage, and spread of a set of candidate non-inferior decisions along the objective-space.

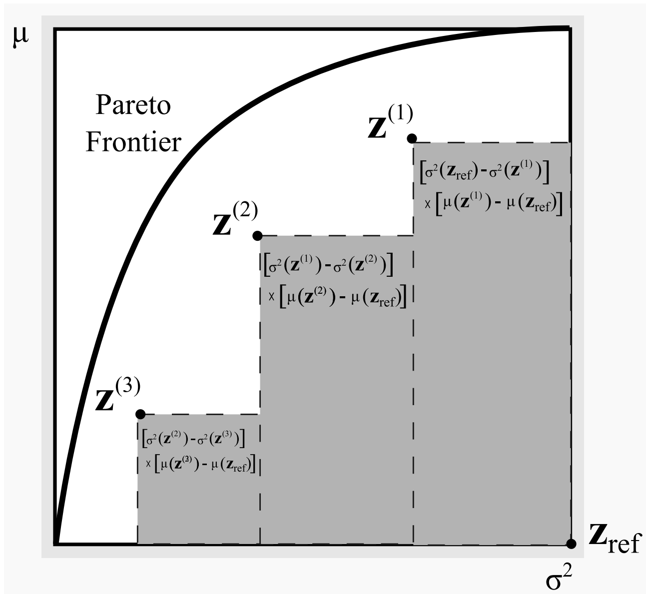
Moreover, $\mathcal{S}$ is the only known indicator that reflects the PD [10], meaning that it is strictly monotonic w.r.t. the PD [86]. In other words, if a finite approximation set A dominates a finite approximation set B, then this fact will be reflected in the hypervolume values of the two sets[2]. Mathematically,

$$A \preccurlyeq B \wedge B \not\preccurlyeq A \Rightarrow \mathcal{S}(A) > \mathcal{S}(B). \tag{3.2}$$

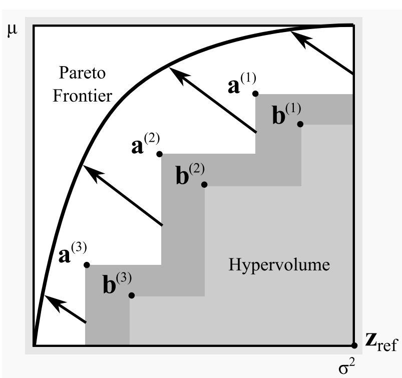
When comparing sets of arbitrary sizes, the following other statements can be made without any further preference information:

1. $\mathcal{A} \lhd \mathcal{B}$ ($\mathcal{A}$ is *better than* $\mathcal{B}$) if $\mathcal{A} \preccurlyeq \mathcal{B}$ and $\mathcal{B} \not\preccurlyeq \mathcal{A}$; or
2. $\mathcal{A} \parallel \mathcal{B}$ (*incomparable*) if $\mathcal{A} \not\preccurlyeq \mathcal{B}$ and $\mathcal{B} \not\preccurlyeq \mathcal{A}$.

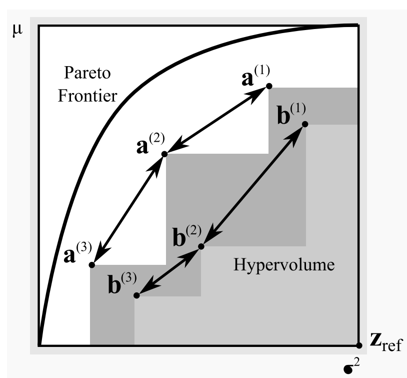
[2]The PD can be easily generalized for comparing two sets of trade-off alternatives: A is said to Pareto-dominate B if every alternative in B is Pareto-dominated by at least one alternative in A.



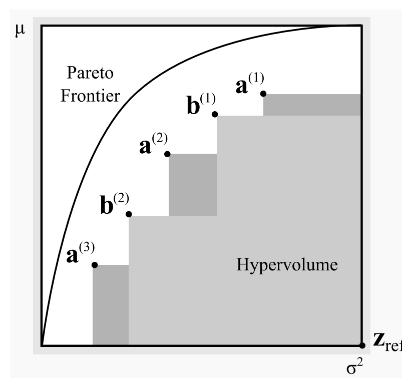
(a) Measured Hypv



(b) Higher proximity



(c) Higher spread



(d) Higher coverage

Figure 3.1: In part (a), the hypervolume of an approximation set of three objective vectors for a mean (μ) maximization/variance (σ^2) minimization bi-objective problem can be computed by summing over the Pareto dominance regions bounded by each alternative, its neighbor in the objective space, and the reference point. The sets of alternatives $\mathcal{A} = \left\{\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \mathbf{a}^{(3)}\right\}$ convey higher (b) proximity to the PF; (c) spread; and (d) coverage than the sets $\mathcal{B} = \left\{\mathbf{b}^{(1)}, \mathbf{b}^{(2)}, \mathbf{b}^{(3)}\right\}$, respectively.

Another property of the Hypv indicator is the so-called $\ntriangleleft$-compatibility [228], i.e.,

$$\forall\ \mathcal{A}, \mathcal{B} \in \pi(\mathcal{Z}) : \mathcal{S}(\mathcal{A}) \geq \mathcal{S}(\mathcal{B}) \Rightarrow \mathcal{A} \ntriangleleft \mathcal{B}, \tag{3.3}$$

what motivates the design of heuristics for maximizing Hypv as a means to improve sets of trade-off options. Given a reference point $\mathbf{z}_{\text{ref}}$, the Hypv of $\mathcal{A}$ is defined as [227]:

$$\mathcal{S}(\mathcal{A}) = \int\limits_{\mathbf{z} \in \mathcal{Z}} \alpha_{\mathcal{A}}(\mathbf{z}) \mathrm{d}\mathbf{z}, \tag{3.4}$$

$$\alpha_{\mathcal{A}}(\mathbf{z}) = \begin{cases} 1, & \text{if } \exists \mathbf{z}' \in \mathcal{A} : \mathbf{z}' \preceq \mathbf{z} \preceq \mathbf{z}_{\text{ref}} \\ 0, & \text{otherwise.} \end{cases} \tag{3.5}$$

Geometrically, the $\mathcal{S}$-Metric measures the hypervolume of the polytope formed by the union of all dominance regions of each $\mathbf{z} \in \mathcal{F}$ ($\mathcal{F}$ is a finite approximation set of objective vectors), bounded by a reference point $\mathbf{z}_{\text{ref}} \in \mathcal{Z}$ and the ideal objective vector (see Fig. 3.1). The indicator can also be described as a composition of mappings of the form

$$\underbrace{\mathcal{R}^d \times \cdots \times \mathcal{R}^d}_{N \text{ decision vectors } (\mathcal{U})} \quad \xrightarrow{\mathbf{f}} \quad \underbrace{\mathcal{R}^m \times \cdots \times \mathcal{R}^m}_{N+1 \text{ objective vectors } (\mathcal{F} \cup \{\mathbf{z}_{\text{ref}}\})} \quad \xrightarrow{\mathcal{S}} \quad \underbrace{\mathcal{R}}_{\text{The hypervolume value}} \tag{3.6}$$

and is computed as

$$\mathcal{S}(\mathcal{F}(\mathcal{U}), \mathbf{z}_{\text{ref}}) = \Lambda \left(\bigcup_{\mathbf{z} \in \mathcal{F}(\mathcal{U})} \{\mathbf{z}' \in \mathcal{Z} : \mathbf{z} \preceq \mathbf{z}' \preceq \mathbf{z}_{\text{ref}}\} \right), \tag{3.7}$$

where $\Lambda = \prod_{i=1}^{m} [a_i, b_i]^m$ is a Lebesgue measure over hyperboxes in $\mathcal{R}^m$. Its maximization over infinite sets of candidate non-inferior solutions has been shown to be equivalent to finding the true PF [86], which has lead to the development of sound hypervolume maximization-based methods such as SMS-EMOA [32] (Section 3.4.1). Before presenting another interesting property of the hypervolume, we define a Nadir point.

Definition 3.2 (Nadir point)**.** *The objective vector $\mathbf{z}_{nad} \in \mathcal{Z}$ is a Nadir point if it is composed of the worst components of the objective vectors representing the extrema of the Pareto Frontier.*

It turns out that, given a finite set of feasible solutions $\mathcal{U}$, if $\mathbf{z}_{\text{ref}} \preceq \mathbf{z}_{\text{nad}}$, then maximizing $\mathcal{S}$ corresponds to finding the Pareto Set $\Omega^\star$ when $|\mathcal{U}| \to \infty$ [86]. It has also been experimentally observed that the maximization of $\mathcal{S}$ for finite sets leads to subsets of the PF presenting high levels of diversity, coverage and spread of points in the objective space [10] (see Fig. 3.1 (b)–(d)).

Because of the theoretical guarantees and implications of the hypervolume maximization for MOO, this thesis devotes substantial emphasis on its usage as a global criterion for automating sequential MCDM. In fact, when maximizing $\mathcal{S}$, we can think of the original MOO problem as a mono-objective optimization problem over sets of mutually non-dominated (or incomparable) decisions. As pointed out by Zitzler et al. [229], *"(...) the search space Ψ consists of all potential Pareto set approximations rather than single solutions, i.e., Ψ is a set of sets"*.

3.1.2 Hypervolume and the convexity of the Pareto Frontier

One important fact concerning the geometry of the PF is that it will be convex if every individual objective f_i are convex functions [85]. Under those conditions, a non-inferior solution $\mathbf{u}^\star \in \Omega^\star$ can be characterized by at least one weight vector $\mathbf{w}$, where $w_j \in [0, 1]$ and $\sum_j^m w_j = 1$, which is associated with the following multi-objective weighted sum formulation [85]:

$$\langle \mathbf{w}, \mathbf{f} \rangle := \min_{\mathbf{u} \in \Omega} \quad \sum_{j=1}^{m} w_j f_j(\mathbf{u}). \tag{3.8}$$

The convexity of the PF in this case allows for the use of scalarization methods for which there are discretization strategies for $\mathbf{w}$ to generate N non inferior decisions, by independently solving N weighted mono-objective problems in the form of Eq. (3.8) [85]. However, one practical difficulty of such scalarization is that a uniform discretization over the weights does not lead to uniformly distributed objective vectors over the PF, except when all objective-functions are linear, in which case the PF is a hyperplane. The maximization of $\mathcal{S}$, on the other hand, is guaranteed to obtain a finite subset of objective vectors belonging to the PF with good additive approximation [41], although the resulting distribution of points is also not guaranteed to be uniform.

3.1.3 Non-convexity of Hypervolume

When one attempts to obtain a finite approximation of the PS (i.e., $\mathcal{U}$) composed of N non-inferior solutions, it is straightforward to verify that the maximization of $\mathcal{S}$ leads to a non-convex problem over Ω^N. One first has to note from Eq. (3.7) that $\mathcal{S}$ can be described by a sum over products of differences between the objective functions evaluated at the ordered elements of $\mathcal{U}$. For instance, considering two objective functions (for $m = 2$, and assuming minimization of both objectives), note that we can write $\mathcal{S}$ as

$$\begin{aligned}\mathcal{S}(\mathcal{U}) = \sum_{i=1}^{N-1} \left[f_1(\mathbf{u}_{i+1}) - f_1(\mathbf{u}_i)\right]\left[z_{2,\mathrm{ref}} - f_2(\mathbf{u}_{i+1})\right] \\ + \left[z_{1,\mathrm{ref}} - f_1(\mathbf{u}_1)\right]\left[z_{2,\mathrm{ref}} - f_2(\mathbf{u}_1)\right].\end{aligned} \tag{3.9}$$

In other words, following the direct application of the Lebesgue measure in Eq. (3.7), the Hypv computation can be decomposed by summing over the areas (products of the differences between individual objective values) of the dominance regions that are exclusively contributed by neighboring pairs of objective vectors, following a lexical order in the objective space (see Fig. 3.1 (a)).

It then becomes clear that the product of differences violates the conditions for $\mathcal{S}(\mathcal{U})$ to be a convex function in Ω^N, even when each individual f_j is a convex function. This is because the product of two non-negative convex functions, say, g and h, of the form $h \circ g : \mathcal{R}^n \mapsto \mathcal{R}$, is convex iff $(g(\mathbf{x}) - g(\mathbf{y}))(h(\mathbf{x}) - h(\mathbf{y})) \geq 0$ for each $\mathbf{x}, \mathbf{y}$ in the domain of g and h [37,63]. Thus, for $\mathcal{S}(\mathcal{U})$ to be convex, it would be required for each f_j to not conflict with any other objective function (i.e., the signs of the gradient vectors components, ∇f_j, should be the same coordinate-wise) and, hence, the Pareto Set associated with the MOO problem would contain the unique global solution, common to each individual optimization criteria, i.e., $\mathbf{u}^* = \mathbf{u}_j^*\ \forall\ j$, where $\mathbf{u}_j^* = \arg\min_{\mathbf{u}} f_j(\mathbf{u})$ and $\mathbf{u}^*$ is the unique non-inferior solution to the MOO problem. It is evident, however, that, because of the non-convexity of $\mathcal{S}$, sophisticated problem-solving techniques are required for hypervolume maximization, such as biologically-inspired metaheuristics, which turns out to be the technique of choice of this thesis (see section 3.4).

3.1.4 Flexibility in Multiple Criteria Decision-Making

In the following, we explore a logical chain of results connecting three key concepts henceforth discussed: maximal Hypv $\Rightarrow$ PD compatibility $\Rightarrow$ preference for flexibility. The set-based extension of the PD is defined as follows [227].

Definition 3.3 (Set-based Pareto Dominance). *For two sets of alternatives $\mathcal{A}, \mathcal{B} \in \pi(\mathcal{Z})$, $\mathcal{A}$ is said to (Pareto) dominate $\mathcal{B}$ ($\mathcal{A} \preccurlyeq \mathcal{B}$) iff* $(\forall \mathbf{z}^{(b)} \in \mathcal{B}), (\exists \mathbf{z}^{(a)} \in \mathcal{A}) : \mathbf{z}^{(a)} \preceq \mathbf{z}^{(b)}$.

In decision theory, no rational DM will deliberately choose elements in the set of dominated options $\bar{E}_A = \{\mathbf{z} \in \mathcal{A} : \exists \mathbf{z}' \in \mathcal{A}, \mathbf{z}' \preceq \mathbf{z}\}$ [126]. Hence, the set of incomparable options that are not dominated by any other option is $\mathcal{A}^\star = A \setminus \bar{E}_{\mathcal{A}}$. Removing $\bar{E}_{\mathcal{A}}$ from $\mathcal{A}$ does not reduce flexibility, whereas removing any $\mathbf{z}^\star \in \mathcal{A}^\star$ from $\mathcal{A}$ does reduce it, and, thus, we regard $\mathcal{A}^\star$ as the *essential set* of $\mathcal{A}$ [169].

Flexibility of Set-Based PD

Recall that $\pi(\mathcal{Z})$ denotes the set of all non-empty subsets of $\mathcal{Z}$. We now show that the set-based PD is consistent with *preference for flexibility* [8]:

Definition 3.4 (Flexibility Consistency). *Let P be a binary relation on $\mathcal{Z}$ ($\mathbf{z}^a P \mathbf{z}^b$ means that $\mathbf{z}^a$ is preferred to $\mathbf{z}^b$). A relation $R \subseteq \pi(\mathcal{Z}) \times \pi(\mathcal{Z})$ is P-consistent with flexibility if:*

$$\forall \mathcal{A} \in \pi(\mathcal{Z}), \forall \mathbf{z} \in \mathcal{Z} \setminus \mathcal{A} \begin{cases} \mathcal{A} \cup \{\mathbf{z}\} \parallel \mathcal{A} \Leftrightarrow \exists \mathbf{z}' \in \mathcal{A} : \mathbf{z}' P \mathbf{z}, \\ \mathcal{A} \cup \{\mathbf{z}\} \, R \, \mathcal{A} \text{ otherwise.} \end{cases}$$

Def. 3.4 establishes that adding a new alternative $\mathbf{z}$ to a set $\mathcal{A}$ is only preferable if $\mathbf{z}$ is essential to $\mathcal{A}$. Moreover, the relation R on $\pi(\mathcal{Z})$ is P-consistent with preference for flexibility iff it satisfies the following axioms [8]:

Axiom 3.1.1 (Restricted Monotonicity). $\forall \mathbf{z}, \mathbf{z}' \in \mathcal{Z}, \mathbf{z} P \mathbf{z}' \Rightarrow \{\mathbf{z}, \mathbf{z}'\} \parallel \{\mathbf{z}\}$ *and* $\neg(\mathbf{z} P \mathbf{z}') \Rightarrow \{\mathbf{z}, \mathbf{z}'\} R \, \{\mathbf{z}\}$.

Axiom 3.1.2 (Incomparability Consistency). $\forall \mathbf{z} \in \mathcal{Z}, \forall \mathcal{A}, \mathcal{B} \in \pi(\mathcal{Z}) : \mathcal{A} \subseteq \mathcal{B}, \mathcal{A} \cup \{\mathbf{z}\} \parallel \mathcal{A} \Rightarrow \mathcal{B} \cup \{\mathbf{z}\} \parallel \mathcal{B}$.

Axiom 3.1.3 (Preference Consistency). $\forall \mathbf{z} \in \mathcal{Z}, \forall \mathcal{A}, \mathcal{B} \in \pi(\mathcal{Z}) : (\mathcal{A} \cup \{\mathbf{z}\} R \, \mathcal{A} \wedge \mathcal{B} \cup \{\mathbf{z}\} R \, \mathcal{B}) \Rightarrow \mathcal{A} \cup \mathcal{B} \cup \{\mathbf{z}\} R \, \mathcal{A} \cup \mathcal{B}$.

Axiom 1 states that the DM is indifferent to adding a non-essential alternative to a singleton, but enjoys enlarging it when the original element is not inferior to the new one; Axiom 2 states that adding a non-essential $\mathbf{z}$ to $\mathcal{A}$ implies it is also non-essential to a set containing A; whereas axiom 3 states that adding an essential $\mathbf{z}$ to a pair of sets implies it is also essential to their union [8]. The orders PD $\preceq$ (P) and the set-based PD $\preccurlyeq$ (R) satisfy axioms 1–3 and, thus $\preccurlyeq$ is $\preceq$-consistent with flexibility, according to Def. 3.4.

Flexibility of the Hypervolume Indicator

If Hypv comparisons are used in a relation $\preccurlyeq_{\mathcal{S}}$ in $\pi(\mathcal{Z}) \times \pi(\mathcal{Z})$, $\preccurlyeq_{\mathcal{S}}$ can also be verified to be $\preceq$-consistent with flexibility. In fact, $\preccurlyeq_{\mathcal{S}}$ refines $\preccurlyeq$ [227], i.e., ranking according to $\mathcal{S}$ values

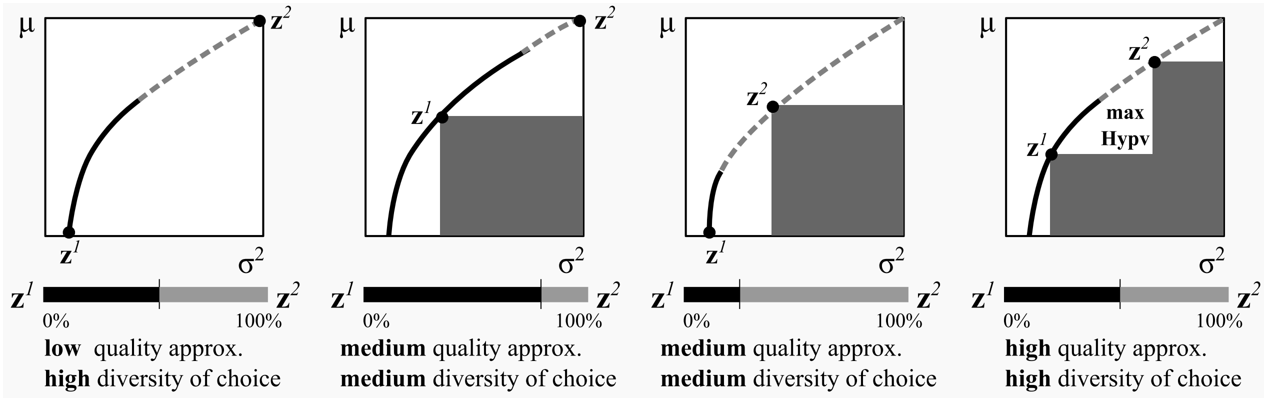


Figure 3.2: PF partitioning associated with the placement of two mutually non-dominated alternatives. In the leftmost figure, the Hypv is minimal when computed from the extrema alternatives and, hence, the approximation quality is the lowest among all possible distributions of two points lying in the PF. On the other hand, diversity of choice in this case is high because both alternatives have approximately the same likelihoods of being chosen when considering the universe of all possible DMs preference specifications along the PF. The optimal μ-distribution obtained with maximal Hypv in the rightmost figure not only achieves high approximation quality, but also high diversity of choice.

refines the partial order obtained by $\preccurlyeq$, in which case Hypv also satisfies axioms 1–3. Generally, any indicator I refines $\preccurlyeq$ if the following holds (see Theorem 3.1 in [227]):

$$\mathcal{A} \preccurlyeq \{\mathbf{z}\} \Rightarrow I(\mathcal{A} \cup \{\mathbf{z}\}) = I(\mathcal{A}) \text{ and} \tag{3.10}$$

$$\mathcal{A} \not\preccurlyeq \{\mathbf{z}\} \Rightarrow I(\mathcal{A} \cup \{\mathbf{z}\}) > I(\mathcal{A}), \tag{3.11}$$

i.e., adding an essential option improves I, but adding a non-essential one does not alter its value.

3.1.5 On Optimal μ-Distributions and Diversity of Choice

Maximal Hypv has been investigated in terms of the distributions of μ alternatives over the PF [9,31]. Note that the set attaining maximal $\mathcal{S}$ value is a subset of the PF [86]. One question is how the μ-distributions over the PF are related with Kumar's [142] notion of diversity of choice. It turns out that the goal of maximal Hypv is to achieve a μ-distribution representative of all admissible choices. Suboptimal Hypv μ-distributions can leave vast portions of the PF underrepresented. The point of taking decisions to achieve future maximal Hypv μ-distributions is thus to have higher quality approximation of the PF with higher diversity of choice for future preference specifications (see Figs. 5.5 and 3.2).

As it can be noted from Fig. 3.2, the Hypv maximization favors distributions of options lying on knee regions of the Pareto Frontier, thus avoiding the extrema.

3.2 Stochastic Multi-Objective Optimization

Stochastic MOO (SMOO) problems can be modeled in a similar way as that in (3.1):

$$\min\ \mathbf{f}(\mathbf{u},\xi) = (f_1(\mathbf{u},\xi)\cdots f_m(\mathbf{u},\xi))^{\intercal},\ \text{s.t. } \mathbf{u}\in\Omega, \tag{3.12}$$

where ξ is a vector of random variables (r.v.'s). In most cases, ξ will represent an exogenous noise process. Note that, as a consequence of each objective function being dependent on ξ, each objective value is also a r.v., and, hence, $\mathbf{z} = \mathbf{f}(\mathbf{u},\xi)$ is a random vector. It is often assumed that Ω is still a deterministic feasible set, although chance constraints can also be incorporated as a proxy for integrating the DM attitude to risk into the SMOO model.

It should be noted that the model in Eq. (3.12) is an abstraction, without a precise mathematical meaning, i.e., it does not specify how multiple, conflicting, stochastic objective functions should be simultaneously optimized. Abdelaziz [1] identifies two reduction techniques for solving Eq. (3.12) in the literature: the so called multi-objective method and the stochastic method. In the former, the SMOO problem is reduced to an MOO one, whereas, in the latter, it is reduced to a Stochastic Optimization (SO) one. One example of the SMOO $\rightarrow$ MOO reduction is:

$$\min\ (\mathbb{E}\left[f_1(\mathbf{u},\xi)\right]\ \cdots\ \mathbb{E}\left[f_m(\mathbf{u},\xi)\right])^{\intercal},\ \text{s.t. } \mathbf{u}\in\Omega, \tag{3.13}$$

where one can approximate the expected PF by taking e.g. the sample means of each objective function over a certain number of repeated evaluations of $\mathbf{u}$. Another possibility is to minimize the second moment of the distribution of the objective functions as well, yielding the model

$$\min\ (\mathbb{E}\left[f_1(\mathbf{u},\xi)\right]\ \cdots\ \mathbb{E}\left[f_m(\mathbf{u},\xi)\right]\ \mathrm{Var}\left[f_1(\mathbf{u},\xi)\right]\ \cdots\ \mathrm{Var}\left[f_m(\mathbf{u},\xi)\right])^{\intercal},\ \text{s.t. } \mathbf{u}\in\Omega. \tag{3.14}$$

A more general model considers minimizing r functionals $\mathcal{V}_j^{(1)}\left[f_j(\mathbf{u},\xi)\right],\cdots,\mathcal{V}_j^{(r)}\left[f_j(\mathbf{u},\xi)\right]$ for each objective function $j = 1,\cdots,m$, which can be either moments of the probability distribution (expectation, variance, skewness, etc.) or e.g. risk measures. Cardoso et al. [50] have generalized this approach by using Monte Carlo simulation to approximate the objective vectors distributions subject to non-Gaussian disturbances. The approach was utilized so that the decision maker can base his/her decisions on any quantile of interest.

The SMOO $\rightarrow$ SO reduction, on the other hand, works by means of aggregation functions, and may take, for instance, the following forms:

$$\max\ u\left(\mathbb{E}\left[f_1(\mathbf{u},\xi)\right],\cdots,\mathbb{E}\left[f_m(\mathbf{u},\xi)\right]\right),\ \text{s.t. } \mathbf{u}\in\Omega, \text{or} \tag{3.15}$$

$$\max\ \mathbb{E}\left[u\left(f_1(\mathbf{u},\xi),\cdots,f_m(\mathbf{u},\xi)\right)\right],\qquad \text{s.t. } \mathbf{u}\in\Omega, \tag{3.16}$$

in which $u : \mathcal{R}^m \mapsto \mathcal{R}$ is the DM utility function (to be maximized). If u is linear, then, due to the linearity of the expectation operator, the two formulations are equivalent.

An alternative SMOO $\rightarrow$ SO reduction model that is still unexplored in the literature is the explicit maximization of the joint cumulative distribution function over the objective space such as, e.g.,

$$\max \Pr\left[\mathbf{f}(\mathbf{u},\xi) \preceq \eta\right],\ \text{s.t. } \mathbf{u}\in\Omega, \tag{3.17}$$

in which η is some desired benchmark performance vector.

For instance, Azevedo and Araújo [13] have proposed a kernel-based method for estimating the joint distribution of a population of candidate sets of non-inferior solutions in the objective space by means of multivariate Pareto order statistics, although the estimated distribution (the so called Non-Dominance Landscape, NDL) was secondary to the optimization process. This approach is promising for heuristically solving SMOOPs, mainly because the estimated distribution allows for the direct computation of the Pareto Non-Dominance Probability (PNDP) of any given solution – i.e., the probability of $\mathbf{u}$ being a non-inferior solution relative to a fixed population of candidate mutually non-dominated solutions, $\mathcal{U}$ [13].

Coelho and Bouillard [57] have instead used the joint probability approach as a constraint in a reliability-based model, while Coelho [56] has used it within an intricate SMOO model, solved with a co-evolutionary metaheuristic:

$$\text{s.t.} \quad \begin{cases} \min_{\mathbf{u},\eta} \eta \\ \Pr\left[\mathbf{f}(\mathbf{u},\xi) \preceq \eta\right] \geq \alpha, \end{cases} \tag{3.18}$$

where both the decision variables, $\mathbf{u}$, and the "*quantiles*", η, must be simultaneously tuned. Assuming minimization, the goal was to search for the lowest quantile (or benchmark performance) which can be dominated by a solution to be searched for in the search space, while guaranteeing that the probability of dominance stays at a minimum, pre-defined critical level, α. While the SMOO $\rightarrow$ SO reduction can be straightforward, it is not suitable for treating undefined preferences, in which one is interested in approximating the PF.

The contribution of this thesis in terms of SMOO modeling, thus, is to show how a higher-level set-based optimization meta-model for maximizing the expected hypervolume over sets of non-inferior solutions can preserve the guarantees of approximating the whole PF with arbitrary accuracy. The thesis also shows how to incorporate uncertainty-awareness into metaheuristics for maximizing $\mathbb{E}\left[\mathcal{S}\right]$ (see **chapter 6**, section 6.3, Theorem 6.3.1).

3.3 Sequential Decision-Making

When time is taken into account, the most simple and intuitive problem-solving strategy is to break the problem into *decision periods* over time and to solve them *independently*, as separate and non-interacting entities. This strategy is known as the *myopic* approach. The limitations arising from the myopic strategies become evident when there are time-linkages between the decision periods, i.e., when some or all decisions taken in the past directly influence the optimality properties of the problem in future periods. This issue was recognized in the late 1950's by Bellman [23], who devised the so called *optimality principle* from a set of recurrence equations governing the right decomposition of decision sequences so as to guarantee that the decisions taken at any period will allow for optimal decisions to be found in subsequent ones.

This has resulted into the development of two important research fields: (i) in computer science, Bellman equations are used to derive an optimal problem decomposition strategy for solving constructive, sequential problems in discrete spaces that can allow for a complete solution being represented as a path in a directed acyclic graph [61]; and (ii) in control engineering, the equations are used to find a unique control law, or *policy*, which can guarantee that each and every decision (or control action) taken at any decision period are optimal. This is possible because Bellman's equation convey a way to track all relevant information flow about future states which allow for an optimal decision at the current state [168].

3.3.1 Optimal Sequential Decision Models

When it comes to applications for which a sequence of decisions must be taken over time, there are three mainstream models of optimal behavior that are commonly adopted in the *reinforcement learning* literature, which is rooted in Bellman's dynamic programming approach [127]: (1) the *finite horizon*; (2) the *discounted infinite horizon*; and (3) the *average-reward* model.

The first model is the most simple one:

$$\min_{\mathbf{u}} \mathbb{E}\left[\sum_{t=1}^{H} f_t(\mathbf{u})\right], \tag{3.19}$$

where the value of the decisions are additive, i.e., it is a sum of the objective function value at each period t. Implicit in this model is the assumption that the DM is indifferent w.r.t. shorter and longer-term rewards, since the importance of $f_t(\mathbf{u})$ is constant over $t = 1, \cdots, H$ (H is the number of decision periods). At the first step, the DM must ideally take the so called *H-step optimal decision*, that is, the best possible decision knowing that there are H remaining decisions to be taken, and so forth. Note that the value of H limits how far ahead the DM should look into the future in order to take a decision. Clearly, if the DM does not have access to predictive knowledge regarding the probability distribution of the future values of f_t, then the decision may not be optimal.

The second, discounted model is the most common in the literature:

$$\min_{\mathbf{u}} \mathbb{E}\left[\sum_{t=1}^{\infty} \gamma^t f_t(\mathbf{u})\right]. \tag{3.20}$$

Note that the model is discounted by $\gamma \in [0, 1)$, what can be interpreted as a way of modeling impatience: the closer γ is to zero, the more eager the DM is to take decisions that perform well in the shorter-term. As pointed out by Kaelbling et al. [127], γ also conveys a trick to bound the infinite sum, making the model more tractable when compared to the finite-horizon case.

The last mainstream model is the limiting case of the infinite model when the discount factor approaches one:

$$\min_{\mathbf{u}} \lim_{H\to\infty} \mathbb{E}\left[\frac{1}{H}\sum_{t=1}^{H} f_t(\mathbf{u})\right], \tag{3.21}$$

where the notion of indifference of the DM regarding shorter or longer-term rewards is even stronger then in the finite-horizon model.

The point that is most worth noting about the optimal models in sequential decision-making problems is the fact that very different sequences of decisions can be obtained, depending on the model choice. Because in this thesis we are dealing with vector-valued functions, none of the aforementioned models can be directly used for MCDM. Moreover, due to the fact we are dealing with the interesting case of undefined preferences over the optimization criteria on the DM's end, we will also assume that the DM cannot in principle reveal her preference regarding the discount factor choice. In fact, we will later argue that one principled way to deal with this issue is to set the discount factor automatically, as a function of the perceived uncertainty which is conveyed in the available predictive knowledge, regarding the future decision periods.

3.3.2 Sequential Multiple Criteria Decision-Making

Guaranteeing optimality properties in decision-making under multiple conflicting objectives is, however, not intuitive, unless the preferences of the DM are fully specified beforehand, reducing the problem back to a mono-objective one. As a problem solving technique, the rigorous mathematical treatment concerning MOO for sequential problems began in the early 1970's, following the work of Klötzler [133] and his collaborators. It became evident from that work the difficulty in proving Bellman optimality for a vector-valued objective function, since usual rules would generally fail when dealing with sets of non-dominated decisions. The Bellman-Pareto equations for the sequential SMOO/MCDM case were derived for one-step transition probabilities after Furukawa [92], who defined an m-dimensional vector-valued function evaluated when going from states i to j, $f_{ij}^{u} : \mathcal{N} \times \Omega \times \mathcal{N} \mapsto \mathcal{R}^m$, in which $\mathcal{N}$ is a countable set of states at which the system may be. Note how the underlying time-linkage is explicitly encoded in f_{ij}^{u}. This was an important result because it showed, under a one-step time-linkage (also known as the *Markov assumption*), the existence of optimal decision-making policies allowing the DM to achieve a non-inferior optimal set of Pareto optimal decisions for each decision period. Three years later, Nollau [161] extended the Bellman-Pareto equation for general state transition functions in sequential SMOO/MCDM. Finally, Sukulimowskil [192] addressed the solution to an MOO/MCDM problem by predicting the consequences of implementing each non-dominated decision from the Pareto set as to the future Pareto Sets that could be obtained.

Bellman equations for dynamic SMOO under the Markov Decision Processes framework began to emerge in the late 1970's through the early 1980's, after various incremental works of Viswanathan et al. [215], Furukawa [91, 92], White [219], Hening [116], and others, both for finite and infinite decision horizons. While those early works set the theoretical ground for exactly solving sequential SMOO problems under general dynamic and stochastic models, they have not lead to effective problem-solving tools. In fact, the combinatorial explosion of the search space as the number of states and decision periods grow (for discrete state-spaces) makes the practical application of such optimality guarantees prohibitive [168]. That is why research activity on various approximation techniques became intense during the last decades.

In Roijers et al. [173], it can be found the latest up-to-date survey on sequential MCDM (assuming an infinite number of decision periods). It brings one of the first attempts to discuss *in a more thoughtful way* in which circumstances a decision-making process should be approached as a multi-criteria one, although the discussion implicitly assumes that the DM preferences can be modeled as a linear scalarization function parameterized by a weight vector. The three scenarios analyzed by Roijers et al. [173] are:

1. The *unknown weights* scenario, in which scalarization is impossible before the optimization but is trivial when a decision needs to be taken;

2. The *decision support* scenario, where scalarization is impossible at all because it can be

too difficult to specify the weights or the precise formulation of the objective functions. This can be the case, e.g., when taking a group decision involving a committee of DMs; and

3. The *known weights* scenario, for which scalarization is possible before the optimization is performed, although, if the objective functions are not convex, the chosen weights may not guarantee that a valid non-inferior decision will be found.

3.3.3 Hypervolume-based Sequential Multiple Criteria Decision-Making

There have been some initial attempts of adopting the hypervolume indicator as a proxy for obtaining non-inferior *policies* and decisions in Multi-Objective Reinforcement Learning (MORL). Those developments are very recent, and, therefore, contemporary to the research reported in this thesis. Wang and Sebag [218] designed a MORL algorithm based on the Monte-Carlo Tree Search (MCTS) method, which consists on searching for non-inferior online policies given the current state of the decision process. The tree is explored heuristically by selecting the decisions (actions) that maximally improve the hypervolume value computed over the upper confidence bounds on the multiple conflicting reward functions over all the history of decisions taken so far. In this way, the decision process converges in the long run to a learned non-inferior policy of maximal hypervolume (in the case of the MCTS, merely a sequence of decisions).

A very similar approach is taken in the recent paper of Van Moffaert et al. [213]. Instead of using a MCTS method, a standard offline Q-learning approach using the same hypervolume contribution reward idea as of Wang and Sebag [218] for selecting a decision at each period was employed. As in Wang and Sebag [218], the hypervolume-based MORL Q-learning algorithm was benchmarked on standard RL toy-problems, such as Deep Sea Treasure, and it has outperformed competing scalarization methods with different discretization strategies in terms of spread, coverage and other qualitative measures of the resulting non-inferior policies.

Both approaches, however, require several runs of the MORL algorithm so that the DM can then analyze the non-inferior policies produced and select the one that is best suited to her preferences. Van Moffaert et al. [213] pointed out that

> *"When the policies obtained by different runs of the algorithm are collected, the user can still make her/his decision on which policies or trade-offs are preferred, but the advantage is that emphasis on particular objectives is not required beforehand."*

It is worth mentioning that the approach taken in this thesis for hypervolume-based MCDM (see **chapter 5**) is completely distinct in principle from the aforementioned contemporary MORL methods. Because we are not interested in obtaining *non-inferior policies*, our methods do not require offline training phases. Actually, our approach much more resembles the multi-objective model based predictive control system of Butans [46], who proposed the so called Pareto Decision Tree (PDT) algorithm for selecting a control action maximizing the hypervolume of the next time step for a multi-objective inverted pendulum control problem. A bottleneck of this approach, however, is the exponential complexity involved in performing re-optimizations. Moreover, the dynamic model was perfectly matched to the controlled deterministic system. In contrast, Bayesian tracking in AS-MOO not only avoids re-optimizations, but also handles uncertainty when the dynamics can be *approximated* by a linear Gauss-Markov assumption.

3.4 Multi-Objective Metaheuristics

When it comes to MOO methods for approximating the PF, evolution-inspired metaheuristics [97] densely populate the state of the art. Evolutionary MOO techniques for coping with noisy and dynamic environments (but not with both at the same time) began to emerge in the early 2000's after Farina et al. [84], Teich [205], Hughes [121], and others. The idea is to take advantage of the *implicit parallelism* [44] inherent to Evolutionary Algorithms (EAs) in order to guide a diverse set of candidate non-inferior decisions toward the PF.

Multi-Objective Evolutionary Algorithms (MOEAs) intend to find in a single run (due to their population-based nature) a finite subset of solutions in Ω^*, diverse enough to represent the entire PF in the objective space. Despite convergence to the true PF with MOEAs be only guaranteed in probabilistic terms, the motivation for using them to solve MOO problems is to take advantage of their stochastic behavior and of the Pareto dominance information over the objective space as the main driving force to guide the search towards the PF [66]. By doing so, one can often obtain a good finite set of mutually non-inferior solutions which tends to be at the same time (*i*) as close as possible to the PF (i.e., attempts to minimize a set-based distance metric in the objective space between the approximated efficient solution set to the PF); (*ii*) as evenly spaced as possible over the objective space; and (*iii*) covering as much as possible the whole extent of the PF hypersurface in the objective space, from one extreme to the others.

3.4.1 The Fast Non-Dominated Sorting Genetic Algorithm

The Fast Non-dominated Sorting Genetic Algorithm (NSGA-II) [66] is one of the most popular choices among MOEAs when it comes to solving real-world problems, mainly because of its ease of implementation; its well-documented results in the literature when applied to several engineering problems; and its algorithmic efficiency. The NSGA-II is a Darwinian evolution-inspired meta-heuristic consisting of several cycles (generations) comprising the following procedures:

1. A ranking procedure over the candidate solutions based on the Pareto dominance;

2. A local density estimation procedure known as crowding distance estimation;

3. A selection procedure based on binary tournaments among candidate solutions, in which one solution wins the tourney if it has the best rank or if it is located in a less crowded region in the objective space when compared with solutions belonging to the same ranks;

4. The probabilistic application of stochastic operators to generate new candidate solutions by crossing over existing solutions selected by independent binary tourneys and by mutating the resulting offspring solutions; and

5. A survival procedure based on truncate selection in which the population of candidate solutions is kept at a constant size from generation to generation by eliminating all offspring

1: Initialize $\mathcal{P}_0 = \{\mathbf{u}_p\}$ with N candidate decisions and set $t \leftarrow 0$
2: **repeat**
3: Generate a new decision $\mathbf{u}'$ from $\mathcal{P}_t$
4: Set $\mathcal{P}_{t+1} \leftarrow \text{Reduce}(\mathcal{P}_t \cup \{\mathbf{u}'\})$ and $t \leftarrow t+1$
5: **until** Termination condition is satisfied

Pseudocode 1: $\mathcal{S}$-Metric Selection EMOA

1: Compute $\{\mathcal{C}_1, \cdots, \mathcal{C}_C\}$ by non-dominated sorting $\mathcal{Q}$ [32]
2: Select $\mathbf{z}_r = \arg\min_{\mathbf{z} \in \mathcal{C}_C} \Delta\mathcal{S}(\mathbf{z}, \mathcal{C}_C)$
3: **return** $\mathcal{Q} \setminus \mathbf{z}_r$

Pseudocode 2: Reduce($\mathcal{Q}$)

and parent solutions which are worst ranked (with the crowding distance estimation values serving as tie-breakers).

When provided with suitable stochastic heuristic search operators and parameter settings, this simple algorithmic outline often leads to superior performance over classical MOO solvers, as extensively reported in the literature for challenging benchmark and real-world problems [66].

3.4.2 $\mathcal{S}$-Metric Selection Evolutionary Algorithm (SMS-EMOA)

The $\mathcal{S}$-Metric Selection Evolutionary Multi-Objective Algorithm (SMS-EMOA) [32] is an indicator-based metaheuristic designed for maximizing the hypervolume ($\mathcal{S}$). As the NSGA-II, it works by partitioning the population of candidate decisions into C classes $\mathcal{C}_1, \cdots, \mathcal{C}_C$, so that elements belonging to the first class ($\mathcal{C}_1$, the class of mutually non-dominated decisions) are not (Pareto) dominated by any other decision and, thus, are always preferred over decisions from the other classes. In general, given a class $\mathcal{C}_j$,

$$(\nexists\ \mathbf{u} \in \mathcal{C}_j), (\mathbf{u} \preceq \mathbf{v},\ \mathbf{v} \in \bigcup_{i=1}^{j-1} \mathcal{C}_i). \tag{3.22}$$

When two decisions belong to the same class, the one yielding the highest $\mathcal{S}$-Metric contribution ($\Delta_{\mathcal{S}}$) to its class is preferred. The $\Delta_{\mathcal{S}}$ value is simply the measure of how much a given objective vector associated to a certain decision contributes to its class hypervolume:

$$\Delta_{\mathcal{S}}(\mathbf{z}, \mathcal{C}) = \mathcal{S}(\mathcal{C}) - \mathcal{S}(\mathcal{C} \setminus \{\mathbf{z}\}). \tag{3.23}$$

By iteratively replacing the least preferred solutions with novel decisions generated via heuristic operators, the hopes are that the maximization of $\mathcal{S}$ can be promoted. Beume et al. [32] argued that $\mathcal{S}(\mathcal{P}_{t+1}) \geq \mathcal{S}(\mathcal{P}_t)$ holds after one iteration of SMS-EMOA. The algorithmic description of SMS-EMOA is given in Pseudocodes 1 and 2.

Until this point, MOEAs operating in stationary and deterministic environments have been discussed. In the following, MOEAs that are capable of operating in time-varying environments are reviewed.

3.4.3 Overview of Reactive and Preemptive Approaches

Previous studies on how metaheuristics should behave in face of environmental changes diverge, though. We can distinguish from three different approaches: (*i*) reactive ones, that only take action for handling changes upon detection e.g. by either raising mutation rates or lowering selective pressure [223]; (*ii*) preemptive[3] (myopic) ones which, having no knowledge regarding when or with which intensity the next change will occur, attempt to alleviate their consequences by e.g. generating diversity throughout all optimization process [12] or through problem-specific ad-hoc strategies [38]; and (*iii*) predictive ones, which try to guess when the next change will occur and how intense it will be by e.g. using regression models [190].

One early example of an ad-hoc preemptive single-objective metaheuristic considered the problems of online task scheduling and vehicle routing. For the first problem, Branke and Mattfeld [38] designed preemptive heuristics for incorporating on their solutions the principle that if the tasks requiring the highest execution times are scheduled for the first production periods (but still respecting priority constraints), then the resulting solutions would exhibit higher *flexibility* upon unexpected changes in the production plan. Note, however, that such strategy is problem dependent and cannot be directly transfered to different problems. As for the second problem [39], the ad-hoc preemptive strategy was to increase the waiting time of the vehicles departing from stations, thus making the routing plan more robust to unexpected rises in passengers demands.

Since such reactive and preemptive approaches do not operate upon the knowledge of the probabilities and the severity with which such atypical disruptive events might occur, it is not always possible for them to foresee the long-term influence of such policies in the system performance.

Multi-Objective Reactive/Preemptive Algorithms

Farina et al. [84] discussed the challenges posed by temporal changes for approximating the dynamic PF. Besides, they grouped sequential MOO problems into four categories, based on whether the PF and the PS would change over time (t):

- Type I – PS changes with t, whereas PF does not;
- Type II – both PS and PF change with t;
- Type III – PF changes, whereas PS does not;
- Type IV – Neither PS or PF change with t.

More precisely, the sequential MOO problems of Type I are those for which, whenever the objective function evaluation changes between subsequent decision periods (i.e., $\mathbf{f}_t(\mathbf{u}) \neq \mathbf{f}_{t+1}(\mathbf{u})$

[3]Our usage of the term "preemptive" should not be mistaken with the traditional usage in computer science, mainly in operational systems, which refers to privileged computational tasks possessing the ability of interrupting and later resuming other tasks. Instead, we use the strict dictionary meaning found e.g. in the Wiktionary.com, which refers to actions "*Made so as to deter an anticipated unpleasant situation*", or e.g., that found in the Thesaurus dictionary, which refers to actions "*designed or having the power to deter or prevent an anticipated situation or occurrence*".

for some fixed $\mathbf{u} \in \Omega$), although the decision vectors that are mapped to the PF may also change (i.e., $\Omega_t^* \neq \Omega_{t+1}^*$), the objective vectors in $\mathcal{F}_t(\Omega_t^*)$ are the same as in $\mathcal{F}_{t+1}(\Omega_{t+1}^*)$; and so forth. Farina et al. [84] designed a set of numerical sequential MOO benchmarks for each category to facilitate comparisons between new dynamic MOEAs (DMOEAs). With regard to the design of DMOEAs specifically suited to changing environments, little work has been done in the literature and only recently the first works discussing e.g. "How performance should be measured for comparing sequential MOEAs?" have appeared [113].

One well-known reactive DMOEA is due to Goh and Tan [96], who presented a competitive/cooperative paradigm for sequential MOO. The authors have adopted a Latin hypercube sampling process to generate diversity in the search space. The sampled solutions, called "*stochastic competitors*", are generated whenever environmental changes are detected. The contribution of the random solutions subpopulation is evaluated under a competition mechanism in which contestants are chosen among other evolving subpopulations. If the stochastic competitors emerge as the winner subpopulation, the defeated subpopulation is restarted in the region in which the random solutions were sampled. In addition, whenever change is detected, part of the non-dominated solutions before the change is stored in memory. The chosen non-dominated solutions correspond to the extrema of each objective and, if there is available space in memory, randomly chosen solutions are also taken to be later reinserted in the population.

Other relatively recent works have considered the application of DMOEAs as well as non-evolutionary MOO metaheuristics for tracking the PF in a variety of problems. For instance, Bingul [33] applied a reactive fuzzy-adaptive genetic algorithm for evolving military strategies in a complex dynamic four-objective combat simulation environment. Deb et al. [68] preemptly solved a bi-objective dynamic hydro-thermal power scheduling problem by inserting random and mutated solutions in the population evolved by NSGA-II.

Azevedo and Araújo [12, 13] also proposed the preemptive insertion of random solutions (albeit with guidance from the estimated joint cumulative distribution of objective vectors in [13]) as a means to continuously generate diversity in order to cope with changes in the objective functions. Zhang [225] proposed a preemptive artificial immune system framework with memory and diversity management for obtaining a three-variable discrete greenhouse control system. Helbig and Engelbrecht [112] considered different reactive archive management policies in a vector-evaluated particle swarm optimization algorithm for a number of sequential MOO benchmarks. Di Barba [74], on the other hand, developed a reactive ϵ-constraint-time-domain method for the dynamic bi-objective design of the shapes of electromagnetic devices.

Note, however, that all of the aforementioned DMOEAs for sequential MOO rely on either the *reactive* or *preemptive* (myopic) paradigms, which makes them strongly dependent on the several heuristics put together to trigger complex algorithmic behavior. Not only such excess of heuristics makes it almost impossible to analyze the DMOEAs behavior, but also they end up leading to lots of ad-hoc parameter tuning. This thesis, on the other hand, makes the case for principled, adaptive metaheuristics that do not require offline parameter tuning in order to handle changes in the objective functions. This is possible because such parameters are described as functions of the underlying perceived uncertainty. For the latest up-to-date literature overview on existing sequential MOEAs, c.f. [115].

On Preemptive Diversity-Based Metaheuristics

Premature convergence is a phenomenon that can be manifested in virtually all population-based search methods (a.k.a. metaheuristics) when optimizing over multimodal search spaces. A multimodal search space is one for which there are several local optima that may be understood as attractors in the search dynamics, i.e., the progress of the population across iterations can be directed to one or more local optima, making individual solutions progressively more similar to each other. This process is known as *diversity loss*[4].

Therefore, the inability of generating and maintaining diversity at the population level is strongly correlated with premature convergence. In dynamical environments, i.e., in problems whose objective-functions are subject to change over time, a low diversified population of candidate solutions will have less chance of adapting to changes [12, 13].

Different metaheuristics provide different parameters and strategies for diversity maintenance. For instance, in genetic algorithms, one can adjust the mutation rates or the selection pressure (e.g. adjusting the number of competing solutions in tournament selection), whereas in swarm intelligence, one can adjust e.g. repulsion factors – in particle swarm optimization – or evaporation rates – in ant colony optimization.

In related areas of machine learning, such as reinforcement learning, research on diversity maintenance is also refereed to as the "exploration vs. exploitation" trade-off, e.g. [55]. Exploitation roughly means intensifying the search towards promising regions, so better solutions can be found more quickly. The price to be paid when exploiting promising regions is that of diversity loss. In order to mitigate diversity loss, an exploration strategy can be employed, in which new candidate solutions are generated (or sampled) around previously unvisited regions of the search space. The key for an efficient search procedure is thus the efficient adaptive handling of such trade-off that arises in diversity management.

3.4.4 Overview of Ad-Hoc Predictive Approaches

The incorporation of prediction-based techniques into EAs can be traced back to the early 2000's, although one could argue that the self-adaptation concept conceived in the 1970's implicitly incorporated predictive information (e.g. [185]). Many ad-hoc prediction-based techniques have been proposed for coping with changing environments using different heuristics. For instance, van Hemert et al. [210] were the first to propose a two population-based approach, in which one population of candidate Pareto optimal decisions is evolved considering only the current available environmental information, whereas the second is evolved using an estimation of the future fitness values. By allowing solutions in the "future" population to migrate at a given rate to the current population, they were able to improve the genetic algorithm ability to track the dynamically changing optimum. That is because, although such mechanism is tailored to static environments, the metaheuristic parameters (mutation rates) are evolved based on a historical assessment of previous success rates on the task of generating improved solutions.

[4]Not every premature convergence event necessarily leads to a local optima. Sometimes, diversity loss (e.g. a constant value appearing in some locus of the decision vector throughout all candidate solutions) incurs in early stoppage of the search process in a region which does not necessarily represent a local optima. This is true specially for discrete search spaces.

Simões and Costa [190] used Markov chains and linear regression in a memory-based Evolutionary Algorithm (EA). The goal was to predict when the next change would happen and to estimate which kind of dynamic regime would follow, assuming periodic dynamics. In this scheme, the most promising representative solutions for the next environment are reinserted into the EA population. Although achieving positive results on numerical benchmark functions (which mostly complied with the assumptions), even assuming the predictors to be accurate, there are no guarantee that the level of similarity between the old and the upcoming environments is sufficient for the representative solutions to readapt, or even that the right building blocks will be propagated to generate novel, adapted solutions.

Koo et al. [135] used a global predictive gradient for estimating the trajectory of the PS. A convex combination of all objective functions with randomly sampled weights is used in the estimation. The solutions are then perturbed along a random direction leading to the predicted dynamical PS. The approach can thus improve the performance of dynamic EMOAs without the need for making unrealistic assumptions about the environmental dynamics. One drawback, however, is that the use of randomly drawn aggregation functions may fail when one or more objective functions are locally non-convex.

Hatzakis and Wallace [110] and Zhou et al. [226] used autoregressive models for tracking the dynamical PS. The main difference between the two approaches resides on the points used for representing the PS. The former uses two anchor points, whereas the later tracks the PS center of mass and the manifold defined by the difference of each non-dominated solution in the PS to its center of mass. Both approaches rely on inserting solutions from the predicted population into the actual population.

The main drawback of the discussed related prediction-based approaches is the inability to estimate the performance of the predicted solutions in the upcoming optimization environments. Also, it is not clear how many and which solutions should be inserted. All the existing predictive metaheuristics rely on problem-dependent ad-hoc strategies and rules that also introduce additional parameters to be tuned offline. Also, the lack of estimation of the solutions performance in future environments leaves the user clueless regarding the expected performance of the solutions obtained by those methods.

Although one could rely on ad-hoc predictive methods for handling uncertain environments, we propose a fourth yet unexplored principled approach: *anticipatory* MOO metaheuristics [14]. The main differences relies on the modeling of the optimization problem solved by such anticipatory methods and their uncertainty-awareness and adaptive capabilities in terms of online parameter tuning. Those properties are achieved through the conception of problem-independent techniques for estimating the future behavior of candidate decisions. Such ideas are developed in **chapter 6**.

The Effects of Prediction in Metaheuristics

In the context of metaheuristics applied for solving sequential decision-making problems, very few works have been dedicated to assess the incorporation of predictive models, which are indispensable to the synthesis of anticipatory metaheuristics. In the following, the two most broad studies about the effects of prediction within metaheuristics are discussed. Bosman [36], and

Bosman and La Poutré [35] reported a study on evolutionary prediction-based metaheuristics for dynamic optimization – problems for which the objective function is parameterized by time. In those works, preliminary results regarding the effects of incorporating predictive knowledge about the behavior of the objective function were reported, for simple metaheuristics.

In the first work [36], the metaheuristic performance was assessed in the following scenarios: (a) no access to predictive knowledge (the blind version); (b) partial access to predictive knowledge; and (c) complete access to perfect predictive knowledge. Bosman designed simplified objective functions over only one decision variable in the real line.

Not surprisingly, the metaheuristic with full access to the future objective values was always able to find the optimal solution throughout the optimization process. The versions with access to partial knowledge were designed by using linear and quadratic regression models for predicting the future objective values. Both predictors were trained using a least squares procedure. Bosman reported that the version utilizing the more flexible predictive model obtained smaller prediction errors and, as a result, the performance of the metaheuristic was only slightly inferior when compared to that with full knowledge of the future. This discrepancy was attributed to the timespan required for the accumulation of enough samples for training the predictor. As expected, the "blind" version obtained very poor performance regarding the distance between the solutions found to the true optimal solutions.

It should be noted that Bosman did not invest in the design of an efficient metaheuristic, but rather in understanding the implications of incorporating predictive knowledge on their performance in very simple toy problems. Besides, only deterministic, single-objective problems were considered. In the follow-up paper [35], the study was extended to stochastic scenarios, in which randomly distributed packages should be collected by a vehicle in the real plane. The results again showed there are significant performance gains when utilizing partial knowledge from the underlying distributions.

3.5 Summary of the Contributions

This chapter's contributions to the thesis are as follows:

1. It presented the foundations of Multi-Objective Optimization (MOO) under uncertainty and the limitation in existing Stochastic MOO (SMOO) models;

2. Early models of sequential MOO and Multiple Criteria Decision-Making (MCDM), including hypervolume-based approaches, were discussed. Early predictive metaheuristics were also covered. In both cases, limitations and drawbacks were addressed;

3. Finally, a logical chain of results connecting three key concepts for the anticipatory proposal in **chapter 5** and **chapter 6** was suggested: maximal hypervolume $\Rightarrow$ PD compatibility $\Rightarrow$ preference for flexibility.

In the next chapter, we not only discuss MOO models for the financial portfolio selection application, but also suggest and experiment with preemptive evolutionary MOO algorithms to improve portfolios stability in dynamic investment environments.

Chapter 4

Preemptive Multi-Objective Strategies for Active Asset Allocation

Apportion what you have into seven, or even eight parts, because you don't know what disaster might befall the land.

– Solomon (Ecclesiates 11:2)

A man who suffers before it is necessary, suffers more than is necessary.

– Lucius Annaeus Seneca

Following the Modern Portfolio Theory (MPT) paradigm inaugurated by Markowitz [151], investors seek to spread risk over uncorrelated (or negatively correlated) assets. The inception of MPT has formalized in mathematical terms a principle that is well-known to humanity perhaps since its very beginnings: the principle of *diversification*. This chapter describes the problem of how to allocate wealth among N risky assets, so as to maximize expected return and minimize the expected risk, known as *portfolio selection* [151]. Although this problem has been long addressed as a single-objective optimization one (by maximizing risk-return ratios), we discuss the advantages of solving its multi-objective counterpart. We also propose in this chapter preemptive strategies for mitigating investment risk based on hypervolume regularization[1] [15].

4.1 Risk Management: The Wisdom of Diversification

The motivation for diversifying investments is straightforward and intuitive: when some investment instruments go wrong, the investor is better off if significant portions of wealth had been also allocated to other well performing instruments. In fact, the loss of a Decision Maker (DM) who invested in the best-performing stocks of 2007 and held them throughout the 2008 financial collapse would be more than 60%, while investors who spread their money as to follow the S&P 500 index lost 37%[2]. Still a poor performance, but comparatively much better.

[1] Parts of this chapter have appeared in Azevedo and Von Zuben [15].

[2] http://usatoday30.usatoday.com/money/perfi/retirement/story/2011-12-08/investment-diversification/51749298/1

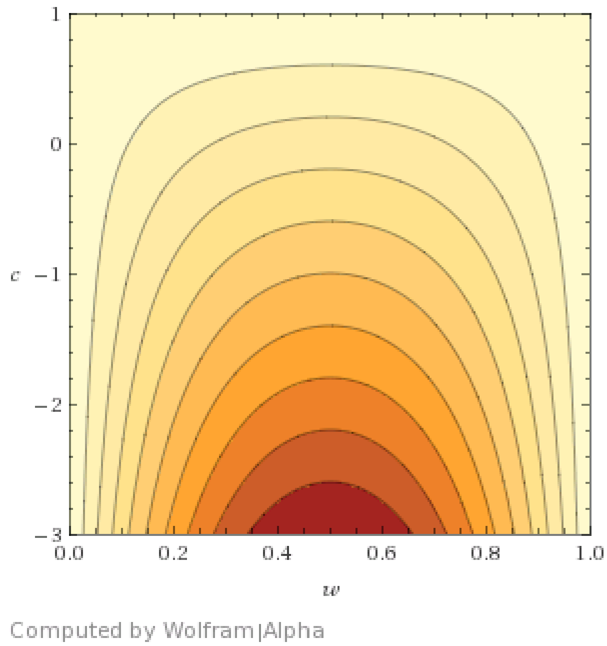


Figure 4.1: The contour plot of risk as a function of the weight w and of $c = 2\rho - 1$.

In order to illustrate how diversification works, let x and y be two normally distributed random variables representing the returns of two stocks, with equal variances, i.e., $\sigma_x^2 = \sigma_y^2 = \sigma^2$, and covariance $Cov(x, y) = \rho\sigma_x\sigma_y = \rho\sigma^2$, where ρ is the linear correlation between x and y. If we take the variance as a measure of risk[3], then we can compute the risk of a portfolio $\mathbf{w} = (w_1\ w_2)^\intercal \in S^1$ as:

$$\begin{aligned} Var[w_1x + w_2y] &= w_1^2\sigma_x^2 + w_2^2\sigma_y^2 + 2w_1w_2\rho\sigma_x\sigma_y \\ &= w_1^2\sigma^2 + w_2^2\sigma^2 + 2w_1w_2\rho\sigma^2 \\ &= (w_1^2 + w_2^2 + 2\rho w_1w_2)\sigma^2 \\ &= (w_1^2 + (1 - w_1)^2 + 2\rho w_1(1 - w_1))\sigma^2 \\ &= (w_1^2 + (1 - w_1)((1 - w_1) + 2\rho w_1))\sigma^2 \\ &= (w_1^2 + (1 - w_1)(1 + (2\rho - 1)w_1))\sigma^2. \end{aligned}$$

Note that, for fixed weights $\mathbf{w}$ and fixed σ^2, the risk is minimized as $\rho \to -1$. In other words, the more the assets returns are negatively correlated (thus yielding opposite behaviors), the more diverse the portfolio is regarded and less risk incurs. Also, taking the correlation ρ and the variance σ^2 as constants, the weights minimizing the risk in this case are $\mathbf{w}^\star = (\frac{1}{2}\ \frac{1}{2})^\intercal$. Supposing $\sigma^2 = 1$, the minimum risk value is then $\frac{c+3}{4}$, where $c = 2\rho - 1$ and $-3 \leq c \leq 1$. The risk of the portfolio in this case is thus completely eliminated when it is equally weighted and

[3]Markowitz modeled risk as the variance of the return distributions in MPT, but later admitted that it has serious flaws and proposed a measure of semi-variance to account only from variations below the average return. In fact, investors are actually more concerned in minimizing potential losses.

the assets are perfectly negatively correlated. Note how the benefits of diversification vanishes as $\rho \to 1$, in which case any allocation yields virtually the same risk (Figure 4.1).

4.2 Trade-offs of Active Investment Management

The MPT is rooted on the idea of diversifying capital across assets so to establish an investment portfolio that lies on the Pareto Frontier (PF), when considering expected return and risk. The impossibility of simultaneously improving both return and risk results from the fact that assets yielding higher returns usually are those leading to the gravest risks and vice-versa. It becomes an arduous task to compose a portfolio to maximize the expected utility of an investor – mainly because eliciting a utility function that accurately represents the investor attitude to risk is challenging and prone to inconsistencies. Moreover, as already discussed in previous chapters, the preferences of the DM can be undefined, what calls for multi-objective approaches to handle the underlying trade-off arising from MPT.

4.2.1 Single-Criteria Downside Investment Rules

Despite the practical appeal of multi-objective techniques, single-criteria decision-making heuristic rules in active investment management, which combine return and risk, are still common among practitioners. Those rules are often referred to as *downside* (or disaster level) methods [104] and are as old as Markowitz's MPT itself. One of the earliest methods came after Roy's 1952 paper [177] and has been known as the Safety First (SF) principle. The downside methods seek to search for the portfolio minimizing the chances of a return realization below some predetermined threshold[4]. Regardless of whether it is possible to determine an investor's actual utility function, Roy argued against expected utility maximization and claimed instead that investors are more concerned about preserving a certain amount of their invested capital [177]:

> *"A man who seeks advice about his actions will not be grateful for the suggestion that he maximizes expected utility."*

The SF principle is realized when the DM chooses a minimal acceptable return rate. The optimal SF portfolio selected is the one satisfying:

$$\mathbf{w}^{\star}_{\mathrm{SF}} = \arg\min_{\mathbf{w}} \Pr\{R_{\mathbf{w}} \leq R_0\}, \tag{4.1}$$

where $R_{\mathbf{w}} = \mathbf{w}^{\intercal}\mathbf{R}$ is the portfolio (random) return, and R_0 is the chosen disaster level (usually set to zero or some small negative rate). In practical scenarios, however, the true returns probability distribution is unknown. Nevertheless, using Tchebycheff's inequality,

$$\begin{aligned}\Pr\{R_{\mathbf{w}} \leq R_0\} &= \Pr\{\boldsymbol{\mu}_{\mathbf{w}} - R_{\mathbf{w}} \geq \boldsymbol{\mu}_{\mathbf{w}} - R_0\} \\ &\leq \frac{\sigma^2_{\mathbf{w}}}{(\boldsymbol{\mu}_{\mathbf{w}} - R_0)^2},\end{aligned}$$

[4] In **chapter 3**, we described general stochastic multi-objective optimization models (e.g. [56]) that were clearly inspired by those old risk management ideas.

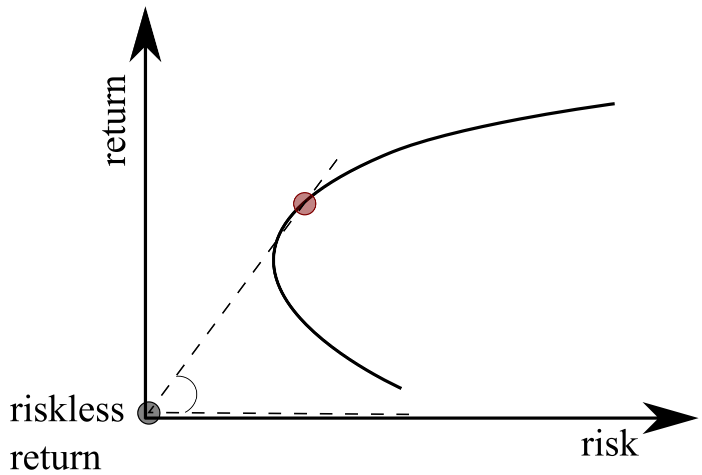


Figure 4.2: The Sharpe Ratio measures the slope between a risk-free asset return rate and the objective vector associated with a portfolio in the PF.

where $\boldsymbol{\mu}_{\mathbf{w}}$ is the mean vector of the returns, whereas $\boldsymbol{\sigma}^2_{\mathbf{w}}$ is the variance (Markowitz's risk). Thus, when the (empirical) distribution of the returns is unknown, the DM following the SF principle can approximate the problem by choosing the portfolio $\mathbf{w}^{\star}_{\mathrm{SF'}}$ such that

$$\mathbf{w}^{\star}_{\mathrm{SF'}} = \arg\min_{\mathbf{w}} \frac{\sigma_{\mathbf{w}}}{(\boldsymbol{\mu}_{\mathbf{w}} - R_0)}. \tag{4.2}$$

Because of the SF influence over theoretical and practical investing, Arthur D. Roy was recently described by Sullivan [201] as *"the forgotten father of portfolio theory"*.

Another remarkable property of the SF principle is that if R_0 is the fixed return rate (e.g. an inflation-indexed government bonds rate) of a risk-free asset, then $\mathbf{w}^{\star}_{\mathrm{SF}}$ is equivalent to the portfolio maximizing another popular downside measure, known as the Sharpe Ratio (SR) [189] (Figure 4.2). In practice, many investors search for portfolios that are shown to maximize average historical return per unity of risk. The SR is then defined as the expected return of $\mathbf{w}$ in excess of the available risk-free return rate divided by the volatility (standard deviation). Thus,

$$\mathbf{w}^{\star}_{\mathrm{SR}} = \arg\max_{\mathbf{w}} \frac{(R_{\mathbf{w}} - R_0)}{\sigma_{\mathbf{w}}}. \tag{4.3}$$

In the experiments reported in section 4.4.5, we use the SR as a decision-making rule for selecting a portfolio from the obtained efficient frontier.

4.3 The Case for Active Sequential Portfolio Selection

When solving sequential optimization problems for which there is no available predictive knowledge, sometimes heuristic rules are used for modifying the solution on the fly, as new data is collected. The demand for such *online* optimization techniques are emerging in modern applications like computational advertising and automated trading in the stock market (*algo trading*) [211]. The famous "*Universal Portfolio*" (UP) algorithm developed in the 1990's by Thomas Cover is one noteworthy example [59].

Without assuming anything about the statistical behavior of the market, Cover has shown that the portfolios selected *online* by his algorithm asymptotically outperform over time the performance of the best fixed portfolio in hindsight. This remarkable theoretical result has been tested – with somewhat positive results – using real-world data for finite investment horizons (e.g. [103]). Cover's UP is perhaps the most important result to date supporting active investment management strategies, because it explicitly states that the portfolios obtained with UP will outperform the best possible buy-and-hold, passive strategy, in the limit. The obvious practical drawback of the UP algorithm is that investors are not willing to wait indefinitely for collecting such promising returns. Another important drawback is that there are little hopes of accurately characterizing the inherent risks of the generated sequence of portfolios.

4.3.1 A Follow-The-Leader Preemptive Strategy

An example of a much simpler and general data-driven online heuristic rule is the following: let $\mathbf{w}_t$ be the weights of our portfolio at time t. We want to maximize the return, which is now time dependent, $R_{\mathbf{w}_t}$. A general procedure for solving the problem online is:

1. For $t = 1, 2, \cdots$, do
   (a) Choose a portfolio $\hat{\mathbf{w}}_t \in S^{N-1}$
   (b) Observe the returns $R_{t,i} (i = 1, \cdots, N)$
   (c) Compute the portfolio return $R_{\mathbf{w}_t} = \sum_{i=1}^{N} w_{i,t} R_i$

Note that $\hat{\mathbf{w}}_t$ must be blindly chosen before the assets returns $R_{t,i}$ are realized. Because there is no predictive knowledge available, the choice of $\hat{\mathbf{w}}_t$ must be based on the available historical data. A simple heuristic rule for Step 1 of this algorithm selects $\hat{\mathbf{w}}_t$ as the portfolio maximizing the (sample) return for all the past $t-1$ known investment periods, i.e.,

$$\forall t, \quad \hat{\mathbf{w}}_t = \underset{\mathbf{w} \in S^{N-1}}{\arg\max} \sum_{t=1}^{t-1} R_t(\mathbf{w}). \tag{4.4}$$

This rule is known as *Follow-The-Leader* (FTL), which is shown to lead to small *regret* for certain classes of convex objective functions (f_t) [187]. The regret is a simple measure taken over the temporal sequence of decisions that were produced by the FTL rule:

$$\text{Regret}_T(\{\hat{\mathbf{x}}_1, \cdots, \hat{\mathbf{x}}_T\}) = \sum_{t=1}^{T} f_t(\hat{\mathbf{x}}_t) - \min_{\mathbf{x} \in \Omega} \sum_{t=1}^{T} f_t(\mathbf{x}), \tag{4.5}$$

in which T is the number of decision periods. Note from Eq. (4.5) that the performance of the decisions is compared to that of the best fixed decision in hindsight [187].

It turns out Markowitz's analysis is also a backward-looking (myopic) strategy, in the sense that if the investment environment changes going forward, the analysis will suggest portfolios that are efficient only when evaluated using the outdated information. Although risk is reduced as the pairwise correlations between assets decrease, such correlations are subject to change over time and therefore should be predicted [80] – and so should the expected returns. Predictive strategies are nevertheless often overlooked in the literature of portfolio selection, perhaps due to the conservative pragmatic view that timing the market is futile. In fact, when using such myopic approach, the DM is impliciting investing (often wrongly) on a static investment environment.

The sequential (or multi-period) paradigm for active portfolio selection has been identified in the recent survey of Kolm et al. [134] as one of the most promising new research trends in this area. In this view, portfolios are seeing as moving targets that should be tracked considering all "*trade-offs between risk-adjusted returns and trading costs*" [134]. Before describing a multi-objective anticipatory proposal for tracking such trade-off portfolios over time (see **chapter 7**), we propose in the following a novel preemptive and stable strategy based on the notion of hypervolume regularization [15].

4.4 A Novel Multi-Objective Preemptive Strategy

In finance engineering, the formulation of the Markowitz's Mean-Variance Problem (MVP) [152] and its diversification principle have been the essence of theoretical investment risk management for nearly half a century. The solution to the MVP leads to an optimal wealth allocation strategy which maximizes the expected return rate under a maximum risk constraint[5].

However, despite its intuitive appeal, the practical application of the Markowitz portfolio approach is not reliable and has been reported to behave poorly in real-world out-of-sample data, e.g. [82, 98]. This is mainly due to the prevalent presence of outliers in real-world data which makes the tails of the estimated returns distribution fatter than those of the Gaussians assumed in the MVP [88]. Hence, approaches to robustly estimate the parameters of the returns distribution have been investigated for mitigating such model risk, e.g. [83, 98].

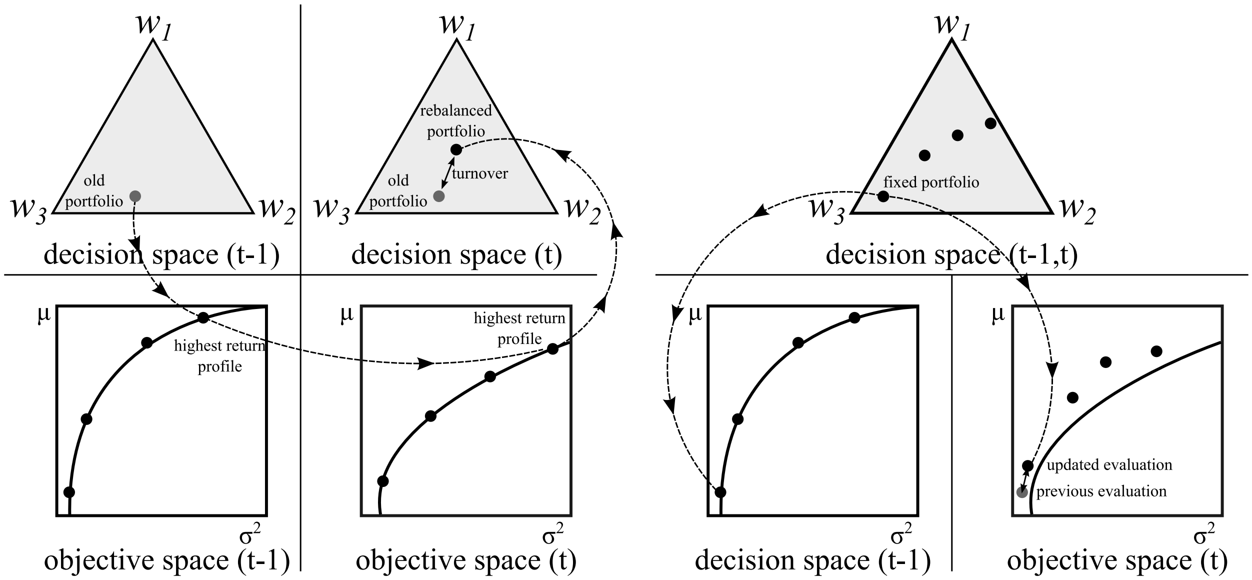
In realistic scenarios, it is reasonable to assume that a multiple account portfolio manager would desire to have a diverse set of mutually non-dominated portfolios at hand in order to serve a vast number of preference profiles. Some of the portfolios turn out to perfectly fit investors seeking higher returns at moderate risk, while others will suit stronger risk aversion profiles. Hence, obtaining a finite approximation of efficient allocations in the PF can be very useful to managers as new accounts arrive and preferences change.

Active asset allocation implies that the invested portfolio should be monitored and constantly rebalanced to reflect changes in expectations of risk and return. Very active strategies, however, generate transaction costs such as brokerage commissions and fees that may impair the investment return in the long term.

A well-known technique to mitigate transaction costs, especially for smaller investors, consists of introducing cardinality constraints in the optimization model so that the resulting portfolios are sparse, i.e., composed of few assets [29, 147]. For instance, Bertsimas and Shioda [29] argue that

> "(...) *asset management companies that manage separate accounts for their clients that only have say $100,000 can only realistically own a small number of securities, since otherwise, the transaction costs would significantly affect performance.*"

[5] Conversely, one can also minimize risk under a minimum expected return constraint.



(a) Stability in Decision Space (b) Stability in Objective Space

Figure 4.3: Two types of stability are investigated: (a) when rebalancing an old portfolio matching a fixed preference investment profile in the objective space (so that it can remain over the Pareto frontier), portfolio compositions yielding lower turnover rates are desired; (b) for a fixed portfolio, upon the update of the joint returns distribution, its performance evaluation (risk and return) should change as little as possible between one investment round to another.

That is because keeping a large number of assets in a portfolio requires more buy/sells operations in order to rebalance it according to updated data when compared with portfolios keeping a lower number of assets. However, the MVP problem becomes NP-hard when practical constraints are considered, such as bounded cardinality or transaction lots thresholds, for which classical convex quadratic programming techniques to analytically compute the efficient frontier no longer work [70]. Although there are exact algorithms for solving the mixed-integer quadratic programs [58] resulting from the constrained portfolio selection versions, the computation of optimal trade-off portfolios cannot be done in polynomial time. In those scenarios, multi-objective metaheuristics have been reported to yield good approximations of the efficient frontier for in-sample data (e.g. [5, 93, 178]), within reasonable (polynomial) computational cost.

When matching a portfolio at the efficient frontier to a certain preference profile, i.e., when selecting a trade-off portfolio according to the investor risk-averse level, account managers want the matched portfolio to remain *stable* between two subsequent investment rounds, upon the arrival of new data, in two senses:

1. The new rebalanced portfolio which matches the preference profile should not differ too much from the previously matched one, in the decision space (Fig. 4.3 (a)); and

2. The expected return and risk of a given portfolio should change as little as possible from one investment environment to another, in the objective space (Fig. 4.3 (b)).

Stable portfolios of the first kind usually possess lower turnover rates than unstable ones.

Higher turnover rates imply higher transaction costs. Hence, buy-and-hold passive strategies leads to minimum transaction costs, but are more susceptible to the market volatility. On the other hand, overly active rebalancing strategies can undermine the potential gains of the portfolio by incurring higher costs. This form of portfolio stability has been undergoing some research activity in the literature (e.g. [109]).

The second form of stability is also important because it implicitly means that the decision is adaptable and *robust over time* [90, 125]. In other words, for a fixed portfolio, as new temporal data arrive, the estimated expected return and risk will likely change much less than what it would for unadaptable portfolios. In this sense, being adaptable means being resilient to new and possibly unpredictable disruptions in the joint returns distribution process.

By proposing a new preemptive, stable multi-objective framework for portfolio optimization and assessing it on real-world data, this investigation leads to the following original contributions:

1. It proposes using a pairwise robust covariance matrix estimator [4] (see **chapter 2**) so that different robust statistics can be assessed at ease for MVP due to its quadratic complexity, when compared to exponential complexity of alternative techniques;

2. It proposes utilizing the robustness-integrating hypervolume [16] in order to obtain stable efficient frontiers, by integrating it into a novel regularized extension of the SMS-EMOA algorithm [32] (see **chapter 3**);

3. It assesses the stability of the obtained portfolios on out-of-sample real-world data, what, surprisingly, is an often neglected procedure in the literature; and

4. It benchmarks active management decision-making strategies under transaction costs, reporting the resulting return over investment and the corresponding turnover rates.

The main goal of those experiments are thus to assess whether the Regularized SMS-EMOA (RSMS-EMOA) proposal [15] outperforms the original SMS-EMOA on producing stable solutions in dynamic environments. For evaluating the merit of each algorithm, we selected two MVP scenarios: one for which unstable maximum likelihood estimators are used, and other for which robust estimators provide a more smooth change between investment rounds. We also compare the RSMS-EMOA with the NSGA-II algorithm, which has been reported to perform satisfactorily for multi-objective, single-period MVP frameworks (e.g. [5]).

It is expected that the RSMS-EMOA proposal is able to incorporate structural preferences over the candidate solutions into the optimization process. In the case of portfolio selection, a few of such desirable structural features may include: lower cardinality, balanced weight distribution, and stability in validation and test data sets.

4.4.1 The Sequential Markowitz's Mean-Variance Problem (MVP)

Let $\mathbf{r} = (r_1 \; r_2 \; \cdots \; r_N)^\intercal \in \mathcal{R}^N$ be a random vector composed of random returns of N risky assets, in which $\boldsymbol{\mu}_\mathbf{r}$ and $\boldsymbol{\Sigma}_\mathbf{r}$ are its mean and covariance matrix, and let $\mathbf{w} = (w_1 \; w_2 \; \cdots \; w_N)^\intercal \in$

$S^{N-1} = \left\{ \mathbf{w} \in \mathcal{R}^N : w_k \geq 0 \; \forall \; k \; , \; \sum_{j=1}^{N} w_j = 1 \right\}$ be the weight vector of the portfolio, denoting the proportion of wealth to be invested in each available asset, in which S^{N-1} denotes the $(N-1)$-simplex. Then, the classical Mean-Variance Problem (MVP) was formulated by Markowitz [151] as the following constrained quadratic program for risk minimization:

$$\mathbf{w}^{\star} = \arg\min_{\mathbf{w}} \left\{ \mathbf{w}^{\intercal} \boldsymbol{\Sigma} \mathbf{w} : \boldsymbol{\mu}^{\intercal} \mathbf{w} \geq \mu_0, \; \mathbf{w} \in S^{N-1} \right\}, \tag{4.6}$$

in which μ_0 is the minimum acceptable expected return for $\mathbf{w}^{\star}$. If the first and second moments are known in advance, the MVP can be directly solved. However, $\{\boldsymbol{\mu}, \boldsymbol{\Sigma}\}$ must be first estimated from the available historical data in real-world applications.

We propose transforming (4.6) into a mixed-binary bi-objective sequential formulation:

$$\min_{\mathbf{A}_t, \mathbf{w}_t} \quad \mathbf{w}_t^{\intercal} \mathbf{A}_t \hat{\boldsymbol{\Sigma}}_t \mathbf{A}_t \mathbf{w}_t \tag{4.7}$$

$$\max_{\mathbf{A}_t, \mathbf{w}_t} \quad \hat{\boldsymbol{\mu}}_t^{\intercal} \mathbf{A}_t \mathbf{w}_t \tag{4.8}$$

$$\text{s.t.} \quad \mathbf{1}_N^{\intercal} \mathbf{A}_t \mathbf{w}_t = 1, w_{t,j} \geq 0 \tag{4.9}$$

$$\operatorname{tr}(\mathbf{A}_t) \leq a_u \tag{4.10}$$

in which $\mathbf{A}_t$ is an $N \times N$ binary diagonal matrix so that $\mathbf{A}_t(j,j) = 1 \iff w_{tj} > 0$, i.e., if and only if the j-th asset belongs to the portfolio, and $\mathbf{A}_t(j,j) = 0 \iff w_{tj} = 0$. Note that the constraint $w_{t,j} \geq 0$ implies that no short-selling is allowed, otherwise, negative weights would be acceptable in the model.

The motivation for including binary variables into problem (4.7) to (4.10) is to allow for an increased search power[6] over the feasible decision space. In other words, by explicitly incorporating the information of whether an asset belongs to the portfolio, we can design heuristic search operators to exploit this information. By directly switching on and off the participation of specific assets in the allocations, we expect to provide the problem solver with the ability of escaping from local optima more quickly than if binary variables were not present in the model.

It is also worth noting that the use of search operators over the binary diagonal matrix $\mathbf{A}_t$ complies with the bounded cardinality constraint of Eq. (4.10) (in which $\operatorname{tr}(\mathbf{A}_t)$ denotes the trace of $\mathbf{A}_t$), allowing for the optimization procedure to quickly reduce (or increase) the cardinality of the portfolios (see Section 4.4.2 for details).

Moreover, incorporating cardinality constraints into the MVP model has lead to the experimental observation that the portfolios with the highest expected returns in the efficient frontier are generally those with lower cardinality and vice-versa, e.g. [53]. Cardinality is thus expected to be correlated with the position of the portfolios when mapped to the Pareto frontier in the objective space.

Furthermore, $\{\hat{\boldsymbol{\mu}}_t, \hat{\boldsymbol{\Sigma}}_t\}$ can be estimated by either using the standard MLE estimators, or by taking instead the sample median and the robust covariance estimator discussed in **chapter 2**. We recall that the motivation for robustly estimating the parameters of the return distributions is to mitigate model misspecification, as it is evident that real-world stock data rarely comply with the Gaussian assumption of Markowitz's MPT.

[6] We follow Deb and Agrawal's [67] definition of search power, which corresponds to the probability of moving to an arbitrary location of the decision space when applying one search operator on any given solution.

Finally, it is worth mentioning that, despite the Gaussian assumption is evidently wrong for financial data (see **chapter 2**, section 2.1.3), the point of investigating the effect of robust estimators and of regularization while still following the Gaussian assumption is to assess the extent to which those two factors can help our resulting portfolio selection algorithms to be more tolerant to outliers, and thus, lead to more stable portfolios.

The Sliding Window Plugin-Rule

When inputting available historical data into our algorithms, we want to somehow capture the changing behavior reflected in the assets prices between subsequent periods of time. This goal is achieved by using a sliding-window approach (see **chapter 2**), also known in the literature as *rolling horizon* [71]. The idea is to go through the historical data using a fixed-size window and advance it at a constant time step, so as to discard outdated information and progressively incorporate more recent data.

Let t be the current time index, that is initially set as $t \leftarrow t_0 (t_0 \geq K)$. The sliding-window approach then considers the following three-step procedure:

1. Given the time series of observed returns from the latest K time-steps for each asset, i.e., $\mathbf{R}_{t-K}^{t-1} = \{\mathbf{r}_k\}$, estimate $\{\hat{\boldsymbol{\mu}}_t, \hat{\boldsymbol{\Sigma}}_t\}$ from $\mathbf{R}_{t-K}^{t-1}$.

2. Solve Eqs. (4.7)-(4.10) after setting the unknown distribution parameters with those estimated from the data in the previous step. This procedure is called the plug-in-rule (see e.g. [83]).

3. Implement the efficient portfolio $\hat{\mathbf{w}}_t^\star$ at time t, respecting the preferences of the DM.

In a buy-and-hold strategy, the chosen efficient portfolio can be hold unchanged for the next T time-steps, in which T is the investment horizon, although in this chapter we select a potentially different portfolio at each decision period. During the investment period $\Delta_t(T) = [t, t+T]$, the Return Over Investment (ROI) associated with $\hat{\mathbf{w}}_t^\star$ at the end of $\Delta_t(T)$ is given by

$$\mathrm{ROI}(\hat{\mathbf{w}}_t^\star, \Delta_t(T)) = \frac{\mathcal{V}_{t+T}(\hat{\mathbf{w}}_t^\star) - \mathcal{V}_t(\hat{\mathbf{w}}_t^\star)}{\mathcal{V}_t(\hat{\mathbf{w}}_t^\star)}, \tag{4.11}$$

in which $\mathcal{V}_t(\mathbf{w})$ is the net asset value of $\mathbf{w}_t$ at time t. In the rolling horizon approach [71], the time index is then set as $t \leftarrow t + T$, the returns matrix $\mathbf{R}_{t-K}^{t-1}$ is updated to include the new observations arrived during $\Delta_t(T)$, discarding the oldest T data points, and the procedure goes on as the portfolios are rebalanced for the subsequent environments.

4.4.2 Single and Multi-Objective Metaheuristics for MVP

When constraints such as bounded cardinality (Eq. (4.10)) are integrated into the MVP, the problem becomes a mixed-binary quadratic programming, which is NP-hard [150]. Moreover, Mansini and Speranza [150] showed that the MVP with minimum transaction lots is NP-complete. Therefore, in practical settings, where such constraints are considered, mixed-binary MVP versions become computationally intractable for exact solution methods, hence, stochastic approximation methods such as metaheuristics are often considered to handle the computational complexity.

Single-objective metaheuristics such as simulated annealing, memetic, genetic, and estimation of distribution algorithms have been applied for the MVP [6, 60, 178]. Those works can be categorized as follows:

- Lagrangian methods, in which the problem becomes $\max_{\mathbf{w}} \boldsymbol{\mu}^\intercal \mathbf{w} - \lambda \mathbf{w}^\intercal \boldsymbol{\Sigma} \mathbf{w}$. In this case, the metaheuristic is executed several times over many values of λ for approximating the efficient frontier, e.g. [5].

- Maximization of the Sharpe Ratio (Sec. 4.2.1), which combines the returns maximization and risk minimization into a single objective function. In this case, one single solution is obtained, e.g. [6].

Fewer studies assessed multi-objective versions (e.g. [5, 70, 93, 109]) but none of them considered sequential, dynamic environments settings, or the integration of robust statistics and regularization terms. Some even failed to report performance in out-of-sample data.

Due to its good reported performance for the standard MVP [5], we use NSGA-II as a baseline for solving (4.7) to (4.10). In the following, we present our preemptive proposal for solving the cardinality-constrained sequential multi-objective MVP, which is an extension of the SMS-EMOA algorithm [32], introduced in **chapter 3**.

4.4.3 Regularized SMS-EMOA

The hypervolume ($\mathcal{S}$, $\mathcal{S}$-Metric) [10] is a theoretically sound performance indicator that was presented in **chapter 3**. An alternative mathematical formulation of this indicator is given as:

$$\mathcal{S}(\mathcal{F}_i, \mathbf{z}_{\mathrm{ref}}) = \int\limits_{\mathbf{z} \in \mathcal{F}(\Omega)} \mathbf{1}_{\mathcal{H}(\mathcal{F}_i, \mathbf{z}_{\mathrm{ref}})} \mathrm{d}\mathbf{z}, \tag{4.12}$$

$$\mathcal{H}(\mathcal{F}_i, \mathbf{z}_{\mathrm{ref}}) = \{\mathbf{z} \mid \exists\, \mathbf{z}' \in \mathcal{F}_i : \mathbf{z}' \prec \mathbf{z} \prec \mathbf{z}_{\mathrm{ref}}\}, \tag{4.13}$$

in which $\mathbf{1}_{\mathcal{H}(\mathcal{F}_i, \mathbf{z}_{\mathrm{ref}})}$ is the characteristic function of $\mathcal{H}(\mathcal{F}_i, \mathbf{z}_{\mathrm{ref}})$, that is 1 if $\mathbf{z} \in \mathcal{H}(\mathcal{F}_i, \mathbf{z}_{\mathrm{ref}})$ and 0 otherwise.

Robustness Integrating Hypervolume ($\mathcal{S}^\phi$)

Its robustness integrating version was defined by Bader and Zitzler as [16]:

$$\mathcal{S}^\phi(\mathcal{F}_i, \mathbf{z}_{\mathrm{ref}}) = \int\limits_{\mathbf{z} \in F_M} \alpha^\phi_{\mathcal{F}_i}(\mathbf{z}) \mathrm{d}\mathbf{z}, \tag{4.14}$$

in which the characteristic function in $\mathcal{S}$ (Eq. (4.12)) is replaced by the attainment function $\alpha^\phi_{\mathcal{F}_i}(\mathbf{z})$:

$$\alpha^\phi_{\mathcal{F}_i}(\mathbf{z}) = \begin{cases} \phi\left(\min\limits_{\substack{\mathbf{z}' \in \mathcal{F}_i, \\ \mathbf{z}' \prec \mathbf{z}}} r(\mathbf{z}')\right), & \text{if } \mathbf{z} \in \mathcal{H}(\mathcal{F}_i, \mathbf{z}_{\mathrm{ref}}) \\ 0, & \text{otherwise.} \end{cases} \tag{4.15}$$

```
Initialize $\mathcal{P}_0 = \{\mathbf{w}_p\}$ with $\mu$ portfolios and set $t \leftarrow 0$
repeat
    Generate a new portfolio $(\mathbf{A}, \mathbf{w})'$ from $\mathcal{P}_t$
    Set $\mathcal{P}_{t+1} \leftarrow$ Regularized Reduce($\mathcal{P}_t \cup \{(\mathbf{A}, \mathbf{w})'\}$) and $t \leftarrow t + 1$
until Termination condition is satisfied
```

Pseudocode 3: Regularized SMS-EMOA for Portfolio Optimization

```
Compute $\{\mathcal{F}_1, \cdots, \mathcal{F}_v\}$ by non-dominated sorting $\mathcal{Q}$ [32]
Select $\mathbf{z}_r = \arg\min_{\mathbf{z} \in \mathcal{F}_v} \Delta\mathcal{S}(\mathbf{z}, \mathcal{F}_v)\phi(r(\mathbf{z}))$
return $\mathcal{Q} \setminus \mathbf{z}_r$
```

Pseudocode 4: Regularized Reduce($\mathcal{Q}$)

The function $r : \mathcal{F}(\Omega) \to [0, 1]$ is designed such that the maximally robust solutions are assigned 0 and the least robust ones are assigned 1. It should be noted that although Bader and Zitzler [16] have named this modified hypervolume indicator as a "robustness integrating" one, it conveys a much more general framework for regularizing solutions according to any desirable structural characteristic (e.g. sparsity of the decision vectors).

The desirability function $\phi : [0, 1] \to [0, 1]$ can then assume various forms to account for structural preferences over the decision vectors. In Eq. (4.14), $\mathbf{z} \in \mathcal{H}(\mathcal{F}_i, \mathbf{z}_{\text{ref}})$ contributes 100% to $\mathcal{S}^\phi$ only if it is fully desired in terms of robustness. For $r(\mathbf{z}) > 0$, though, the contribution of $\mathbf{z}$ to $\mathcal{S}^\phi$ is discounted by the desirability assigned to the most "robust" solution $\mathbf{z}' \in \mathcal{F}_i$ that dominates $\mathbf{z}$.

We propose integrating r into Pseudocode 4 as a multiplicative regularization term by means of the desirability function, i.e. $\Delta\mathcal{S}(\mathbf{z}, \mathcal{F}_v)\phi(r(\mathbf{z}))$. This leads to the Regularized SMS-EMOA (RSMS-EMOA) proposal (Pseudocode 3), which is equivalent to maximizing $\mathcal{S}^\phi$ within classes. The choice of r allows for integrating selective pressure towards the following three desirable aspects of portfolio optimization: diversification, cardinality reduction, and stability. For the original SMS-EMOA, we define $r_0(\mathbf{z}) = 0, \ \forall \ \mathbf{z} \in \mathcal{F}^N$. For all versions of the algorithm, the desirability function is a linear decreasing mapping: $\phi(r(\mathbf{z})) = 1 - r(\mathbf{z})$.

Entropy Regularization

In order to simplify the exposition, we abuse notation so that the objective vector $\mathbf{z} = \mathbf{f}(\mathbf{A}, \mathbf{w}) \in \mathcal{R}^2$ is composed of the expected risk and return of the portfolio, computed as

$$\mathbf{z} = (\mathbf{w}^\intercal \mathbf{A} \hat{\mathbf{\Sigma}} \mathbf{A} \mathbf{w} \quad \hat{\boldsymbol{\mu}}^\intercal \mathbf{A} \mathbf{w})^\intercal. \tag{4.16}$$

Moreover, we allow r to take as parameter either the objective function or the decision parameter, in which case both notations are equivalent, i.e., $r(\mathbf{z}) \equiv r(\mathbf{A}, \mathbf{w})$.

With the goal of improving the diversification of the candidate portfolios, we define the regularization function r_E as the normalized (negative) entropy.

$$r_E(\mathbf{A}, \mathbf{w}) = 1 + \frac{1}{\text{tr}(\mathbf{A})} \sum_{i=0}^{N} a_{ii} w_i \log a_{ii} w_i, \tag{4.17}$$

in which portfolio $(\mathbf{A}, \mathbf{w})$ attains maximum diversification when its weights are equally distributed among the assets that take part in it, i.e., $w_i = \frac{1}{\mathrm{tr}(\mathbf{A})}\ \forall\ i$ such that $a_{ii} = 1 \implies r_E(\mathbf{A}, \mathbf{w}) = 0$. Also, $\mathrm{tr}(\mathbf{A}) = 1 \implies r_E(\mathbf{A}, \mathbf{w}) = 1$, i.e., imbalanced portfolios are less desired. It is worth noting that r_E is insensitive to the portfolio cardinality and, therefore, if an equally-weighted two-assets portfolio attains maximum diversification, so do the N-assets index portfolio.

Note that an equally weighted portfolio is not necessarily the one that minimizes risk, but the practical goal of introducing this regularization term into the RSMS-EMOA algorithm is to penalize extremely imbalanced portfolios that induce higher risks.

Cardinality Regularization

Besides diversification, another desirable feature are low cardinality portfolios, inducing sparse weight vectors. Such portfolios may be easier to manage in practice, but may also conflict with risk minimization. The regularization function r_C is given as

$$r_C(\mathbf{A}, \mathbf{w}) = \frac{1}{N}\mathrm{tr}(\mathbf{A}), \tag{4.18}$$

in which N-assets portfolios are the least desired ones.

Objective Space Stability Regularization

Finally, we focus on the second form of portfolio stability discussed in Sec. 4.4. This is important because it means that the solution is adaptable over time, even if the market enters a more volatile phase (see e.g. [109]). As new data arrive, the estimated expected return and risk of such stable portfolios will likely change less than what it would do for unadaptable portfolios.

In order to leverage stability in the objective space, we estimate the change in the risk and return of the candidate portfolio in a validation dataset which is temporally adjacent to the training set – and therefore expected to reflect changes in the returns distribution. Let $\mathbf{R}_{t-K}^{t-1}$ be the returns data matrix in period $\Delta_{t-K}(K)$. We then partition that matrix into $\mathbf{T} = \mathbf{R}_{t-K}^{t-H-1}$ and $\mathbf{V} = \mathbf{R}_{t-H}^{t-1}$ $(T \ll K)$, in which $\mathbf{T}, \mathbf{V}$ are the training and validation returns datasets for period $\Delta_{t-K}(K)$. Let $\mathbf{z}_T$, $\mathbf{z}_V$ be the objective vectors of the portfolio evaluated in $\mathbf{T}$ and in $\mathbf{V}$, respectively. Then,

$$r_S(\mathbf{z}) = 1 - \frac{1}{1 + ||\mathbf{z}_T - \mathbf{z}_V||^2}, \tag{4.19}$$

in which the objective vectors that remain stable under the same time varying conditions are most desired.

4.4.4 Experimental Setup

We assessed two RSMS-EMOA(r) versions[7] in our experiments. The RSMS-EMOA(r_S) utilizes the regularization function of Eq. (4.19) and tries to regularize the portfolios so that their estimatives are stable in subsequent environments. For the second version, we mixed r_E and

[7]Source codes at http://www.researchgate.net/profile/Carlos_Azevedo2

r_C to form $r_{EC} = (1 - \alpha)r_E + \alpha r_C$, yielding the RSMS-EMOA(r_{EC}), whose goal is to achieve lower cardinality portfolios with a diversified wealth allocation, as a plausible ad-hoc preemptive strategy for mitigating risk and for making the portfolios easier to manage.

We set $\alpha = 1/2$ for the experiments so to provide a better balance between the tensions of reducing portfolio cardinality and of simultaneously improving their diversification.

Because NSGA-II, SMS-EMOA, and RSMS-EMOA(r) all share the same evolutionary algorithmic framework, we were able to render all parameters values the same for each version. The setup utilized is as follows: population size was set as $\mu = 100$; mutation rate of 0.2; crossover probability of 1.0; and binary tournament selection. Binary selection in a multi-objective context means that (a) two portfolios are randomly selected from the population of candidate solutions; (b) if either one Pareto dominates the other, it is chosen for undergoing crossover and mutation; (c) if they are mutually non-dominated, then the one yielding the highest $\mathcal{S}$-Metric contribution (or Crowding Distance for the NSGA-II) is declared the winner.

Because our focus is on analyzing the effects of robust parameter estimation and regularization, we utilized a very simple uniform crossover in which two offspring inherit each weight of the portfolios from either parents with 1/2 probability. For mutation, we randomly choose between two operators: (1) modify the non-zero weights; or (2) modify the binary diagonal matrix ($\mathbf{A}$). If operator (1) is selected, then, with probability 1/2, we either increase or decrease the investment on a randomly chosen asset by a uniformly drawn factor from 10 to 50%. If (2) is selected, then, with probability 1/2, we either add or remove a randomly chosen asset from the portfolio. If it is removed, we simply set its weight (and its entry in $\mathbf{A}$) to zero. If it is added, we set its entry to one in $\mathbf{A}$ and randomly set its weight within a $\pm 10\%$ range from an equally-balanced weight, $1/\mathrm{tr}(\mathbf{A})$. After any operation, the portfolio weights are renormalized.

Other parameters for the simulations were: 100 generations for each algorithm in each environment; 30 runs; 89 FTSE 100 assets, using the adjusted close prices from 20/11/2006 to 20/11/2012 (including the financial crisis period); three investment horizons ($H \in \{25, 50, 100\}$); the training periods length K were chosen such that $H/K \approx 0.13$, meaning that the rate of change between environments is constant for all horizons. This rate of change was chosen to yield model a slowed-paced time-varying investment environment.

We also used Mann-Whitney paired significance tests at a 5% level where appropriate. In addition, we narrowed the scope of the experimental study, relaxing the cardinality constraint and allowing all 89 assets to take part in the portfolios, but added a minimum investment constraint of 0.5%, which is easily handled by pruning the asset which violates that condition and renormalizing the portfolio.

Finally, we adopted a seeding approach in which, whenever the environment changes, the previous efficient portfolios serve as starting points for the optimization in the subsequent optimization period. In this way, every portfolio in the population is rebalanced between investment environments. In the following, we describe additional metrics used for discussing and comparing the performance of each algorithm.

Ranks Correlation Between Periods

We compute the Spearman Correlation (SC) coefficient between the portfolios in the efficient frontier, between the expected return estimated before and after each rebalancing operation. Differently from Pearson correlation, the SC coefficient computation is based on the relative rank orders between the observed samples. Thus, it can capture some non-linear correlations as well, provided that the relationship between the two measured variables is monotonic. Because it is based on order statistics, it is also considered a more robust estimate when compared to Pearson correlation.

Stable frontiers are expected to possess high rank correlation, keeping the relative order of the portfolios so that e.g. low risk portfolios will likely remain so in a subsequent investment period, when evaluated under new data.

Turnover Rates

For measuring the Turnover Rate (TR) of a given investment strategy s over time, we follow DeMiguel and Nogales [71] and define TR as the average absolute difference in the portfolios weight allocations, traded between subsequent investment periods:

$$\mathrm{TR}(\mathbf{w}^s) = \frac{1}{T-K-1} \sum_{t=K}^{T-1} \sum_{j=1}^{N} \left| w^s_{(t+1),j} - w^s_{t,j} \right| , \tag{4.20}$$

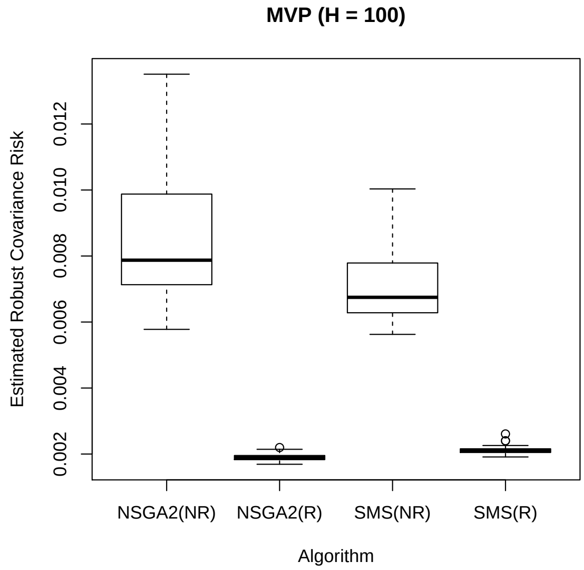
in which s is an index for identifying the portfolio. High turnover has direct consequences for trading, since it generates high transaction fees and costs. Are stable portfolios in the objective space also stable in the decision space? By measuring turnover from our experiments, we expect to be able to answer this question.

Decision-making Rule

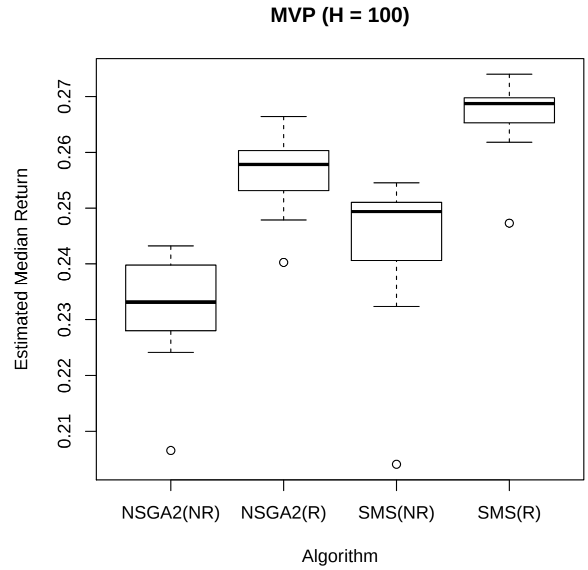
In real-world portfolio management, it is common for investors to search for the portfolio which is the most rewarding for each risk unit taken. Therefore, at the beginning of each investment period, we identify the efficient portfolio which possesses the highest Sharpe Ratio (SR) value (Eq. (4.3)), $\mathbf{w}^\star_{SR}$, and implement it in a rolling horizon, out-of-sample data simulation, rebalancing the portfolio for 6, 22, and 56 periods of $H = 100$, 50, and 25 days, respectively. Those rebalancing periodicities where chosen so to approximately respect the defined rate of change of $H/K \approx 0.13$. For the SR computation, we set $R_0 = 0$. Besides the maximum SR value achieved by $\mathbf{w}^\star_{SR}$, we also report the turnover rates, entropy, and the cardinality of the portfolio. The results are contrasted to those of the naive equally-weighted portfolio, $\mathbf{w}_{\mathrm{Index}}$, implementing a buy-and-hold strategy for the whole test period.

4.4.5 Results and Discussion

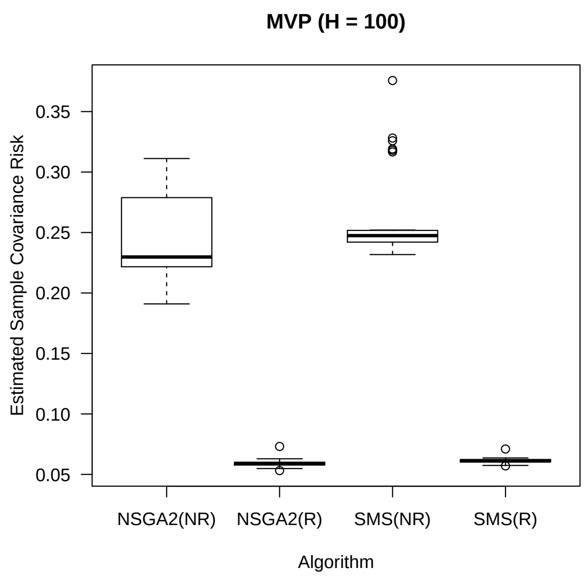
We separate the results reporting in two parts: first, we point out the observed benefits of solving the robust MVP over the original non-robust version in terms of portfolio diversification and stability; then, we discuss the effects of integrating regularization terms under the RSMS-EMOA(r) framework.



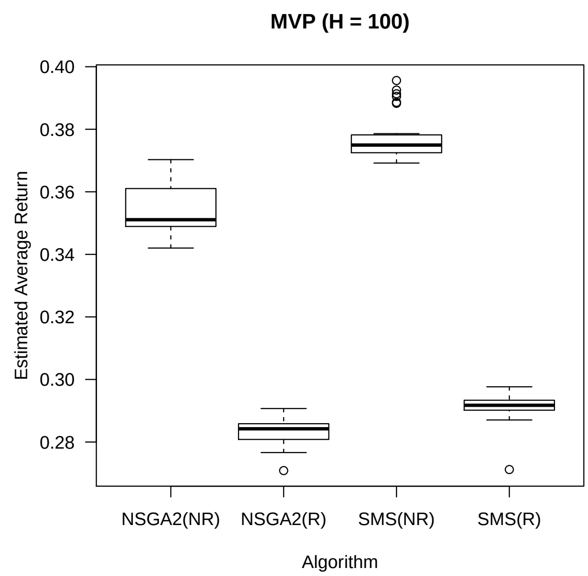
(a) Robust risk.



(b) Robust return.



(c) Maximum-likelihood risk.



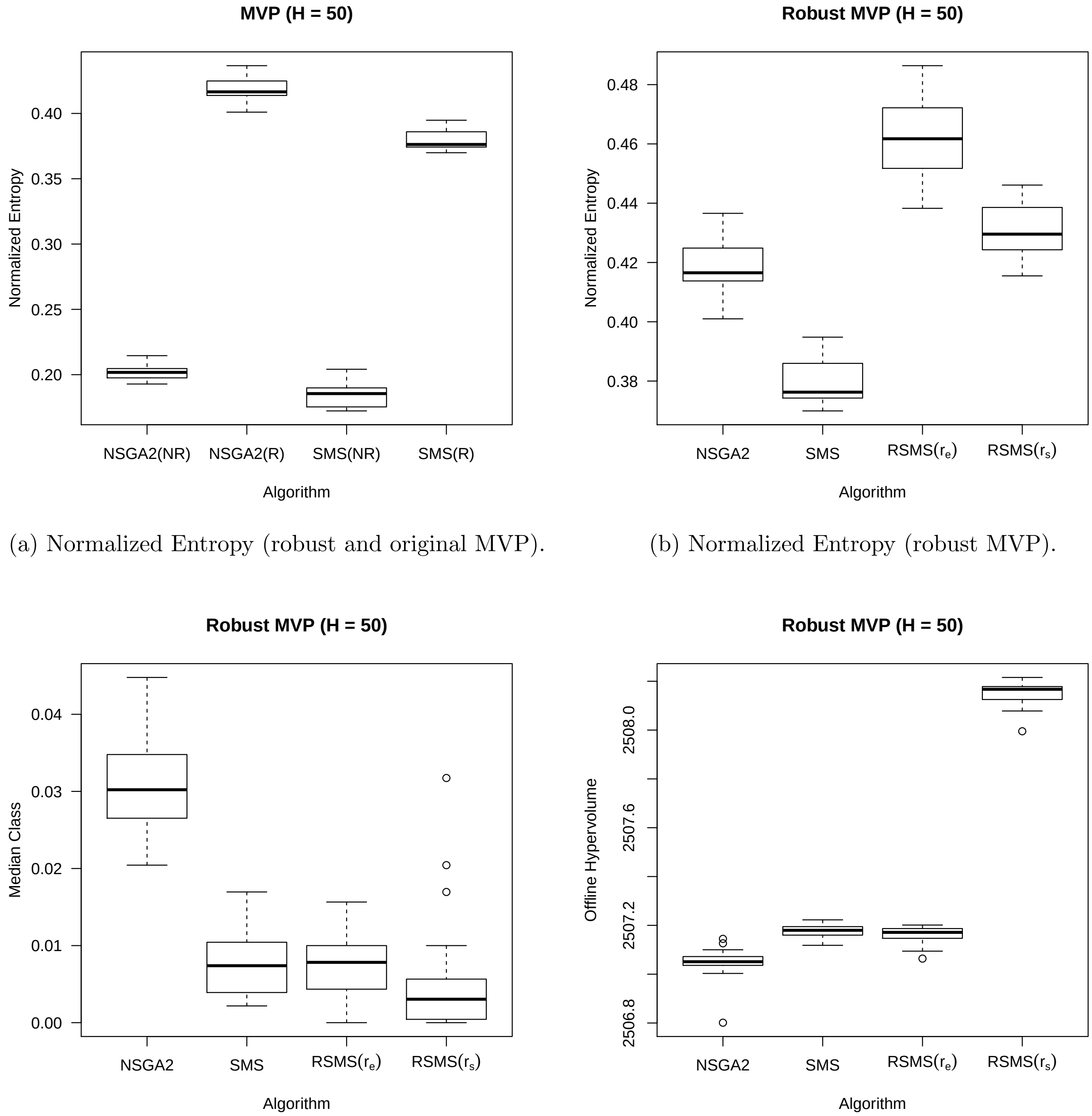
(d) Maximum-likelihood return

Figure 4.4: Boxplots reporting the average estimated risk and return for all obtained efficient portfolios, for algorithms solving the robust and the original MVP. (a)–(b) Robustly estimated risk/return; (c)–(d) Maximum-likelihood estimated risk/return.

The Effects of Robust Estimation

The first interesting observation we can draw regards what happens to the robustly estimated return and risk[8] when considering the portfolios obtained by the algorithms solving the non-

[8]All boxplots represent offline measures, meaning that we collect the statistics at the end of each training period, before rebalancing the portfolios.



(a) Normalized Entropy (robust and original MVP).

(b) Normalized Entropy (robust MVP).

(c) Median Class.

(d) Offline hypervolume.

Figure 4.5: Boxplots reporting relevant performance measures for the algorithms solving the robust and the original MVP.

robust MVP, and vice-versa. For instance, we observe that, despite solving MVP by utilizing the robust covariance matrix for estimating risk, both NSGA-II(R) and SMS-EMOA(R) produced portfolios with much lower risk when evaluated using the non-robust sample covariance estimator (Figure 4.4(c)). The same did not happen with regard to return (Figure 4.4(d)), as those algorithms evolved efficient portfolios with lower estimated (sample average) return, although the estimated median return was statistically significantly higher than for the non-robust port-

Table 4.1: Ex-post performance for the maximum Sharpe Ratio portfolio from the efficient frontier.

Algorithm	$\mathrm{Max}_{\mathrm{SR}}$	Turnover	Cardinality	Entropy
NSGA-$\mathrm{II}_{\mathrm{NR}}$	1.219	0.866	**7.370**	0.725
NSGA-II_{R}	3.628	0.749	26.610	0.854
SMS-$\mathrm{EMOA}_{\mathrm{NR}}$	1.193	0.876	**6.00**	0.743
SMS-$\mathrm{EMOA}_{\mathrm{R}}$	3.652	**0.717**	25.550	0.859
RSMS-EMOA(r_{EC})	3.692	**0.720**	26.036	**0.861**
RSMS-EMOA(r_S)	**4.106**	0.791	25.440	0.847
Index ($\mathbf{w}_{\mathrm{Index}}$)	–	0.000	89.000	1.000

folios. This result suggests that (*i*) solving the original MVP may lead to more risk-inclined portfolios; and (*ii*) the robust MVP may lead to less risk-inclined portfolios.

Another remarkable effect of integrating the robust statistics into the original MVP was observed for the portfolios diversification. While the algorithms solving the non-robust MVP, NSGA-II(NR) and SMS-EMOA(NR), obtained low cardinality (see Table 4.1) and low entropy portfolios, the opposite happened with NSGA-II(R) and SMS-EMOA(R), which favored medium cardinality portfolios with more balanced wealth allocations (Fig. 4.5(a)). This observation strongly complies with the observed lower estimated risk for the robust versions previously noted.

Regarding the effects of solving the robust and original MVP formulations on stability in the objective space, we note from Fig 4.6(d)–(f) that the robust versions did not always lead to the most stable Pareto frontiers, as measured by the rank correlation, although for the FTSE 100 data with the shortest investment horizon, $H = 25$ (Fig 4.6(d)), which requires the most frequent rebalancing operations (56 in total), the NSGA-II(R) obtained the most stable frontiers (with statistical significance), followed by SMS-EMOA(R), what is also a remarkable outcome.

Finally, from the ex-post simulations, in which the maximum SR ($\mathrm{Max}_{\mathrm{SR}}$) portfolio is implemented in each period, we observe from Table 4.1 that the non-robust versions were outperformed by all other robust versions in terms of stability in the decision space, as measured by the turnover rates. Here, we find intuitive the fact that, due to the lower cardinality portfolios evolved by NSGA-II(NR) and SMS-EMOA(NR), more effort is needed to rebalance the portfolio in order to handle changes in the non-robust return/risk estimates as new data arrive, which was somewhat expected.

The Effects of Regularization

Among the algorithms solving the robust MVP formulation, the RSMS-EMOA(r_{EC}) was successful on obtaining the highest diversified portfolios (with statistical significance), as expected (Fig. 4.5(b)), followed by the other regularized version, RSMS-EMOA(r_S). The plain SMS-EMOA(R), on the other hand, appeared to have greedily traded off diversification by return maximization, what may explain the lowest entropy portfolios for the robust MVP while attaining higher hypervolume[9] values than those achieved by NSGA-II(R) (Fig. 4.5(d)). In terms of offline hypervolume, RSMS-EMOA(r_S) outperformed its contenders, followed by the other SMS-EMOA versions, which outperformed NSGA-II. The striking performance of RSMS-EMOA(r_S) in terms of hypervolume is coherent with the fact that it also obtained the portfolios with the lowest median class values among all versions. After the non-dominated sorting procedure is applied for all population, each portfolio is assigned a class value, ranging from zero to the number of partitions found. It can be seen that RSMS-EMOA(r_S) is the version that keeps the highest number of mutually non-dominated portfolios in the population, what may contribute for a more effective exploitation of the efficient frontier, thus greatly improving its population hypervolume.

The version that seeks to incorporate stability in the objective space, RSMS-EMOA(r_S), has achieved its goal (see Figure 4.6(a)–(c)), although tied with NSGA-II(R) for $H = 25$, and, therefore, for investors seeking portfolios which are less likely to be affected by market volatility over time in terms of the estimated return and risk, it is the recommended choice among the algorithms we tested. What is interesting about this result, however, is that such objective space stability leads to higher instability in the decision space (as measured by turnover rates, Table 4.1), when compared to the other algorithms solving the robust MVP.

This apparent conflict between the two forms of stability makes sense when recalling that RSMS-EMOA(r_S) also achieved the highest hypervolume values: when attempting to stabilize the return and risk estimatives between rebalancing periods, the algorithm may have become more effective at improving over local Pareto sets in the decision space, given that it showed the highest levels of improvement in hypervolume between training periods. However, this goal was attained at the cost of more dramatically modifying the weights of the seeded population of efficient portfolios from one investment environment to another.

4.4.6 Concluding Remarks

The regularized, robust algorithms proposed in this chapter stands under the class of preemptive approaches for model (the robust part) and structural (the regularization part) risk mitigation. The results obtained with the proposed rolling horizon multi-objective portfolio framework for the robust Mean-Variance Problem (MVP) support the following remarks:

- The plain SMS-EMOA have outperformed NSGA-II on approximating the MVP efficient frontier for real-world stock data, both under robust and non-robust estimators for the mean and covariance matrix of returns, but the Pareto frontiers obtained with NSGA-II were the most stable ones;

- The consideration of robust statistics in the MVP formulation, specifically the pairwise robust covariance matrix estimation, allowed for more stable portfolios in the decision space, as significantly lower turnover rates have been observed;

[9] Because we did not know what to expect from the range of values that would be taken by both return and risk, we set the reference point to arbitrary high values (in magnitude) as (50,-50) for risk/return.

- The proposed RSMS-EMOA framework was effective on incorporating desired structural aspects in the final portfolios, such as reduced cardinality and improved diversification and stability; and

- The two concepts of stability discussed for both spaces (objective and decision) are desirable, but, as indicated in our results, they may be conflicting for different investment profiles.

Finally, despite the focus on the MVP application, we claim that the RSMS-EMOA can be used in virtually any problem for which regularization plays an essential role, such as on avoiding overfitting in machine learning multi-objective model building (e.g. in neural networks architecture optimization).

Further Developments and Investigations

Motivated by such positive results, we suggest assessing an RSMS-EMOA version explicitly incorporating the turnover rate (measured in reduced validation sets) as the regularization term. The higher computational effort of such task, however, would not pose a problem, since this step can be done offline within reasonable time, even for databases containing thousands of assets. Investigation on the evaluation of such preemptive algorithms in controlled scenarios with generative models should also be pursued.

4.5 Summary of the Contributions

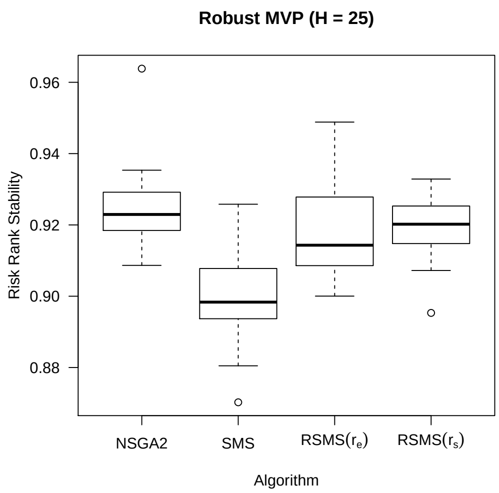
This chapter's contributions to the thesis are as follows:

1. It presented and motivated the principle of diversification, as well as several notions of risk mitigation when deciding how to perform investment allocation in financial instruments;

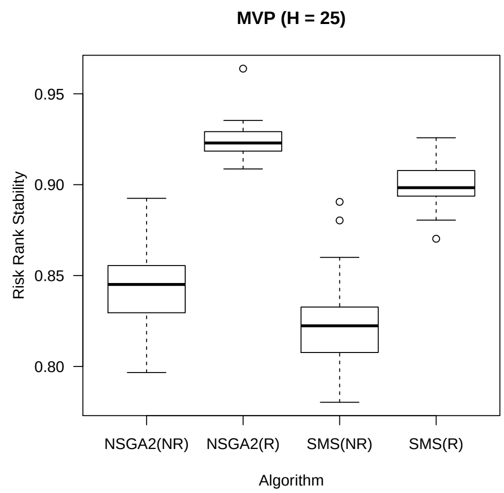
2. The plausibility of active investment management strategies was analyzed, and the potential benefits of using preemptive strategies were assessed;

3. Finally, an original multi-objective proposal for sequential portfolio selection under cardinality constraints was experimentally investigated using real-world stock data, with encouraging results.

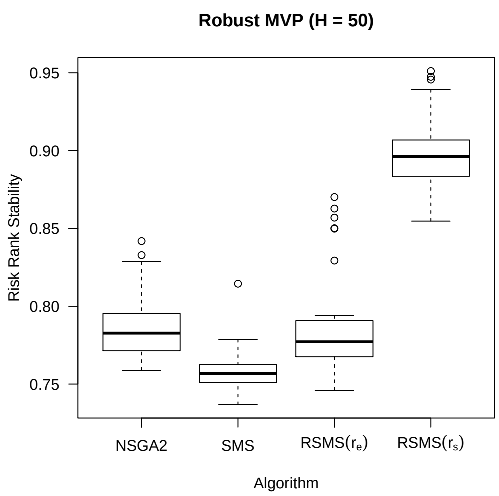
The next chapter describes anticipatory multi-objective optimization and decision-making models, providing one of the main contributions of this thesis.



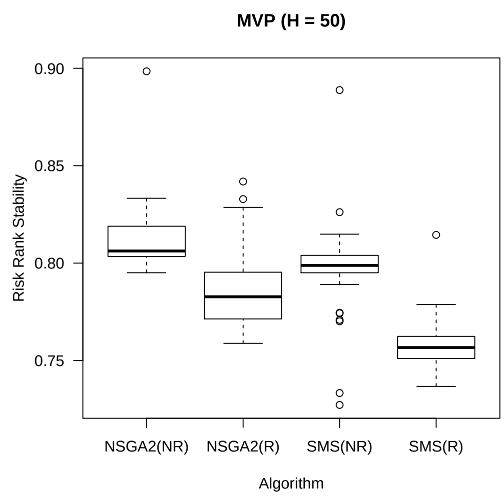
(a) $H = 25$.



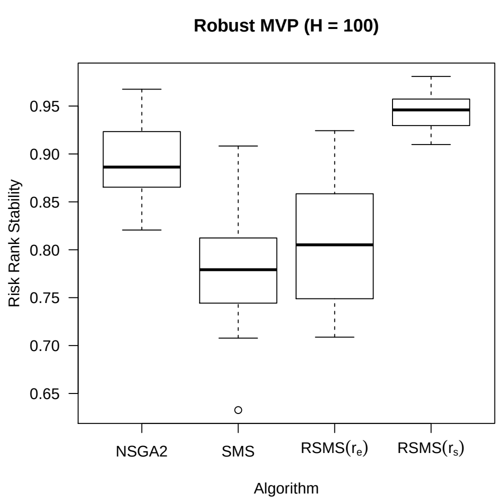
(b) $H = 25$.



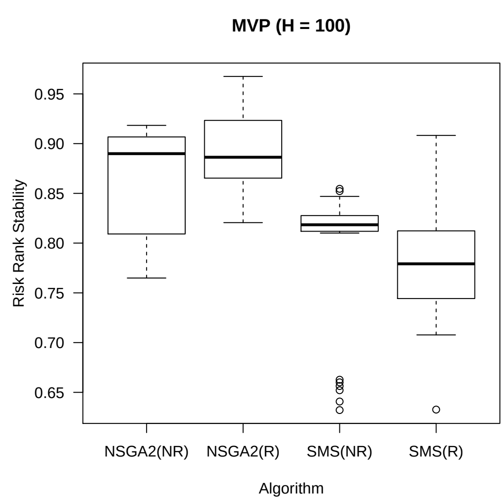
(c) $H = 50$.



(d) $H = 50$.



(e) $H = 100$.



(f) $H = 100$.

Figure 4.6: Boxplots reporting the observed Spearman rank correlation. (a)–(c) Algorithms solving the robust and the original MVP; (d)–(f) Algorithms solving only the robust MVP, including the regularized versions.

Chapter 5

Hypervolume-Based Anticipatory Multiple Criteria Decision-Making

We choose the thoughts we allow ourselves to think, the passions we allow ourselves to feel, and the actions we allow ourselves to perform.

– **Benjamin Franklin**

An intense anticipation itself transforms possibility into reality; our desires being often but precursors of the things which we are capable of performing.

– **Samuel Smiles**

The role of flexibility and anticipation for decision-making is discussed in this chapter and two novel anticipatory models for online Stochastic Multi-Objective Optimization (SMOO) are proposed. The implications of such flexible Anticipatory SMOO (AS-MOO) models for online multi-criteria decision-making are addressed. In addition, a proposal for effectively incorporating prediction into multi-objective problem-solving tools is presented, leading to a novel class of anticipatory multi-objective metaheuristics. The proposed metaheuristics are designed to approximately solve the AS-MOO models discussed hereinafter. Finally, an online anticipatory learning methodology is devised for self-adjusting time preferences according to observed historical prediction errors and to predictive future uncertainty estimated from Bayesian tracking tools that are integrated into MOO metaheuristics.

The models and algorithms developed in this chapter thus provide the basis for the automation of Multiple Criteria Decision Making (MCDM) processes when it is unclear how the trade-offs between the decision criteria *change over time* and when choice consequences are subject to *stochastic uncertainty*.

The autonomous behavior that emerges from the proposed flexible anticipatory metaheuristics is achieved by modeling and automating a Decision Maker (DM, or agent) whose preferences become achieving future improved *flexible and predictable* sets of alternatives to choose from in later periods, for what the search and implementation of solutions predicted to maximize the expected hypervolume over sets of future non-dominated alternatives is pursued.

5.1 Challenges of Solving Sequential Stochastic Multi-Objective Optimization

The traditional premise when designing optimization systems for decision-making under uncertainty, as discussed in previous chapters, is that all relevant parameters for characterizing the problem's random variables are readily available, as well as the preferences of the Decision Maker (DM) and his/her attitude to optimizing near and long-term performance. This is however an unrealistic statement in practical scenarios, wherein: (1) the system must instead resort to noisy historical observations for model building and parameter estimation; (2) the DM has often little to zero knowledge about the relative importance of the optimization criteria; and (3) the DM has no subsidies for deciding whether near-term *performance levels* are more valuable than long-term ones relative to the current decision period.

As shown in **chapter 4**, the former challenge can be addressed by data-driven optimization approaches to enable the DM to overcome the dependency on unrealistic modeling assumptions. This was the case when the parameters of the Gaussian distribution chosen to represent the returns empirical distribution were estimated using robust statistics. Such estimatives were then plugged-in directly into the optimization model. The challenges (2) and (3), however, are extremely overlooked in the literature and little progress has been made to address them in the context of sequential MCDM. We argue in this thesis that addressing challenges (2)–(3) should require as little involvement as possible from the DM. That is because eliciting the DM complete preferences and eagerness to near-term optimized performance under little knowledge about the problem can be unattainable at best and too much speculative and deceptive at worst.

The focal innovations of this thesis for automating MCDM under uncertainty are thus devoted to tackle those two challenges by integrating anticipation into Multi-Objective Optimization (MOO) solvers. The intended goal is to design online anticipatory MOO solvers to search for the *Anticipated Maximal Flexible Choices* (AMFCs). The implementation of flexible choices is justified as to allow the DM to not compromise into the unknown. In other words, AMFCs pave the way for the DM to progressively assume control of the decision-making process, as full knowledge becomes available in the future. On the modeling side, we presume that more principled yet tractable online Stochastic MOO (SMOO) models are attainable. On the algorithmic side, the premise is that the optimization tools should require the least involvement from the DM in terms of *parameter tuning*.

We begin the exposition by addressing the modeling aspects of those novel anticipatory MCDM strategies. The third challenge is handled in this thesis through a strategy to automatically determine the DM eagerness to near-term optimized performance that is incorporated into the proposed anticipatory MOO algorithms via the concept of *Online Anticipatory Learning* (OAL, see **chapter 6**, section 6.1).

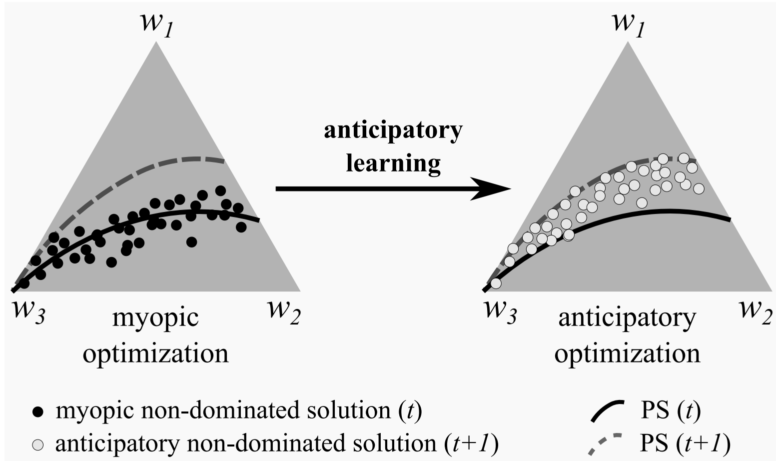


Figure 5.1: Effects of prediction in the search space. On the left, the evolved candidate solutions in the 2-simplex using myopic optimization overfit the current Pareto Set, whereas, on the right, anticipatory learning allows for better adaptability and the evolved solutions are in-between the current and the future Pareto Set. The assumption is that fixed optimal trade-off solutions can become dominated between two decision periods. Note however, there might be portions of the dynamical Pareto Set wherein one or more solutions can still be non-dominated between consecutive periods.

5.2 The Anticipatory Stochastic Multi-Objective Optimization (AS-MOO) Model

The expected effects of adopting an anticipatory strategy in a time-varying optimization problem are depicted in Fig. 5.1, as opposed to solving a myopic SMOO formulation. Suppose w_1, w_2, w_3 are three decision variables in the 2-simplex S^2 (i.e., $w_1 + w_2 + w_3 = 1$ and $w_1, w_2, w_3 \geq 0$). Suppose also that we are solving a two-period decision multi-objective optimization problem. Optimizing over the predicted value of hypervolume of the current candidate non-inferior solutions set for $t + 1$ allows for obtaining trade-off solutions that are in between the Pareto Sets (PS) for the decisions periods t and $t+1$, depending on the *anticipation rate* (or discount factor) chosen by the DM. When compared to a myopic strategy making no use of internal predictive models, an anticipatory strategy is therefore expected to provide mutually non-dominated solutions representing a better trade-off between near and long-term performance levels. Therefore, the obtained solutions can be said to *anticipate* the upcoming environment change. If the directions toward which the PS is moving over time are unpredictable, however, then prediction should have no effect at all in the search process. In **chapter 6**, section 6.1, this adaptive behavior is showcased in the proposed online anticipatory learning.

We thus make the case for a online MCDM strategy that can find candidate solutions requiring as little commitment as possible from the DM in terms of preference specification. Effectively, the anticipated decision mitigate the need for the DM to disclose his/her preferences over the conflicting optimization criteria at early, uncertain decision periods. The proposed strategy is thus a means for *robustly postponing preference specification* when the DM is uncertain about what the underlying trade-offs resulting from each optimization criteria are as well as how they are changing over time. The decision identified under the proposed MCDM strategy is the one for which it is *predicted* to attain, on average, the best possible diverse range of trade-off solutions in subsequent decision periods, when the DM can be in a better position to make an informed decision grounded on more established preferences.

5.2.1 Notation and Definitions

There are two possible formulations for AS-MOO: the (a) *Time-Linkage* (TL); and the (b) *Time-Linkage Free* (TLF) one; the difference being whether the objective functions at time t depend on the decisions taken at time $t-k$ ($t \geq k > 0$). Before stating the two formulations, some notation is needed. We propose modeling the dynamics of the objective functions by means of a hidden random state vector ($\mathbf{x}_t$), subject to stochastic dynamics over time. For instance, in the TL regime, one can assume the state vector to depend linearly on both its previous state and on the decision taken at time $t-1$, where the stochastic part may be modeled as additive noise represented in a vector of random disturbances ξ_t (and e.g. $\mathbf{x}_0 \sim \mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$):

$$\mathbf{x}_{t+1}|\mathbf{u}_t = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t + \xi_t, \tag{5.1}$$

whereas, for the TLF regime, one may simply take $\mathbf{B} = \mathbf{0}$ and write:

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + \xi_t. \tag{5.2}$$

Definition 5.1 (Time-Linkage regime)**.** *A TL regime is one wherein either:*
a) *the temporal evolution of the state vector* $\mathbf{x}_t$ *is influenced by external actions or decisions* $\mathbf{u}_{t-k}$ *taken at previous periods; and/or*
b) *the objective functions* $\mathbf{f}$ *evaluation of a given decision* $\mathbf{u}_t$ *depends on the decisions* $\mathbf{u}_{t-k}$ *taken at previous periods.*

Definition 5.2 (Time-Linkage Free regime)**.** *A TLF regime is one wherein both:*
a) *the temporal evolution of the state vector* $\mathbf{x}_t$ *is independent of external actions or decisions* $\mathbf{u}_{t-k}$ *taken at previous periods; and*
b) *the objective functions* $\mathbf{f}$ *evaluation of a given decision* $\mathbf{u}_t$ *is independent of the decisions* $\mathbf{u}_{t-k}$ *taken at previous periods.*

The classification of a temporal regime depicted in Definitions 5.1 and 5.2 is divided in two parts: (a) how past decisions affect the evolution of the state vector; and (b) how they affect the objective functions evaluation. In the TLF regime, both the state vector and the objective functions are independent of previous decisions, whereas in the TL regime, two situations can happen: firstly, if past decisions influence the evolution of the state vector (i.e., if $\mathbf{B} \neq \mathbf{0}$), then they also necessarily influence the current objective functions evaluation, since $\mathbf{f}$ is parameterized by the state vector $\mathbf{x}_t$. On the other hand, even when the state vector evolution is independent of past decisions, the objective functions evaluation may still be influenced by past decisions. This is the case, for instance, when there are resource costs for adapting previous decisions towards new ones [179]. The consideration of such costs may alter the optimal *anticipatory* *trajectories* in the search space over time, when compared to the scenario in which the costs are ignored.

The TL regime can be identified by the notation $\mathbf{x}_t|\mathbf{u}_{t-1:t-k}$, for $0 < k < t$, which denotes the current hidden state vector *given* a sequence of past decisions $\mathbf{u}_{t-1}, \cdots, \mathbf{u}_{t-k}$. However, in both the TL and TLF regimes, the state may not depend on past decisions. Hence, when $\mathbf{x}_t|\mathbf{u}_{t-1:t-k} \equiv \mathbf{x}_t$, the TL regime is distinguished from the TLF one by the conditional *objective vector* notation $\mathbf{z}_t|\mathbf{u}_{t-1:t-k}$, which denotes that the objective functions evaluation of a current candidate decision $\mathbf{u}_t$ depends on a sequence of decisions taken at previous periods of the decision-making process. In this thesis, we consider the *Markov property*:

$$\begin{aligned} \mathbf{x}_t|\mathbf{u}_{t-1:t-k} &\equiv \mathbf{x}_t|\mathbf{u}_{t-1}, \\ \mathbf{z}_t|\mathbf{u}_{t-1:t-k} &\equiv \mathbf{z}_t|\mathbf{u}_{t-1}, \end{aligned} \tag{5.3}$$

that is, the current state vector and objective vector conditioned on the latest k decisions depend only on the near-termly preceding decision. Moreover, we consider that the past state vectors encode all historical data that is relevant for computing the value of the stochastic objective functions at any given decision period. In general, $\mathbf{x}_{t-k}$ encodes the parameters of the underlying probability distributions generating the observed data. For instance, in the case of portfolio selection, $\mathbf{x}_{t-k}$ represent the mean vectors of a multivariate Gaussian modeling the returns distribution. On the other hand, the current state vector $\mathbf{x}_t$ may represent a distribution over the parameters of the data generating process because, generally, a decision must be taken before the current environment can be known with certainty.

Given a fixed decision $\mathbf{u}_t$, the vector of objective functions can be denoted in four different scenarios, depending on the temporal regime we are addressing:

1. In the *TLF regime* as $\mathbf{z}_t = \mathbf{f}(\mathbf{u}_t, \mathbf{x}_t)$, wherein neither the state and the objective vectors depend on previous decisions;

2. In the *TL regime* as $\mathbf{z}_t = \mathbf{f}(\mathbf{u}_t, \mathbf{x}_t|\mathbf{u}_{t-1})$ to account for the dependency of the state vector on the previous decision;

3. In the *TL regime* as $\mathbf{z}_t|\mathbf{u}_{t-1} = \mathbf{f}(\mathbf{u}_t, \mathbf{x}_t, \mathbf{u}_{t-1})$ to account for the dependency of the objective vector on the previous decision, in which case the objective functions are parameterized by both the current candidate decision being evaluated, $\mathbf{u}_t$, and the decision taken at the preceding period, $\mathbf{u}_{t-1}$; or

4. In the *TL regime* as $\mathbf{z}_t|\mathbf{u}_{t-1} = \mathbf{f}(\mathbf{u}_t, \mathbf{x}_t|\mathbf{u}_{t-1}, \mathbf{u}_{t-1})$ to account for the dependency of both the state and objective vectors on the previous decision. The objective vector notation in this case is the same as in the scenario 3.

Note: When hereafter referring to the TL regime, we assume the scenario 3, which best describes the portfolio selection application that will be presented in **chapter 7**.

The vector-valued objective function $\mathbf{f}$ can thus be decomposed in the scenario 3 as follows:

$$\mathbf{z}_t|\mathbf{u}_{t-1} = \mathbf{f}(\mathbf{u}_t, \mathbf{x}_t, \mathbf{u}_{t-1}) = \mathbf{g}(\mathbf{u}_t, \mathbf{x}_t) + \mathbf{h}(\mathbf{u}_t, \mathbf{u}_{t-1}), \tag{5.4}$$

where $\mathbf{g}$ encodes the original m conflicting objective functions of the MCDM problem and $\mathbf{h}$ denotes the costs incurring over each optimization criteria when evaluating the decision $\mathbf{u}_t$ given that the decision $\mathbf{u}_{t-1}$ had been taken.

Remark: While the functional form of $\mathbf{f}$ is static, it is parameterized by the hidden random state vector, what implies that the statistics from repeated evaluations of $\mathbf{f}$ for a fixed decision vector $\mathbf{u}_t$ may evolve over subsequent decision periods.

We denote the *Stochastic Pareto Frontier* (SPF) in the TLF regime as

$$\mathcal{Z}_t^\star|\mathbf{u}_{t-1} = \mathcal{F}_t(\Omega_t^\star, \mathbf{x}_t) = \{\mathbf{f}(\mathbf{u}_t^\star, \mathbf{x}_t) : \mathbf{u}_t^\star \in \Omega_t^\star\}, \tag{5.5}$$

where $\Omega_t^\star$ is the Pareto Set (PS) at time t and $\mathbf{f}$ is evaluated for all $\mathbf{u}_t^\star \in \Omega_t^\star$. It turns out our SPF definition is based on the Pareto dominance applied over the objective mean vectors:

Definition 5.3 (Stochastic Pareto Frontier)**.** *The Stochastic Pareto Frontier (SPF) is the set* $\mathcal{Z}_t^\star = \mathcal{F}(\Omega^\star)$ *composed of all* random *objective vectors satisfying the property*

$$\forall\ \mathbf{z}^\star \in \mathcal{Z}^\star, \nexists\ \mathbf{z}' \in \mathcal{F}(\Omega) \text{ such that } \mathbb{E}[\mathbf{z}'] \preceq \mathbb{E}[\mathbf{z}^\star]\,. \tag{5.6}$$

Put differently, the definition implies that, although the Pareto set is deterministically formed by all non-inferior solutions concerning the mean objective vectors, the SPF is a *random set* whose support are subsets of the objective space $\mathcal{Z} = \mathcal{F}(\Omega)$. An example of a SPF approximation with four mutually non-dominated random objective vectors is shown in Fig. 5.2.

Moreover, if $\mathcal{U}_t^N = \{\mathbf{u}_{t,1}, \cdots \mathbf{u}_{t,N}\}$ is a finite candidate set of N mutually non-dominated solutions at time t, then $\mathcal{Z}_t^N = \mathcal{F}_t^N(\mathcal{U}_t^N, \mathbf{x}_t)$ can be written in short as the set[1] $\mathcal{Z}_t^N = \{\mathbf{z}_{t,i}\}_{i=1}^N$. The notation used for the TL case is analogous: $\mathcal{Z}_t^N|\mathbf{u}_{t-1} = \{\mathbf{z}_{t,i}|\mathbf{u}_{t-1}\}_{i=1}^N$, for a decision $\mathbf{u}_{t-1}$ implemented at time $t-1$.

Before presenting the proposed AS-MOO models, we demonstrate a novel way of performing partial preference elicitation requiring minimal accountability and involvement for the DM.

5.2.2 Representing Partial Preferences

The assumption is that there is no access to the complete specification of the DM preferences[2]. In order to model the DM partial preferences, we therefore assume that it is possible to define a set of m constraints in the objective space (m is the number of objective functions). Hence, the DM preferences are encoded in our model as a set of m separating hyperplanes in the objective space (Fig. 5.3). Each hyperplane can be directly specified by the DM at each decision period. The set γ_t is thus used to encode the m resulting half-spaces specifying the regions on the trade-off hypersurface that might be of interest to the DM.

With our partial preference specification approach, the DM effectively reduces the problem to finding the optimal trade-off solutions whose objective vectors fall into the interior of the half-spaces, bounded by the separating hyperplanes – what leads us to the following definition:

[1]Whenever possible, we omit the state vector $\mathbf{x}_t$ from the notation, to simplify the exposition. Thus, the SPF $\mathcal{Z}_t^\star$ is simply denoted as $\mathcal{F}_t(\Omega_t^\star)$. Moreover, we denote a finite subset of the true Pareto set composed of N mutually non-dominated solutions as $\mathcal{U}_t^{N\star}$. Therefore, the corresponding finite subset of the true SPF is denoted as $\mathcal{Z}_t^{N\star}$.

[2]Otherwise, the optimization criteria could be aggregated and approached as a single-objective one.

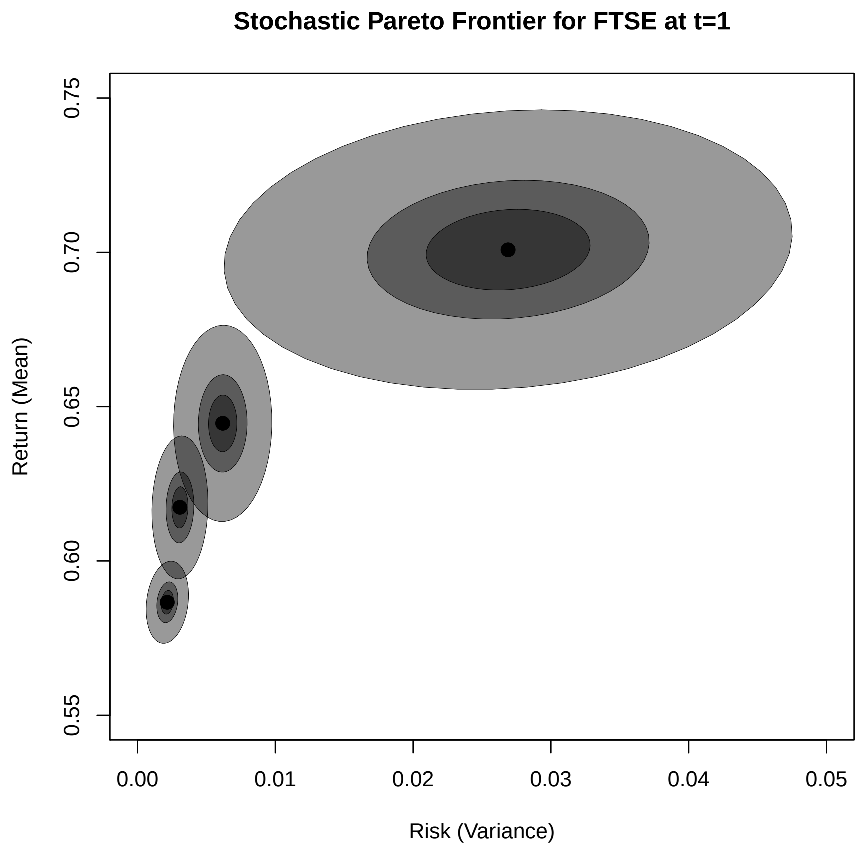


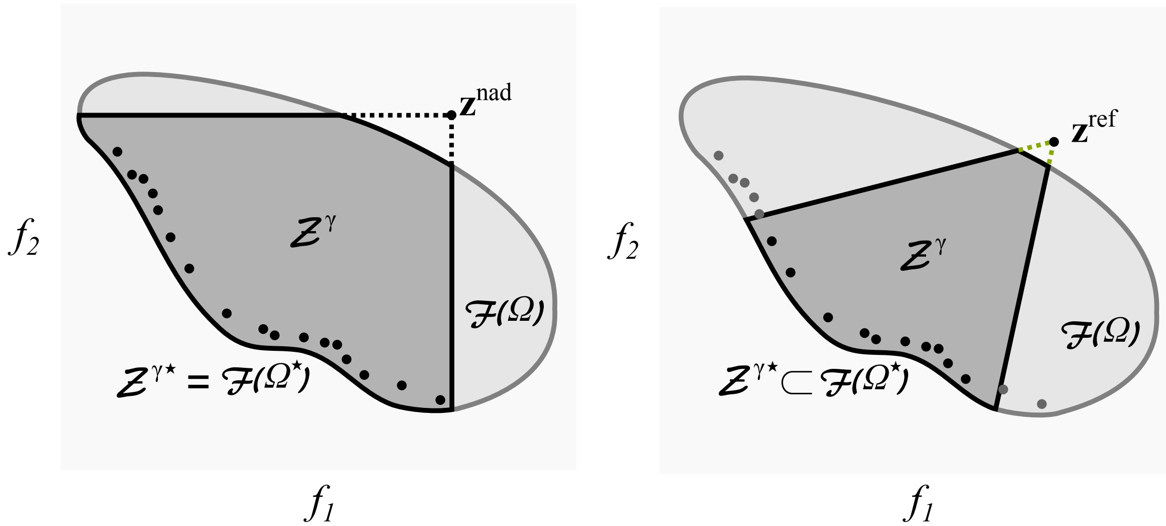
Figure 5.2: Example of a SPF approximation evolved with the proposed algorithm for the portfolio selection application. The objective functions are risk (variance, to be minimized) and mean return (to be maximized). When the evolved mutually non-dominated portfolios are evaluated under several market scenarios sampled from the ground joint assets return distribution, different risk/return evaluations are obtained. The objective function evaluations for each portfolio are then collected to estimate the parameters of the corresponding bivariate Gaussians, by computing the sample mean vectors and covariance matrices. The visualization of the SPF provides insights not only about the stability of the risk/return estimation of each portfolio to prices variability, but also about the underlying correlations between risk/return, which are noticeably different, depending on the location of the portfolio along the SPF.

Definition 5.4 (Preferred Feasible Region)**.** *The so-called* Preferred Feasible Region *(PFR) in the objective space $\mathcal{Z}^\gamma$ is the intersection between the interior of the half-spaces specified in γ and the original feasible set in the objective space.*

By optimizing over the PFR, any candidate decision $\mathbf{u}$ whose image in the objective space, $\mathbf{z} = \mathbf{f}(\mathbf{u})$, does not fall within the PFR should be disregarded. A special case occurs when defining the PFR in terms of a notable point in the original feasible space $\mathcal{Z} = \mathcal{F}(\Omega)$. We recall Definition 3.2 in **chapter 3** of a *Nadir point* $\mathbf{z}^{\text{nad}}$ and write

$$z_j^{\text{nad}} = \sup_{\mathbf{u}^\star \in \Omega^\star} f_j(\mathbf{u}^\star), \tag{5.7}$$

where $\Omega^\star$ is the Pareto Set ($\mathbf{z} = \mathbf{f}(\mathbf{u}^\star \in \Omega^\star)$ thus belongs to the Pareto Frontier, $\mathcal{F}(\Omega^\star)$). Let



(a) Canonical PFR.

(b) Non-canonical PFR.

Figure 5.3: The set of m half-spaces γ specifies the PFR $\mathcal{Z}^{\gamma}$ defined as the intersection of the half-spaces with the original feasible objective space region, $\mathcal{F}(\Omega)$.

$\mathcal{H}_j$ be the half-space composed of all objective vectors whose j-th component is less than or equal to the value of the j-th component of the Nadir point, i.e.,

$$\mathcal{H}_j = \left\{ \mathbf{z} : z_j \leq z_j^{\text{nad}} \right\}. \tag{5.8}$$

Denote the union of all the m corresponding half-spaces as

$$\boldsymbol{\mathcal{H}} = \bigcup_{j=1}^{m} \mathcal{H}_j = \left\{ \mathbf{z} : \mathbf{z} \preceq \mathbf{z}^{\text{nad}} \right\}. \tag{5.9}$$

Then, if the DM sets $\gamma = \{\mathcal{H}_1, \cdots, \mathcal{H}_m\}$, the resulting PFR is then defined as $\mathcal{Z}^{\gamma} = \boldsymbol{\mathcal{H}} \cap \mathcal{F}(\Omega)$. It turns out that this scenario corresponds to the complete absence of preference information and is then termed in this thesis as the *canonical* PFR. Conversely, any preference specification γ_t that is not defined in this way leads to a *non-canonical* PFR (Fig. 5.3 (b)), in which case the hyperplanes can be defined with respect to any reference point $\mathbf{z}^{\text{ref}}$.

Remark: For the canonical PFR defined in terms of the Nadir point, the Pareto Frontier (PF) will coincide with the optimal trade-off frontier of the PFR (i.e., $\mathcal{Z}^{\gamma\star} = \mathcal{F}(\Omega^{\star})$)[3].

The rationale behind the PFR approach is to require very little to no accountability for the DM when specifying preferences under uncertainty, as little effort would be required to specify hyperplanes to bound the regions of interest in the objective space, when compared to the precise assignment of the importances of each individual objective function. In fact, in the absence of any useful knowledge about the MOO problem characteristics, the DM may simply

[3]The extrema points of the Pareto Frontier (PF) trivially falls within $\mathcal{Z}^{\gamma}$ in this case, and since there is no point of the PF dominating the extrema, it follows the whole PF must be contained in the PFR.

abstain from such responsibility and set γ_t and decide to optimize over the canonical PFR, without a priori preference specification. Another motivation for setting a PFR is the potential for building straightforward, simple human-computer interfaces based on the visualization[4] of N objective vectors corresponding to N trade-off solutions that were available in previous periods, in which case the DM can directly specify the required half-spaces by just drawing lines into 2-dimensional projections onto planes resulting from all paired combinations of m objective functions.

5.2.3 Chance Constraints Over Partial Preferences

Because we are dealing with data-driven online SMOO problems, the objective vectors are actually *random objective vectors*, i.e., the candidate decision vectors are evaluated under different *scenarios*, that may be either simulated (sampled) from a probabilistic model randomly taken from the already available data using bootstrap techniques. It is thus a natural consequence that no candidate decision will satisfy the preferences constraints that are implicit in the chosen PFR with full certainty. The best one can do is to talk about how likely the decision is to satisfy the preferences constraints. The DM may thus require the candidate decisions to satisfy his/her preferences at a minimum probability level:

Definition 5.5 (ϵ-feasibility)**.** *Given a decision vector* $\mathbf{u}$ *with an associated random objective vector* $\mathbf{z} = \mathbf{f}(\mathbf{u})$*, we consider* $\mathbf{u}$ *to be* ϵ*-feasible with regard to the DM's PFR iff*

$$Pr\{\mathbf{z} \in \mathcal{Z}^{\gamma}\} \geq \epsilon. \tag{5.10}$$

Remark: If the stochastic uncertainty in the objective space is modeled by fitting multivariate Gaussians for each objective vector, then Eq. (5.10) can be computed by querying the joint cumulative distribution function.

5.2.4 Time-Linkage Formulation

We present in this section the proposed optimization model for the TL regime. The solution of the TL-AS-MOO model is a finite approximation for the dynamic Pareto Set (PS), $\mathcal{U}_t^{N\star}|\mathbf{u}_{t-1} \subset \Omega_t^{\star}$, obtained by maximizing the expected $\mathcal{S}$-Metric (Hypervolume, Hypv) [32] computed over the weighted sum of the temporal random objective vectors of the current candidate set of N mutually non-dominated solutions, $\mathcal{U}_t^N|\mathbf{u}_{t-1}$, and of the sequence of $H-1$ future optimal trade-off sets, $\mathcal{U}_{t+1}^{N\star}|\mathbf{u}_t, \cdots, \mathcal{U}_{t+H-1}^{N\star}|\mathbf{u}_{t+H-2}$ (H is the number of remaining decision periods and, hence, the number of steps ahead the DM must anticipate is $H-1$).

The model explicitly assesses the future discounted objective vectors, accumulating them with the observed objective vectors $\mathcal{Z}_t^N|\mathbf{u}_{t-1}$ in the current optimization environment. This combination is mediated by discount factors, λ_{t+h} (for $1 \leq h \leq H-1$), which limit the influence of future performance. Note that the $\mathcal{S}$-Metric in this thesis is a function of N random objective vectors and, thus, it turns out $\mathcal{S}$ is also a random variable.

[4]For stochastic objective vectors, the visualization may include level curves representing the underlying distributions centered at each mean objective vector.

Note: The notation $\mathcal{U}_t^N|\mathbf{u}_{t-1}$ indicates that the current candidate set of mutually non-dominated solutions is optimized under the influence of the immediately preceding decision $\mathbf{u}_{t-1}$, which is considered in the evaluation of the current objective functions. For the sake of simplicity, we omit the conditional part whenever the TL context is clear. Thus, $\mathcal{U}_t^N|\mathbf{u}_{t-1} \equiv \mathcal{U}_t^N$.

5.2.5 Anticipating Maximal Flexible Choices

The key to improve the range of future options is to anticipate flexible alternatives. Consider a finite objective set $\mathcal{Z}_t^N$ and let $\mathcal{Z}_{t+h}^{N\star}|\mathbf{u}_t^{\max}$ denote a future stochastic Pareto frontier approximation solving the AS-MOO model at time $t+h$, given that $\mathbf{u}_t^{\max} \in \mathcal{U}_t^N$ is chosen.

Definition 5.6 (Anticipated Maximal Flexible Choice, AMFC)**.** *The maximal future Hypv alternative* $\mathbf{u}_t^{max}$ *satisfies*

$$\mathbf{u}_t^{max} = \arg\max_{\mathbf{u}_t \in \Omega_t} \mathbb{E}\left[\mathcal{S}\left(\sum_{h=1}^{H-1} \lambda_{t+h} \underbrace{\mathcal{Z}_{t+h}^{N\star}|\mathbf{u}_t}_{\textit{future trade-offs}}\right)\right], \tag{5.11}$$

and is denoted as the Anticipated Maximal Flexible Choice (AMFC).

From all N choices $\mathbf{u}_t \in \Omega_t$ in Eq. (5.11), $\mathbf{u}_t^{\max}$ is foreseen to lead to future SPF approximations with maximal expected Hypv, where H is the anticipation horizon. The proposed anticipatory decision-making model limit the influence of future objective distributions, mediated by the *anticipation rates* λ_{t+h}. Hypv is thus measured over a temporal convex combination between each ordered objective distribution over time. Given two arbitrary sets, $\mathcal{A}$ and $\mathcal{B}$, of ordered elements, the summation is defined for a lexical order over the decision criteria as, e.g., $\mathcal{A} + \mathcal{B} = \{a^{(1)} + b^{(1)}, \cdots, a^{(N)} + b^{(N)}\}$, so as the multiplication by a scale factor: $\lambda\mathcal{A} = \{\lambda a^{(1)}, \cdots, \lambda a^{(N)}\}$.

Remark: The maximal future Hypv alternative $\mathbf{u}_t^{\max}$ in Eq. (5.11) can be thus interpreted as:

> *Among all candidate alternatives in the current feasible search space Ω_t, which is the one whose implementation would allow for the best possible future cost-adjusted trade-off performances, $\mathcal{Z}_{t+h}^{N\star}|\mathbf{u}_t$, in terms of expected hypervolume over finite sets of N future Pareto-efficient solutions?*

5.2.6 The TL-AS-MOO Mathematical Model

Consider a finite candidate set of trade-off objective vectors $\mathcal{Z}_t^N$ (with $|\mathcal{Z}_t^N| = N$). Let also $\mathcal{Z}_{t+h}^{N\star}|\mathbf{u}_t^{\max}$ denote a future finite trade-off set solving the TL-AS-MOO model at time $t+h$, given that a decision $\mathbf{u}_t^{\max} \in \mathcal{U}_t^N$ maximizing the future $\mathcal{S}$-Metric is taken[5]. Then the TL-AS-MOO model aims to obtain the set $\mathcal{Z}_t^{N\star}$ satisfying:

[5]For notation convenience, we sometimes use $\mathbf{z}_t \in \mathcal{Z}_t^{N\star}$ and $\mathbf{u}_t \in \mathcal{U}_t^{N\star}$ interchangeably, but, in fact, $\mathbf{u}_t$ is the actual alternative being searched for.

$$\mathcal{Z}_t^{N\star} = \arg \max_{\substack{\mathcal{Z}_t^N | \mathbf{u}_{t-1}^{\max} \\ \subset \mathcal{Z}_t | \mathbf{u}_{t-1}^{\max}}} \mathbb{E}\left[\mathcal{S}\left(\underbrace{\mathcal{Z}_t^N | \mathcal{Z}_{t+1:t+H-1}^{N\star}}_{\text{anticipatory trade-off sets}}\right)\right], \tag{5.12}$$

where,

$$\underbrace{\mathcal{Z}_t^N | \mathcal{Z}_{t+1:t+H-1}^{N\star}}_{\text{anticipatory trade-offs}} \equiv \lambda_t \underbrace{\mathcal{Z}_t^N | \mathbf{u}_{t-1}}_{\text{current trade-offs}} + \sum_{h=1}^{H-1} \lambda_{t+h} \underbrace{\mathcal{Z}_{t+h}^{N\star} | \mathbf{u}_{t+h-1}^{\max}}_{\text{future trade-offs}},$$

$$\text{s.t.} \begin{cases} \lambda_t + \sum_{h=1}^{H-1} \lambda_{t+h} = 1, \lambda_t \geq 0, \lambda_{t+h} \geq 0, \\ \mathbf{x}_t = \mathbf{T}(\mathbf{x}_{t-1}, \xi_{t-1}), \\ (\mathbf{z}_t^{(a)}, \mathbf{z}_t^{(b)} \in \mathcal{Z}_t^N) \wedge (\mathbf{z}_t^{(a)} \neq \mathbf{z}_t^{(b)}), \mathbf{z}_t^{(a)} \parallel \mathbf{z}_t^{(b)}, \\ \Pr\{\mathbf{z}_{t+h} \in \mathcal{Z}_{t+h}\} \geq \epsilon, \ \forall\ h \in \{0, \cdots, H-1\}. \end{cases}$$

Remark: The maximal anticipatory Hypv set $\mathbf{Z}_t^{N\star}$ in Eq. (5.12) can be thus interpreted as:

> *Among all candidate sets of N mutually non-dominated vectors in the feasible objective space $\mathcal{Z}_t$, which is the one attaining the best possible* ***combined*** *cost-adjusted trade-off performance in terms of expected hypervolume evaluated in the current and in the $H-1$ subsequent optimization environments?*

In order to assess $\mathcal{Z}_t^{N\star}|\mathbf{u}_{t-1}$, however, one must first obtain *in advance* $\mathbf{u}_t^{\max}, \cdots, \mathbf{u}_{t+H-2}$ and $\mathcal{Z}_{t+1}^{N\star}, \cdots \mathcal{Z}_{t+H-1}^{N\star}$, i.e., the AS-MOO problem at time t depends on the solutions obtained for periods $t+1, \cdots, t+H-1$. That means the TL-AS-MOO model is a recurrence model.

Note: The set $\mathcal{Z}_{t+h}^{N\star}|\mathbf{u}_{t+h-1}^{\max}$ represents the future SPF at time $t+h$ maximizing the expected $\mathcal{S}$ values computed for even further periods at $t+h+1, \cdots, t+h+H-1$, following future preceding choices leading to maximal subsequent future Hypv for the whole anticipation horizon. It is worth noting that we assume the *Markov property*. The set $\mathcal{Z}_{t+h}^{N\star}|\mathbf{u}_t$ then denotes the decision-tree path (Fig. 5.4):

$$\mathcal{Z}_{t+1}^N|\mathbf{u}_t \to \mathcal{Z}_{t+2}^N|\mathbf{u}_{t+1} \to \cdots \to \mathcal{Z}_{t+h}^N|\mathbf{u}_{t+h-1} \equiv \mathcal{Z}_{t+h}^N|\mathbf{u}_t. \tag{5.13}$$

Note: In Eq. (5.12), the DM willingness for near-term optimized performance – i.e., for maximal hypervolume non-dominated sets given the current available historical data – will be maximal when $\lambda_t = 1$, being equivalent to not looking ahead into the future at all, whereas, for $\lambda_t = 0$, the DM is willing to achieve maximum performance in future decision periods, even if that means the non-dominated trade-off solutions do not yield the best possible performance when evaluated from the available historical data.

Note: In Eq. (5.12), we have the following cost-adjusted trade-off objective vector:

$$\mathcal{Z}_{t+h}^{N\star}|\mathbf{u}_{t+h-1}^{\max} = \{\mathbf{f}(\mathbf{u}_{t+h}^{(1)}, \mathbf{x}_{t+h}, \mathbf{u}_{t+h-1}^{\max}), \cdots, \mathbf{f}(\mathbf{u}_{t+h}^{(N)}, \mathbf{x}_{t+h}, \mathbf{u}_{t+h-1}^{\max})\}, \tag{5.14}$$

where $\mathbf{u}_{t+h-1}^{\max}$ is the preceding implemented future AMFC (Eq. (5.11)). This model implies that the future implemented AMFCs must also be estimated (see Eq (5.15)).

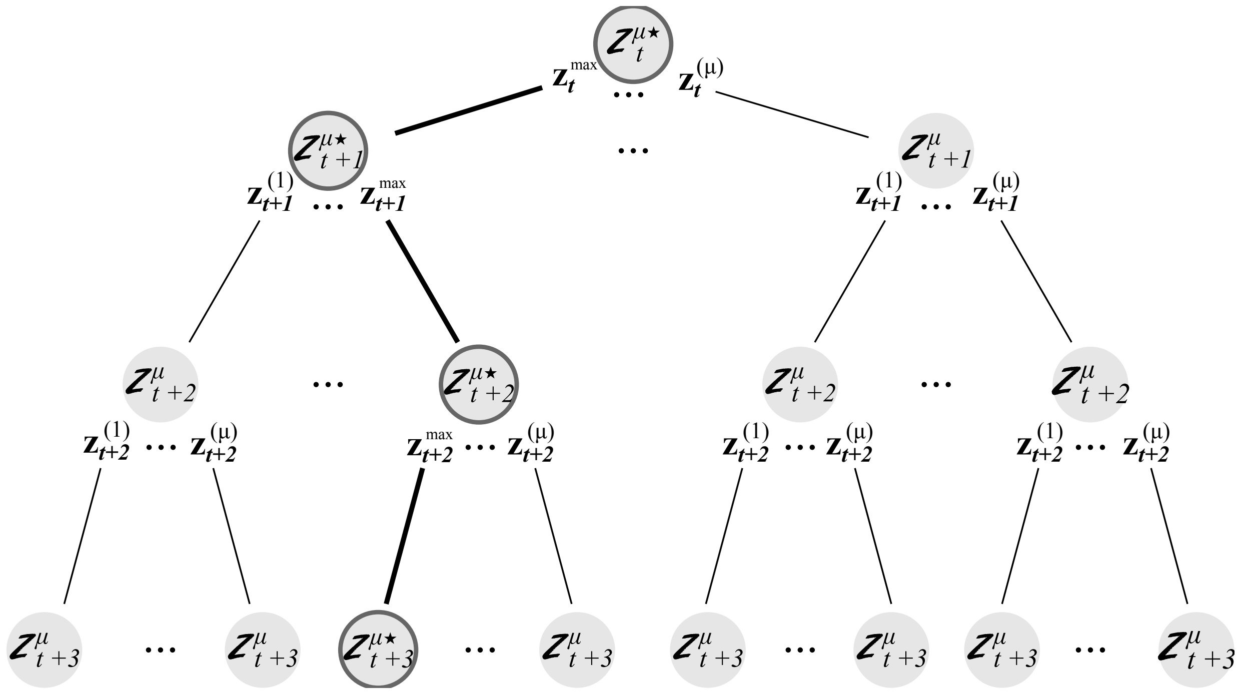


Figure 5.4: In the AS-MOO methodology, a Pareto-efficient alternative $\mathbf{z}_t^{\max}$ is chosen among N options in $\mathcal{Z}_t^N$. This choice influences the future stochastic Pareto frontier approximations that can be formed to maximize the expected Hypv over discounted future efficient objective vectors, $\mathcal{Z}_{t+h}^{N\star}|\mathbf{z}_t^{\max}$.

5.2.7 Approximate Predictive Bayesian Solutions to TL-AS-MOO

In the TL-AS-MOO solvers proposed in this thesis (see **chapter 6**, section 6.5), however, the future sets of Pareto-optimal decisions, $\mathcal{U}_{t+h}^{N\star}|\mathbf{u}_t$, are replaced with the *predicted sets* $\hat{\mathcal{U}}_{t+1}^{N\star}|\mathbf{u}_t$, $\cdots, \hat{\mathcal{U}}_{t+H-1}^{N\star}|\mathbf{u}_t$ obtained from *tracking* the candidate trade-off solution vectors over time – along with the associated objective vectors – with *internal Bayesian models*. The future AMFCs are also replaced with the predictions $\hat{\mathbf{u}}_{t+1}^\star, \cdots, \hat{\mathbf{u}}_{t+H-1}^\star$, for which no online re-optimizations are necessary. Therefore, our anticipatory methodology is a means for *approximately* solving AS-MOO models.

Moreover, upon obtaining $\hat{\mathcal{U}}_t$ after approximately solving the TL-AS-MOO model for the current decision period at t, the *Anticipated Maximal Flexible Choice* $\hat{\mathbf{u}}_t^\star \in \hat{\mathcal{U}}_t^{N\star}$ is chosen from the optimal trade-off set which was the solution to Eq. (5.12), i.e. $\mathcal{U}_t^{N\star}$, what leads to the following definition.

Definition 5.7 (Estimated Maximal Flexible Choice, EMFC)**.** *The Estimated Maximal Flexible Choice (EMFC) $\hat{\mathbf{u}}_t^\star$ for the current decision period is the one chosen from the finite set of mutually non-dominated solutions $\hat{\mathcal{U}}_t^{N\star}$ approximately satisfying Eq. (5.12) such that*

$$\hat{\mathbf{u}}_t^\star = \underset{\hat{\mathbf{u}}_t \in \hat{\mathcal{U}}_t^{N\star}}{\arg\max}\, \mathbb{E}\left\{ \mathcal{S}\left(\sum_{h=1}^{H-1} \lambda_{t+h} \underbrace{\hat{\mathcal{Z}}_{t+h}^{N\star}|\hat{\mathbf{u}}_t}_{\textit{predicted trade-offs}} \right) \right\}, \tag{5.15}$$

where $\hat{\mathcal{Z}}_{t+h}^{N\star}|\hat{\mathbf{u}}_t$ *is the predicted maximal Hypv stochastic Pareto frontier approximation of N alternatives that would be obtained at time* $t+h$*, given that an alternative* $\hat{\mathbf{u}}_t \in \hat{\mathcal{U}}_t^{N\star}$ *is taken at a current decision period.*

It is also worth noting we assume the state vector $\mathbf{x}_t$ to be *unknown* at the current decision period. However, it is possible to come up with the estimate $\hat{\mathbf{x}}_t$ from the *available historical data stream* $\{\mathbf{x}_0, \cdots, \mathbf{x}_{t-1}\}$. In other words, the whole sequence of previous states is known and can be inputted as a *training set* into an algorithm for approximately solving TL-AS-MOO. The AMFC ($\hat{\mathbf{u}}_t^\star$) obtained after solving the TL-AS-MOO model at the current decision period is then implemented before the data at t can be observed. Put another way, the anticipatory decision $\hat{\mathbf{u}}_t^\star$ is to be implemented in an *unseen environment*.

5.2.8 Effects of Not Committing to the Anticipated Maximal Flexible Choice

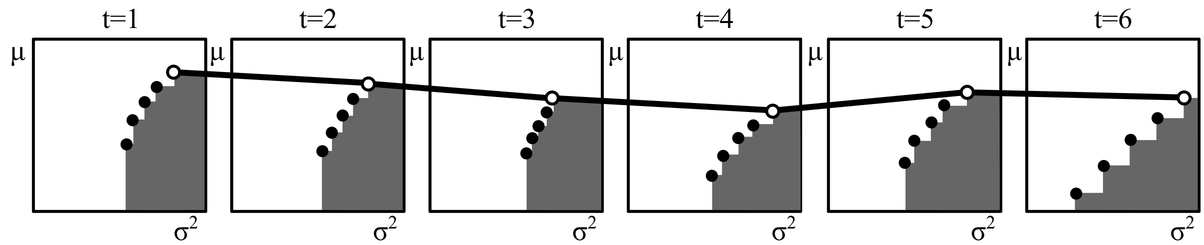
We intuitively examine in this section the nature of the system trajectories resulting from the TL-AS-MOO flexible choices. It turns out that a DM without a complete a priori preference specification could choose a candidate solution randomly drawn from $\hat{\mathcal{U}}_t^{N\star}$, say, $\hat{\mathbf{v}}_t$. In this case, regardless of the form assumed by the state transition function $\mathbf{T}$, it is clear from Eq. (5.11) that $\hat{\mathbf{v}}_t$ is not guaranteed to lead to the maximization of the expected hypervolume in future decision periods, implying in suboptimal future Stochastic Pareto Frontiers (SPFs) in terms of *expected hypervolume*. This means that: (1) the resulting trajectories can be biased towards favoring one objective over another; (2) higher costs can be induced; and (3) the diversity and quality of future admissible options can be compromised.

The implementation of biased decisions in the TL-AS-MOO model can therefore reduces the range of future options to the DM. It follows that, if a DM with undefined preferences deviates from the TL-AS-MOO decision-making process depicted in Eqs. (5.12)–(5.15) (i.e., if he/she abstains from predicting the future consequences of implementing each available trade-off solution), as he/she narrows down his/her preferences in later decision periods, the resulting sets of available non-dominated solutions at those future instants are not expected to attain maximal hypervolume. As a consequence, they are also not expected to attain as much diversity and coverage over the future SPFs as they could have for a DM following Eqs. (5.12)–(5.15), thus undermining the range of trade-off solutions compliant with *future preferences* (see Fig. 5.5).

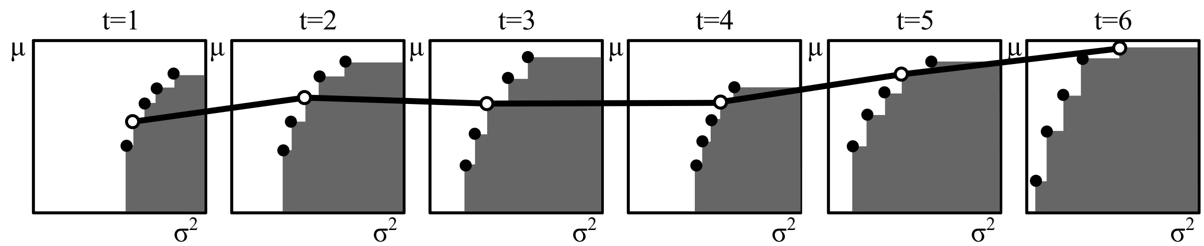
Finally, it should be made clear that, because of the constrained size of the flexible-Pareto Set, implementing the EMFC by solving Eq. (5.15) implies that the DM is in fact choosing future alternative paths over others under the influence of noise and of available dynamical predictive models.

5.2.9 Anticipating Future Preferences

What remains is to discuss in light of the TL-AS-MOO model the effects of advancing through time on the preferences chance constraints of Eq. (5.12), when computing $\mathcal{U}_{t+h}^{N\star}$ for $h = 1, \cdots, H-1$ (or when predicting $\hat{\mathcal{U}}_{t+h}^{N\star}$). It turns out there is no way to compute in advance the future Preferred Feasible Regions (PFRs) $\mathcal{Z}^{\gamma_{t+1}}, \cdots, \mathcal{Z}^{\gamma_{t+H-1}}$, since the TL-AS-MOO model does not



(a) An example of a myopic trajectory.



(b) An example of an anticipatory maximal Hypv trajectory.

Figure 5.5: Example of (a) myopic; and (b) anticipatory sequences of choices in an objective space composed of mean return maximization (N) and variance (risk) minimization (σ^2). In (a), a myopic DM always choose the most risky Pareto-efficient investment, hoping to maximize near-term wealth. The implemented portfolios influence the future diversity of choice measured by the Hypv (gray area). In (b), a DM anticipating the future maximal Hypv portfolios obtains more efficient and diverse future range of options, and may end up earning higher returns over time by taking safer, flexible choices.

prescribe how a DM should update his/her preferences in response to the observed performance of his/her decisions. Therefore, because the PFRs at every future decision period are unknown, and no reasonable dynamical model is available in principle, the assumption used in TL-AS-MOO is that the PFRs remain constant at future periods, i.e., the DM does not alter his/her uncertainty over time about the relative importances of each objective function.

The investigation of behavioral models that could explain how a DM would adjust the PFR over time as a function of the observed past outcomes is therefore an interesting open problem. Despite the absence of such a model, in the approximate TL-AS-MOO solvers proposed in this thesis, this problem can be analogously handled as in how the objective vectors and trade-off solutions are predicted: by tracking the hyperplanes coefficients in $\gamma_1, \cdots \gamma_t$ and by predicting $\hat{\gamma}_{t+1}$, and so forth. If any trend regarding the DMs preferences evolution is captured (e.g. progressively narrowing preferences down towards a specific direction), this information can enable the TL-AS-MOO model to obtain approximate ϵ-feasible Anticipated Maximal Flexible Choices w.r.t. the predicted PFRs $\hat{\mathcal{Z}}^{\gamma_{t+h}}$.

5.2.10 Time-Linkage Free Formulation

The TL-AS-MOO formulation assumed that one decision per period should be taken. There are, however, scenarios for which a *fixed decision* is intended to operate for multiple periods. It is so e.g. when the costs to modify the solution in execution are prohibitive, what leads to

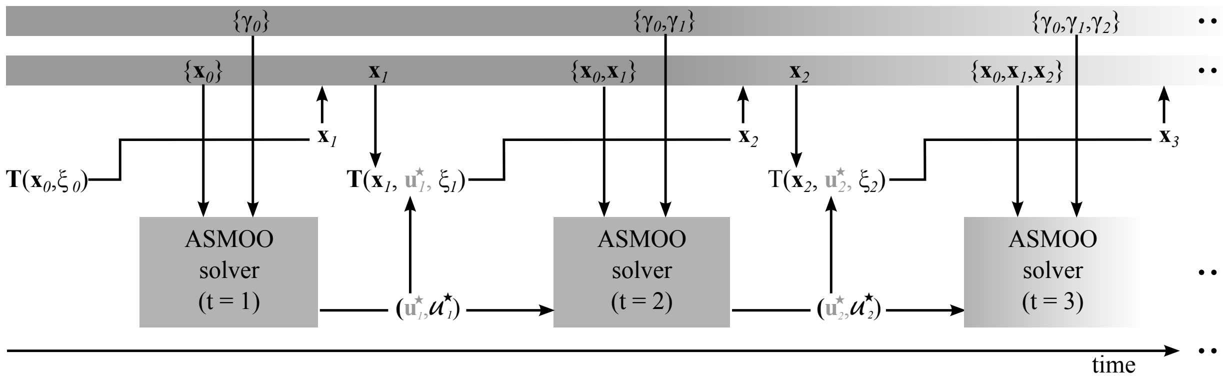


Figure 5.6: Outline of the AS-MOO methodology for online MCDM under uncertainty.

lower frequency re-optimizations. There may also be time constraints that make it impossible to keep the pace of environmental change for optimally adapting current decisions [52]. In such scenarios, a fixed decision should therefore be made robust against all possible disruptive future variations in the operational environment.

The *TLF-AS-MOO model* thus conveys a similar concept as that of the recently proposed single-objective optimization model known as *Robust Optimization Over Time* (ROOT) [90,125], whose aim is to find solutions requiring minimal to no change for maintaining good performance in future environments. The intuitive concept of ROOT was described in Jin et al. [125] as the ability of finding solutions "*(...) whose quality is acceptable over a certain time interval, although they* [may] *not be the global optima at any time instant*".

The task of obtaining fixed decisions is conceptually easier when the dynamics of the problem is not affected by the DM. The TLF-AS-MOO model therefore does not prescribe a decision-making strategy, since that is irrelevant for the dynamics of the problem. Nevertheless, a DM solving a TLF-AS-MOO problem is interested in obtaining a *fixed* finite approximation to the changing Pareto Set composed of N mutually non-dominated solutions that is *robust* against predicted changes in the optimization environment over time. The model is defined as:

$$\mathcal{U}_t^{N\star} = \underset{\mathcal{U}_t^N \subset \Omega_t}{\arg\max}\, \mathbb{E}\left\{ \mathcal{S}\left(\lambda_t \mathcal{F}_t^N\left[\mathcal{U}_t^N\right] + \sum_{h=1}^{H-1} \lambda_{t+h} \mathcal{F}_{t+h}^N\left[\mathcal{U}_t^N\right] \right) \right\}, \tag{5.16}$$

$$\text{s.t.} \begin{cases} \mathbf{x}_t = \mathbf{T}(\mathbf{x}_{t-1}, \xi_{t-1}), \\ \mathbf{u}_t \not\prec \mathbf{v}_t \ \wedge \ \mathbf{v}_t \not\prec \mathbf{u}_t \ (\text{for } \mathbf{u}_t, \mathbf{v}_t \in \mathcal{U}_t^N \text{ and } \mathbf{u}_t \neq \mathbf{v}_t), \\ \Pr\left\{\mathbf{z}_{t+h} | \mathbf{u}_{t+h-1} \in \mathcal{Z}^{\gamma_{t+h}}\right\} \geq \epsilon, \ \forall\ (0 \leq h \leq H-1), \\ \quad \text{where } \mathbf{z}_{t+h} = \mathbf{f}(\mathbf{u}_{t+h}, \mathbf{x}_{t+h}) \in \mathcal{F}_{t+h}^N(\mathcal{U}_{t+h}^N). \end{cases} \tag{5.17}$$

Contrasting the two AS-MOO models with each other, it is clear from Eqs. (5.16) and (5.17) that not only the vector-valued objective function $\mathbf{f}$ in the TLF variant does not depend on past decisions, but also the future random objective vectors within the right-hand side of the summation are computed over the *fixed* candidate approximation trade-off set, $\mathcal{U}_t^N$. Note that the term in Eq. (5.16) denotes the future set of random objective vectors when the candidate

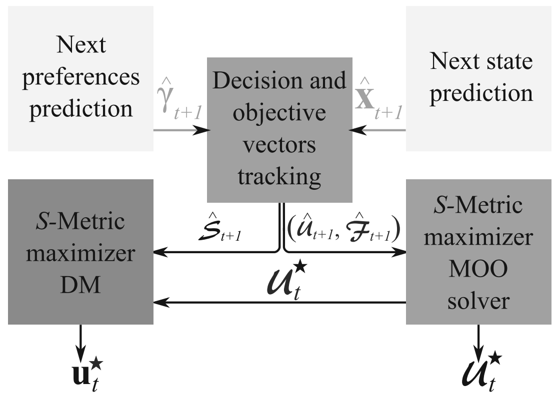


Figure 5.7: Interrelations between the five components of the proposed AS-MOO method.

trade-off solutions in $\mathcal{U}_t^N$ are evaluated over the upcoming environments, i.e.,

$$\mathcal{F}_{t+h}^N(\mathcal{U}_t^N) = \{\mathbf{f}(\mathbf{u}_{t,1}, \mathbf{x}_{t+h}), \cdots, \mathbf{f}(\mathbf{u}_{t,N}, \mathbf{x}_{t+h})\}. \tag{5.18}$$

For those reasons, the TLF-AS-MOO model does not imply a recurrence equation and the only way to assess the future objective values for a current candidate decision is hence by means of *prediction*. As likewise indicated in section 5.2.4, this thesis proposes tracking the evolution of a given objective vector over time with Bayesian models, by taking advantage of the observed historical data stream $\{\mathbf{x}_0, \cdots \mathbf{x}_{t-1}\}$ at disposal.

Figure 5.2.10 outlines the information flow in both two proposed AS-MOO models, wherein it is emphasized the unawareness of an AS-MOO solver about the exact value of the current state vector $\mathbf{x}_t$, whose outcome is only revealed after an anticipatory decision is taken (in the TL regime) and/or a finite approximation to the Pareto set is obtained. Also noteworthy is how the available historical data is accumulated over time, allowing the AS-MOO solver to more accurately track the objective vectors/decision vectors/preferred feasible regions bounds distributions.

Remark: The AS-MOO problem is repeatedly solved for each decision period upon the input of new data from the optimization environment.

A general sketch of the AS-MOO solver for approximating the dynamic PF and selecting a trade-off solution from the anticipatory SPF can be visualized in Fig. 5.2.10, assuming a one step ahead prediction (anticipation horizon of, $H = 2$). The result of one application of an AS-MOO solver at a decision period t is a set of mutually non-dominated ϵ-feasible candidate solutions ($\mathcal{U}_t^{N\star}$) maximizing expected hypervolume for $H-1$ steps ahead (for both TL and TLF variants) and the indication of the AMFC $\mathbf{u}_t^\star \in \mathcal{U}_t^{N\star}$ that is expected to maximally preserve the DM partial preference specification for the periods $t+h$ ($1 \leq h \leq H$), for the TL variant.

There are five possible modules that can be integrated into AS-MOO, namely: (1) a predictor for the future preference specification of the DM; (2) a predictor for the next state of the optimization environment; (3) a Bayesian estimation module for tracking the decision and objective vectors over time and to compute the corresponding predictive distributions; (4) an

$\mathcal{S}$-Metric MOO maximizer; and (5) a procedure for identifying the AMFC that maximizes the expected hypervolume for the future decision periods.

The light gray blocks (1), (2), and (5) are optional, since in the first case, the (partial) preferences of the DM can be given a priori. For instance, when there are multiple DMs with a vast range of preferences using the proposed anticipatory MCDM system, the canonical PFRs can be assumed throughout. The second module is optional because it can be very costly and challenging to identify a proper dynamical model for the environment, what requires the AS-MOO methodology to anticipate robust decisions by relying on its own *internal prediction models* which operate over the decision and objective spaces. That is exactly what module (3) is designed for. The methods used in the third module are described in **chapter 6**, section 6.1, whereas the anticipatory methods for approximately performing the tasks in module (4) and module (5) are described in section 6.5.

5.3 Summary of the Contributions

This chapter's contributions to the thesis are as follows:

1. It argued for anticipation as a means to handle uncertainty in Multi-Criteria Decision-Making (MCDM);

2. It presented novel and useful definitions such as Stochastic Pareto Frontier (SPF) and Preferable Feasible Region (PFR) to allow for a better and more comprehensible modeling of Stochastic Multi-Objective Problems (SMOOPs);

3. Two novel Anticipatory SMOO (AS-MOO) models have been formulated in terms of the maximization of hypervolume over time;

4. The concept of preference for flexibility has been explicitly incorporated into the AS-MOO model by suggesting an automated online decision-making strategy for choosing a solution which is foreseen to yield to future stochastic Pareto Frontiers of maximal hypervolume;

5. Prediction has been susggested to reduce the exponential computational costs of traversing the decision-tree which is implicit when solving the time-linkage AS-MOO model.

In the next chapter, we provide Bayesian anticipatory learning models and anticipatory multi-objective meta-heuristics for evolving a population of candidate trade-off choices approximating the anticipatory stochastic Pareto Frontier. Moreover, a procedure to identify the Anticipated Maximal Flexible Choice (AMFC) is proposed.

Chapter 6

Learning to Anticipate Flexible Trade-off Choices

Humility is the only true wisdom by which we prepare our minds for all the possible changes of life.

– **George Arliss**

The important thing is this: to be able at any moment to sacrifice what we are for what we could become.

– **Charles Du Bos**

While future hypervolume maximization can postpone the assignment of relative importances over the decision criteria, the DM hesitation *about time performance* can be handled via Online Anticipatory Learning (OAL). OAL refer to methods for (i) *self-adjusting* the DM willingness to near-term performance; and (ii) for incorporating predictive knowledge into Anticipatory Stochastic Multi-Objective Optimization (AS-MOO) solvers, mediated by the perceived *temporal predictability* based both on historical trajectory errors and predictive knowledge.

The assumption in OAL is that the more unpredictable the future, the more eager for immediate performance the DM should be, since, by construction, optimizing solutions for early performance is *safer* in this case. Conversely, the more certain the future, the more the DM is willing to capitalize on foreseen opportunities. OAL thus make use of Bayesian tracking in *both the objective and the search spaces* to approximate AS-MOO solutions for which improved performance on the targeted environments is predicted, in terms of PD.

In AS-MOO, the set $\Lambda_{t:t+H-1} = \{\lambda_t, \cdots, \lambda_{t+H-1}\}$ encodes time preferences. In dynamic models, such discount factors are often heuristically determined *a priori* as a monotonically decreasing function – often decaying at an exponential rate – over time, in order to model a DM that prefers more near-term optimized performance. As previously discussed in **chapter 3**, section 3.3.1, the parameterization of an optimization model can drastically alter the obtained sequence of decisions satisfying the chosen parameters. Furthermore, eliciting $\Lambda_{t:t+H-1}$ when N incomparable solutions are to be *simultaneously* searched for is challenging: the future outcomes of some alternatives may be more predictable than those of others. Hence, instead of pursuing a *global* set, we propose self-adjusting N independent sets, $\Lambda^{(1)}_{t:t+H-1}, \cdots, \Lambda^{(N)}_{t:t+H-1}$, so that portions of the stochastic Pareto frontier can be approximated under varying time preferences.

In the following, we present the proposed OAL methods for incorporating predictive knowledge into (a) the objective (or *performance*) space; and (b) the search space of an AS-MOO solver. We hereafter refer to any $\lambda \in \Lambda_{t:t+H-1}$ as the *anticipation rate*.

6.1 Online Anticipatory Learning in the Objective Space

Since the current environment $\mathbf{x}_t$ is unknown, for the sake of applying OAL in the objective space, we need a way to estimate the current objective values (performance levels) $\mathbf{z}_t = (f_1(\mathbf{u}_t, \mathbf{x}_t)\ \cdots\ f_m(\mathbf{u}_t, \mathbf{x}_t))^\intercal$ for a fixed trade-off solution vector[1], based on the *observed* trajectory $\mathbf{Z}_{t-k:t-1} = \{\mathbf{z}_{t-k}, \cdots, \mathbf{z}_{t-1}\}$ (see Fig. 6.1). The underlying stochastic uncertainty regarding the true value of $\mathbf{z}_t$ is modeled as a multivariate Gaussian distribution over the $j = 1, \cdots, m$ values received from each of the m objective functions, i.e., $\mathbf{z}_t \sim \mathcal{N}(\mathbf{m}_{\mathbf{z}_t}, \mathbf{\Sigma}_{\mathbf{z}_t})$, for any decision period at $t-k, \cdots, t, \cdots, t+H-1$. For a given candidate trade-off solution, we thus propose *predicting* the unknown objective vector $\mathbf{z}_t$ by using the Kalman Filter (KF) estimation presented in **chapter 2**, see section 2.3.4.

6.1.1 KF Prediction in the Objective Space

The notation $\hat{\mathbf{z}}_{t+h}|\mathbf{z}_{t-1}$ (for $h = 1, \cdots, H-1$) implies that the prediction step of the KF is repeatedly applied h times to yield a long term estimation about the performance levels distribution at future optimization environments:

$$\hat{\mathbf{z}}_t|\mathbf{z}_{t-1} \to \hat{\mathbf{z}}_{t+1}|\hat{\mathbf{z}}_t \to \cdots \to \hat{\mathbf{z}}_{t+h}|\hat{\mathbf{z}}_{t+h-1} \equiv \hat{\mathbf{z}}_{t+h}|\mathbf{z}_{t-1}. \tag{6.1}$$

The main idea is to use the KF to predict the distributions of $\hat{\mathbf{z}}_t|\mathbf{z}_{t-1}$ and then to combine the prediction with the current estimated distributions for $\mathbf{z}_t$ *mediated by the underlying uncertainty* measured from the current and predicted states (see section 6.1.2).

We assume simple linear dynamics for the objective vectors: $\mathbf{z}_t \sim \mathbf{z}_{t-1} + \mathbf{B}\mathbf{u}_{t-1} + \dot{\mathbf{z}}_{t-1}$, i.e., the next state is distributed as the previous state plus a small (deterministic) displacement, that we call *velocity*, $\dot{\mathbf{z}}_t$. The dynamic also accounts for an external action or decision $\mathbf{u}_{t-1}$, for which $\mathbf{B} = \mathbf{0}$ in the time-linkage free regime. We thus want to track the extended state vector via KF estimation:

$$\mathbf{z}^+_t = (z_{1,t} \cdots z_{m,t}\ \dot{z}_{1,t} \cdots \dot{z}_{m,t})^\intercal, \tag{6.2}$$

where m is the number of objective functions. It can be noted that the KF will not only estimate the current value of each objective function for a given decision $\mathbf{u}_t$, but also how each of them are changing over time (i.e., by estimating the velocities). All those information are encoded in the estimation for the joint distribution of $\mathbf{z}^+_t$ (Fig. 6.1).

Hence, the dynamical model is represented as $\mathbf{z}^+_t = \mathbf{A}\mathbf{z}^+_{t-1}$, where

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_{m\times m} & \mathbf{I}_{m\times m} \\ \mathbf{0}_{m\times m} & \mathbf{I}_{m\times m} \end{pmatrix}_{2m\times 2m}, \tag{6.3}$$

[1]Recall that in the TL regime, we would write $\mathbf{z}_t|\mathbf{u}_{t-1} = (f_1(\mathbf{u}_t, \mathbf{x}_t, \mathbf{u}_{t-1})\ \cdots\ f_m(\mathbf{u}_t, \mathbf{x}_t, \mathbf{u}_{t-1}))^\intercal$ instead.

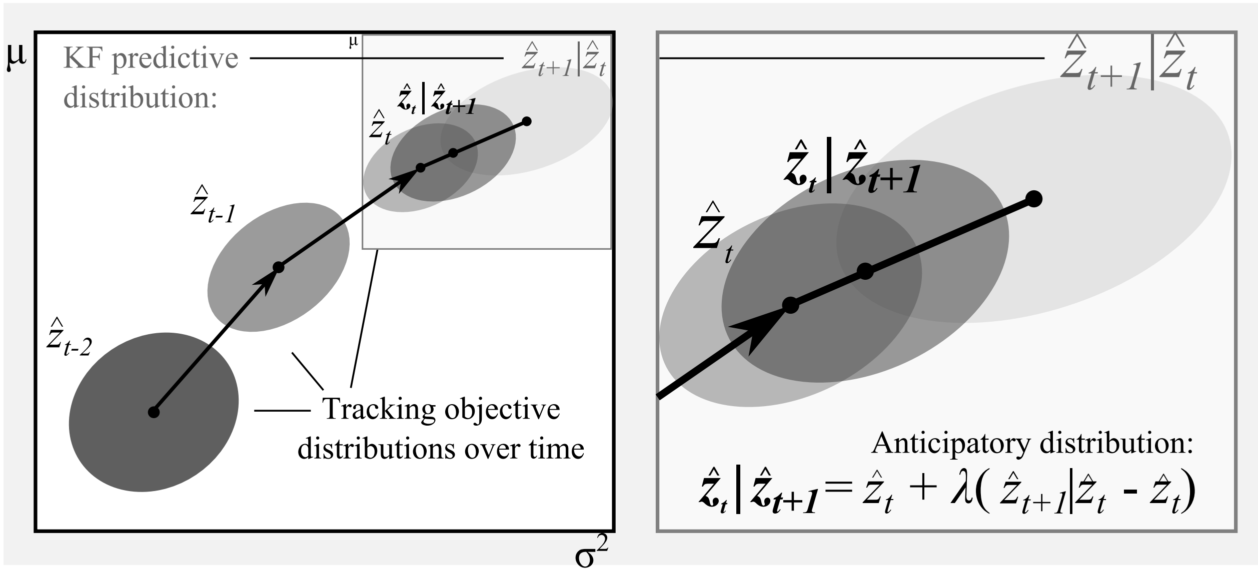


Figure 6.1: On the left, the KF tracks performance for a fixed investment portfolio and estimates the predictive distribution $\hat{\mathbf{z}}_{t+1}|\hat{\mathbf{z}}_t$. The portfolio evolves from a low-risk/low-return estimative to a high-risk/high-return prediction. On the right, OAL is detailed. Given the up-to-date estimation $\hat{\mathbf{z}}_t$ and the predictive $\hat{\mathbf{z}}_{t+1}|\hat{\mathbf{z}}_t$, online anticipatory learning produces an anticipatory distribution $\hat{\mathbf{z}}_t|\hat{\mathbf{z}}_{t+1}$ expressed as the convex combination between $\hat{\mathbf{z}}_t$ and $\hat{\mathbf{z}}_{t+1}|\hat{\mathbf{z}}_t$.

whereas the measurement function $\mathbf{M}$ outputs the noisy values for the first half of $\mathbf{z}^+$. A point worth mentioning is that it is not always possible to apply the KF directly to the decision variables, especially when dealing with problems where the search space has some structure, what is the case of the financial portfolio applications discussed in **chapter 4** (see also section 6.2). The Pseudocode 5 describes how the KF operates to estimate and predict the current and future objective vector distributions associated with a given, *fixed* candidate trade-off solution $\mathbf{u}_t$.

Remark: In the TL regime, the KF prediction takes into account the historical sequence of implemented Estimated Maximal Flexible Choices (EMFCs, see Eq (5.15), **chapter 5**) $\mathcal{U}^{N\star}_{t-K:t-1}$ and computes the historical random objective vectors for a given fixed decision vector $\mathbf{u}$ as $\mathbf{z}_{t-k} = \mathbf{f}(\mathbf{u}_{t-k}, \mathbf{x}_{t-k}, \mathbf{u}^{\star}_{t-k-1})$, for $k = 1, \cdots, K$.

6.1.2 Preference for Long-Term Predictability

The key for self-adjusting the $H \times N$ anticipation rates is to model a DM exhibiting *preference for long-term predictability* w.r.t. the PD over temporal performance. It is assumed that prediction is completely incorporated by a DM as the probability of either of the following events approaches one (i.e., Pr{event} $\to$ 1}):

1. The predicted future objective distribution *dominates* the current objective distribution;
2. The predicted future objective distribution *is dominated by* the current objective distribution; or

Input: The prediction horizon H
Input: A decision vector $\mathbf{u} \in S^{d-1}$
Input: A historical state vector data stream $\mathcal{X}_{t-K:t-1} = \{\mathbf{x}_{t-K}, \cdots, \mathbf{x}_{t-1}\}$
Output: A H step ahead prediction $\hat{\mathbf{z}}_{t+H}$

procedure KFPREDICTION(H, $\mathbf{u}_t$, $\mathcal{X}_{t-K:t-1}$)
  for $k = K$ **to** 1 **do**
    Predict $\hat{\mathbf{z}}_{t-k}|\dot{\mathbf{z}}_{t-k-1}$ using Eq. (2.26)
    Update $\hat{\mathbf{z}}_{t-k}|\mathbf{z}_{t-k}$ using Eq. (2.32)
  end for
  for $h = 0$ **to** H **do** ▷ H steps-ahead prediction
    Predict $\hat{\mathbf{z}}_{t+h}|\hat{\dot{\mathbf{z}}}_{\mathbf{t+h-1}}$ using Eq. (2.26)
  end for
  return $\hat{\mathbf{z}}_{t+H}|\mathbf{z}_{t-1}$
end procedure

Pseudocode 5: Kalman Filter Tracking and Prediction

3. The predicted future objective distribution *is incomparable to* the current objective distribution.

Prediction is thus ignored when the DM cannot tell above the chance level whether one of the three events (1)–(3) is satisfied, in which case near-term performance is preferred. Therefore, prediction is only useful to AS-MOO when it can yield safe information about the PD relation of temporal performance. In order to self-adjust the degree to which prediction is incorporated into AS-MOO, it suffices to estimate the probability for the incomparability case.

It is worth noting that, in the KF prediction, the total variance of $\hat{\mathbf{z}}_{t+h}|\mathbf{z}_{t-1}$ for $h \geq 1$ is always larger then that of $\hat{\mathbf{z}}_t|\mathbf{z}_{t-1}$ [207]. In fact, the more one looks ahead into the future, the higher the total variance is. In order to take advantage of this property, we define $\Lambda_{t:t+H-1}$ as a function of the underlying *temporal uncertainty* measured over the distributions of the current ($\hat{\mathbf{z}}_t|\mathbf{z}_{t-1}$) and the future ($\hat{\mathbf{z}}_{t+h}|\mathbf{z}_{t-1}$) objective vectors predicted with the KF.

Because the Pareto Dominance is the figure of merit to determine when a decision is preferred to another, we propose defining the uncertainty measure regarding $\hat{\mathbf{z}}_t|\mathbf{z}_{t-1}$ and $\hat{\mathbf{z}}_{t+h}|\mathbf{z}_{t-1}$ as a function of the *Temporal Non-Dominance Probability* (TIP) between the two objective vector predictions.

Definition 6.1 (Temporal Non-Dominance Probability)**.** *The probability of the current and the future predicted random objective vectors being mutually non-dominated – and therefore not comparable to each other is expressed as*

$$
\begin{aligned}
p_{t,t+h} &= Pr[\hat{\mathbf{z}}_t \parallel \hat{\mathbf{z}}_{t+h}|\hat{\mathbf{z}}_t] & (6.4)\\
&= Pr[\hat{\mathbf{z}}_t|\mathbf{z}_{t-1} \not\preceq \hat{\mathbf{z}}_{t+h}|\mathbf{z}_{t-1} \textit{ and } \hat{\mathbf{z}}_t|\mathbf{z}_{t-1} \not\succeq \hat{\mathbf{z}}_{t+h}|\mathbf{z}_{t-1}]\,, & (6.5)
\end{aligned}
$$

and is denoted Temporal Incomparability Probability (TIP).

Because a high TIP conveys *high predictability* regarding time incomparability, and considering the postponement of the decision criteria preference specification conveyed in preference for flexibility, we assume the DM prefers longer-term performance in this case. In fact, Kumar [142] interpreted flexibility as an adaptive response to unpredictability.

Predictability is also high when TIP is low, in which case the DM can be confident that either (a) performance will improve so as to be preferable in the PD sense; or (b) performance will deteriorate and, thus, such alternative should be avoided. If, on the other hand, the estimated TIP is roughly 1/2, *unpredictability* is maximal, which is indicative that the overlap between the temporal distributions is large. The anticipation rates can then be determined as the complement of the estimated temporal uncertainty:

$$\lambda_{t+h} = \frac{1}{H-1}\left[1 - \mathcal{H}(p_{t,t+h})\right], \tag{6.6}$$

where H is a binary entropy function, with $\mathcal{H}(1/2) = 1$ and $\mathcal{H}(0) = \mathcal{H}(1) = 0$, e.g.,

$$\mathcal{H}(p_{t,t+h}) = -p_{t,t+h}\log p_{t,t+h} - (1 - p_{t,t+h})\log(1 - p_{t,t+h}), \tag{6.7}$$

in which the parameter $p_{t,t+h}$ stands for the temporal non-dominance probability (definition 6.1). The entropy function thus quantifies the temporal uncertainty about the comparability of the current and future estimated objective vectors.

Remark: The factor $\frac{1}{H-1}$ normalizes the anticipatory sum term in AS-MOO Eqs. (5.12) and (5.16) (**chapter 5**) so $\lambda_t + \sum_{h=1}^{H-1}\lambda_{t+h} = 1$.

Assuming minimization of two conflicting objective functions, the TIP can be expressed as:

$$\begin{aligned} p_{t,t+h} \;=\; & \Pr\left[\hat{z}_{t+h,1}|\hat{z}_{t,1} < \hat{z}_{t,1}|z_{t-1,1} \text{ and } \hat{z}_{t+h,2}|\hat{z}_{t,2} > \hat{z}_{t,2}|z_{t-1,2}\right] + \\ & \Pr\left[\hat{z}_{t+h,1}|\hat{z}_{t,1} > \hat{z}_{t,1}|z_{t-1,1} \text{ and } \hat{z}_{t+h,2}|\hat{z}_{t,2} < \hat{z}_{t,2}|z_{t-1,2}\right], \end{aligned}$$

and can be directly computed by querying the Gaussians marginal cumulative distributions.

6.1.3 Preference for Minimal Predictive Error Trajectories

Another proposal implemented in this thesis for enabling anticipatory learning in our AS-MOO solvers is to self-adjust the anticipation rates in $\Lambda_{t:t+H-1}$ according to the sum of KF observed *historical squared residuals* (or innovation terms, see **chapter 2**, section 2.3.4, Eq. (2.29)):

$$\tilde{\mathbf{Z}}_{1:t-1} = \sum_{k=1}^{t-1} ||\tilde{\mathbf{z}}_{t-k}||^2, \tag{6.8}$$

where $\tilde{\mathbf{z}}_{t-k}$ is the difference between the measured objective vector and the predictive mean objective vector at $t-k$. Let $\tilde{\mathbf{Z}}^i_{1:t-1}$ be the sum of historical residuals for the objective vector of the i-th candidate decision in $\mathcal{U}_t$. Let also $\tilde{\mathbf{Z}}^{\min}_{1:t-1}$ and $\tilde{\mathbf{Z}}^{\max}_{1:t-1}$ be the minimum and maximum sum of residuals observed within $\mathcal{U}_t$. Then, the anticipation rates $\lambda_{t+1}, \cdots, \lambda_{t+H-1}$ for a given candidate decision $\mathbf{u}_{t,i} \in \mathcal{U}_t$ can be determined in terms of its normalized sum of historical residuals:

$$\lambda_{t+h} = \frac{1}{H-1}\left(1 - \frac{\tilde{\mathbf{Z}}^i_{1:t-1} - \tilde{\mathbf{Z}}^{\min}_{1:t-1}}{\tilde{\mathbf{Z}}^{\max}_{1:t-1} - \tilde{\mathbf{Z}}^{\min}_{1:t-1}}\right). \tag{6.9}$$

Remark: The normalized sum of residuals falls within $[0, 1]$, and the factor $\frac{1}{H-1}$ assures that the anticipation rates sum to the unity.

The reasoning behind Eq. (6.9) is virtually the same as that behind Eq. (6.6): the higher the normalized sum of KF historical residuals is, the less confident the DM is about the *accuracy* of the KF predictive distribution $\hat{\mathbf{z}}_{t+h}|\mathbf{z}_{t-1}$ and vice-versa. This strategy implements a corrective retrospective assessment providing a means for self-adjusting the DM confidence and willingness to wait for future optimal performance in terms of how *reliable* the past KF predictions were for the fixed decision $\mathbf{u}_{t,i}$.

6.1.4 The Anticipatory Learning Rule in the Objective Space

Regardless of how the anticipation rates are determined, their self-adjustment is intended to automatically control the extent to which the KF future objective vectors predictions are accounted for guiding the search process towards robust candidate decisions. This principle is implemented in the following *anticipatory learning rule*[2]:

$$\hat{\mathbf{z}}_t|\hat{\mathbf{z}}_{t+1:t+H-1} = \mathbf{z}_t + \sum_{h=1}^{H-1} \lambda_{t+h}(\hat{\mathbf{z}}_{t+h}|\mathbf{z}_t - \mathbf{z}_t) \tag{6.10}$$

$$= \left(1 - \sum_{h=1}^{H-1} \lambda_{t+h}\right) \mathbf{z}_t + \sum_{h=1}^{H-1} \lambda_{t+h}\hat{\mathbf{z}}_{t+h}|\mathbf{z}_t \tag{6.11}$$

(for $0 \leq \lambda_{t+h} \leq 1$). The equation states that the adjusted *anticipatory objective distribution* is a convex combination between the current $\mathbf{z}_t$ and the predictive $\hat{\mathbf{z}}_{t+h}|\mathbf{z}_t$ distributions for the objective vector, where $\lambda_t = 1 - \sum_{h=1}^{H-1} \lambda_{t+h}$. The anticipatory distribution $\hat{\mathbf{z}}_t|\hat{\mathbf{z}}_{t+h}$ conveys the important notion that the current estimative for the objective vector depends on the estimated predictive distribution $\hat{\mathbf{z}}_{t+h}$, in the same vain of Rosen's [174] interpretation of anticipatory systems as those whose current states are mediated prediction (see **chapter 1**).

It turns out that, in the KF estimation, the current and future objective vectors are distributed as multivariate Gaussians, and, hence, the anticipatory objective distribution of $\hat{\mathbf{z}}_t$ is also multivariate Gaussian.

Remark: The linear combination of two multivariate Gaussians, $\mathbf{x}$ and $\mathbf{y}$, through matrices $\mathbf{A}$ and $\mathbf{B}$ is given as:

$$\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{y} \sim \mathcal{N}(\mathbf{A}\mathbf{m}_x + \mathbf{B}\mathbf{m}_y, \mathbf{A}\boldsymbol{\Sigma}_x\mathbf{A}^\intercal + \mathbf{B}\boldsymbol{\Sigma}_y\mathbf{B}^\intercal), \tag{6.12}$$

and, thus, the incorporation of prediction into $\boldsymbol{\Sigma}$ is more conservative (e.g. with $T = 2$):

$$\hat{\mathbf{z}}_t|\hat{\mathbf{z}}_{t+1} \sim \mathcal{N}\left(\lambda_t\mathbf{m}_{\mathbf{z}_t} + \lambda_{t+1}\mathbf{m}_{\hat{\mathbf{z}}_{t+1}|\mathbf{z}_t}, \lambda_t^2\boldsymbol{\Sigma}_{\mathbf{z}_t} + \lambda_{t+1}^2\boldsymbol{\Sigma}_{\hat{\mathbf{z}}_{t+1}|\mathbf{z}_t}\right), \tag{6.13}$$

where $\lambda_t = 1 - \lambda_{t+1}$.

[2]The subscript notation $\mathbf{z}_{t+1:t+H-1}$ denotes a sequence $\mathbf{z}_{t+1}, \cdots, \mathbf{z}_{t+H-1}$.

It follows that the proposed anticipatory learning rule (see Fig. 6.1) is expressed in the parameters of the anticipatory objective distribution as (assuming $T = 2$):

$$\mathbf{m}_{\hat{\mathbf{z}}_t|\hat{\mathbf{z}}_{t+1}} = \mathbf{m}_{\mathbf{z}_t} + \lambda_{t+1}\left(\mathbf{m}_{\hat{\mathbf{z}}_{t+1}|\mathbf{z}_t} - \mathbf{m}_{\mathbf{z}_t}\right), \tag{6.14}$$

$$\mathbf{\Sigma}_{\hat{\mathbf{z}}_t|\hat{\mathbf{z}}_{t+1}} = \mathbf{\Sigma}_{\mathbf{z}_t} + \lambda_{t+1}^2\left(\mathbf{\Sigma}_{\hat{\mathbf{z}}_{t+1}|\mathbf{z}_t} - \mathbf{\Sigma}_{\mathbf{z}_t}\right), \tag{6.15}$$

what results in a linear combination between the parameters of the anticipatory objective distribution and those of the current estimated objective distribution. Since a Gaussian distribution requires the covariance matrix $\mathbf{\Sigma}$ to be positive semidefinite, it is worth verifying the following remark:

Remark: Any nonnegative linear combination of positive semidefinite matrices is also positive semidefinite [118].

Let $\mathbf{A}_i \in \mathcal{R}^{n\times n}$ be a series of symmetric positive semidefinite matrices. Then, $\mathbf{x}^\intercal\mathbf{A}_i\mathbf{x} \geq 0,\ \forall\ \mathbf{x} \in \mathcal{R}^n$, for $i = 0, \cdots, N$. It turns out that multiplying $\mathbf{A}_i$ by a nonnegative constant $c \geq 0$ yields

$$c\mathbf{A}_i = c(\mathbf{x}^\intercal\mathbf{A}_i\mathbf{x}) \geq 0, \tag{6.16}$$

and adding any two $\mathbf{A}_i, \mathbf{A}_j$ results in

$$\forall\ \mathbf{x} \in \mathcal{R}^n, \overbrace{(\mathbf{x}^\intercal\mathbf{A}_i\mathbf{x})}^{\geq 0} + \overbrace{(\mathbf{x}^\intercal\mathbf{A}_j\mathbf{x})}^{\geq 0} \geq 0 \Rightarrow \mathbf{x}^\intercal(\mathbf{A}_i + \mathbf{A}_j)\mathbf{x} \geq 0. \tag{6.17}$$

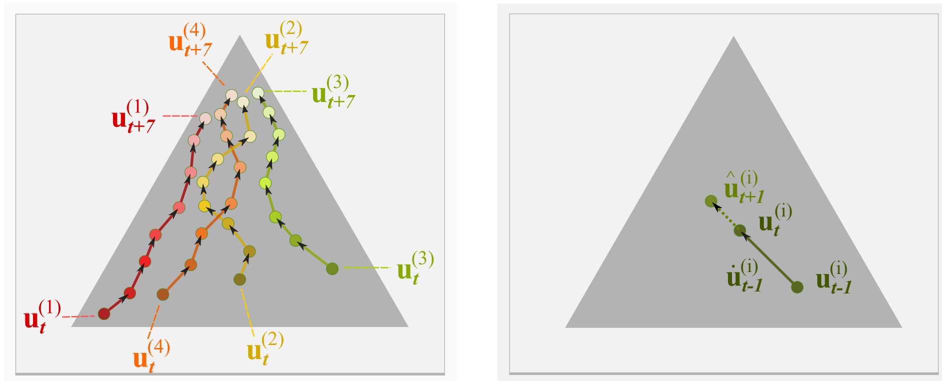
Thus, $c_1\mathbf{A}_1 + \cdots + c_N\mathbf{A}_N \succeq \mathbf{0}$, and hence, if the original covariance matrices in the anticipatory learning rule of Eq. (6.15) are not ill-conditioned, $\mathbf{\Sigma}_{\hat{\mathbf{z}}_t|\hat{\mathbf{z}}_{t+1:t+H-1}}$ is guaranteed to be positive semidefinite. Note that "$\succeq$" defines a partial order in the space of real-valued symmetric matrices, known as the Loewner order [22]:

$$\mathbf{A}_1 \succeq \mathbf{A}_2 \iff \mathbf{A}_1 - \mathbf{A}_2 \quad \text{is positive semidefinite.} \tag{6.18}$$

Figure 6.1 illustrates the effects of applying the proposed OAL rule for combining the distributions of two temporal estimatives for the objective vector of a given candidate trade-off solution in a two-objective space.

6.2 Online Anticipatory Learning in Simplex Decision Spaces

Anticipation in the search space is what enables the solution to a TL-AS-MOO problem. The goal of OAL in the search space is to learn to which directions each Pareto-efficient decision is moving so as to predict what they will be like in future periods (see Fig. 6.2). That is to say, what are the changes required to transform an outdated decision that was Pareto-optimal in a previous decision period into a Pareto-optimal decision for the current (or future) period?



(a) Decision vectors trajectory. (b) Prediction in the search space.

Figure 6.2: Dynamics of mutually non-dominated solution vectors over time.

6.2.1 Unconditional Expectation as a Surrogate for Anticipatory Trade-off Sets

Inspecting the AS-MOO Eq. (5.12) in **chapter 5**, it is clear that the anticipatory sets $\mathcal{U}_{t+h}^{N\star}|\mathbf{u}_t^{\max}$ must be computed in advance. While this would ideally require an exponential $O(N^H)$ number of re-optimizations to identify exaclty the Anticipated Maximal Flexible Choice $\mathbf{u}_t^{\max}$ (see **chapter 5**, Fig. 5.4), we propose using instead the following estimative:

$$\mathbb{E}\left[\hat{\mathcal{U}}_{t+h}^{N\star}\right], \tag{6.19}$$

where $\hat{\mathcal{U}}_{t+h}^{N\star}$ is an unconditional predictive distribution resulting from the Bayesian updating of a prior over the evolving Pareto set. Note from Eq. (6.19) that we alleviate the need to identify $\mathbf{u}_t^{\max}$ by assuming that this information is implicitly learned when tracking the trajectory of the Pareto Estimated Maximal Flexible Choices set over time (Eq. (5.15), **chapter 5**). Although an unconditional predictive distribution provides a less accurate estimative, the computational cost to obtain such estimatives is negligible when compared to the cost of recursively re-optimizing. The unconditional expected value is then used as a *surrogate* to the conditional anticipatory Pareto Sets for approximately solving a TL-AS-MOO problem.

6.2.2 Correspondence Mapping for Multiple Decision Vectors Tracking

The proposed tracking methods for estimating the finite Pareto candidate solutions set approximations trajectories over time conveys the same path described in Eq. (6.1):

$$\hat{\mathcal{U}}_t^{N\star}|\mathcal{U}_{t-1}^{N\star} \rightarrow \hat{\mathcal{U}}_{t+1}^{N\star}|\hat{\mathcal{U}}_t^{N\star} \rightarrow \cdots \rightarrow \hat{\mathcal{U}}_{t+h}^{N\star}|\hat{\mathcal{U}}_{t+h-1}^{N\star} \equiv \hat{\mathcal{U}}_{t+h}^{N\star}|\mathcal{U}_{t-1}^{N\star}. \tag{6.20}$$

Nevertheless, the proposed tracking method is applied *independently* for each non-inferior solution in $\hat{\mathcal{U}}_{t+h}^{N\star}$. In the case of online anticipatory learning in the objective space, discussed in section 6.1, independently tracking and predicting the evolving objective vectors is not an issue because those tasks are carried out w.r.t. *fixed* candidate solutions. But in the search space case, that begs the following question: how to make the correspondence between each EMFC in subsequent optimization environments, i.e., which decision in $\hat{\mathcal{U}}_{t+h-1}^{N\star}$ corresponds to which decision in $\hat{\mathcal{U}}_{t+h}^{N\star}$?

In order to handle the correspondence mapping of individual solutions evolving over time, we first sort all candidate trade-off solutions at any given decision state w.r.t. one of the objective functions. Hence, $\mathbf{u}_{t+h}^{(i)}$ represents the candidate solution which is ranked in the i-th position. For example, let $\hat{\mathcal{U}}_t^{N\star} = \{\hat{\mathbf{u}}_{t,1}, \hat{\mathbf{u}}_{t,2}, \hat{\mathbf{u}}_{t,3}\}$ be the set of candidate solutions[3] at time t and $\hat{\mathbf{m}}_{\mathbf{z}_{t,1}} = (1\ 1)^\intercal, \hat{\mathbf{m}}_{\mathbf{z}_{t,2}} = (0\ 2)^\intercal, \hat{\mathbf{m}}_{\mathbf{z}_{t,3}} = (2\ 0)^\intercal$ the corresponding mean objective vectors. When sorting in increasing order the elements in $\hat{\mathcal{U}}_t^{N\star}$ by their first objective values, we obtain the order

$$\hat{\mathbf{u}}_t^{(1)} = \hat{\mathbf{u}}_{t,2}, \hat{\mathbf{u}}_t^{(2)} = \hat{\mathbf{u}}_{t,1}, \hat{\mathbf{u}}_t^{(3)} = \hat{\mathbf{u}}_{t,3},$$

which indicates that the best candidate solution ($\hat{\mathbf{u}}_t^{(1)}$) w.r.t. the first optimization criterion is $\hat{\mathbf{u}}_{t,2}$ and so forth. Figure 6.2(a) depicts the independent trajectory of four ordered decision vectors over time in a hypothetical scenario when the optimization is performed in the 2-simplex.

After sorting the candidate solutions in the *predicted* approximated Pareto set $\hat{\mathcal{U}}_{t+1}^{N\star}$, we are then able to track the approximated *ordered trade-off solutions* evolution over time. For instance, the resulting prediction chain for the i-th ranked trade-off solution is:

$$\hat{\mathbf{u}}_t^{(i)} | \mathbf{u}_{t-1}^{(i)} \to \hat{\mathbf{u}}_{t+1}^{(i)} | \hat{\mathbf{u}}_t^{(i)} \to \cdots \to \hat{\mathbf{u}}_{t+h}^{(i)} | \hat{\mathbf{u}}_{t+h-1}^{(i)} \equiv \hat{\mathbf{u}}_{t+h}^{(i)} | \mathbf{u}_{t-1}^{(i)}. \tag{6.21}$$

6.2.3 A Sliding Window Dirichlet Dynamical Model

As discussed in **chapter 2**, section 2.3.5, the Dirichlet Distribution (DD) [160] is a natural choice for modeling random variables whose sample spaces are defined over the $(d-1)$-simplex. In this thesis, despite the decision variables being deterministic (i.e., we are not interested in modeling a stochastic DM), we propose representing the N candidate decision vectors in S^{d-1} as the *mean vectors* (Eq. 2.36) of N independent DDs evolving over time.

We also recall from **chapter 2** that the variances (Eq. 2.37) in the DD depend on the scale parameter (a.k.a. concentration, α_C): $\alpha_C = \sum_{l=1}^{d} \alpha_l$. The higher α_C is, the lower the variances of each decision variable are, and vice-versa. Furthermore, when discussing the application of the DD as a prior for the multinomial distribution, we concluded that the posterior of the multinomial proportions, conditioned on evidence on the counts for each category, is also a DD with the concentration parameter being updated by adding the vector of observed counts, $\mathbf{c}$.

Here we provide a different scenario for the DD posterior other than to accumulate counts. We use $\boldsymbol{\alpha}$ to encode the *uncertainty* about the evolving decision variables in the $(d-1)$-simplex. Depending on the optimization environment described by the incoming data, we assume that the uncertainty can either decrease or increase over time and, thus, the decrease in total variance of the DD upon arrival of new data is not necessarily strictly monotone. We then propose updating the posterior distributions in this scenario is by removing old observations from the Dirichlet process history. We thus define a *forgetting factor* K so as to subtract observations collected

[3]The hat notation in this case indicates that $\hat{\mathcal{U}}_t^{N\star}$ is an estimative of the optimal anticipatory Pareto Set, while the true optimal finite Pareto Set is still denoted as $\mathcal{U}_t^{N\star}$.

from decision periods *previous* to $t-K$. Let the posterior uncertainty about the candidate trade-off solutions[4] be represented as

$$\mathbf{u}_0^{(i)} \sim \mathcal{D}(\boldsymbol{\alpha}_0^{(i)}), \mathbf{u}_1^{(i)}|\mathbf{u}_0^{(i)} \sim \mathcal{D}(\boldsymbol{\alpha}_1^{(i)}|\mathbf{u}_0^{(i)}), \cdots, \mathbf{u}_t^{(i)}|\mathbf{u}_{t-1}^{(i)} \sim \mathcal{D}(\boldsymbol{\alpha}_t^{(i)}|\mathbf{u}_{t-1}^{(i)}). \tag{6.22}$$

Let $\boldsymbol{\alpha}_t^{(i)}|\mathbf{u}_{t-1}^{(i)}$ denote the update of concentration $\boldsymbol{\alpha}_t^{(i)}$ given that evidence on the proportions decision vector $\mathbf{u}_{t-1}^{(i)}$ was collected. Assuming a forgetting factor (window size) of $K=2$, and an even-handed concentration parameter at $t=0$, the envisioned strategy yields the following recursive parameter update of the DD posterior over time:

$$\begin{aligned}
\boldsymbol{\alpha}_0^{(i)} &= s(\frac{1}{d} \ \cdots \ \frac{1}{d})^\intercal, \\
\boldsymbol{\alpha}_1^{(i)}|\mathbf{u}_0^{(i)} &= \boldsymbol{\alpha}_0^{(i)} + s\mathbf{u}_0^{(i)} = s(\frac{1}{d} + u_{1,0}^{(i)} \ \cdots \ \frac{1}{d} + u_{d,0}^{(i)})^\intercal, \\
\boldsymbol{\alpha}_2^{(i)}|\mathbf{u}_1^{(i)} &= \boldsymbol{\alpha}_1^{(i)}|\mathbf{u}_0^{(i)} + s\mathbf{u}_1^{(i)} - \boldsymbol{\alpha}_0^{(i)} = s(u_{1,0}^{(i)} + u_{1,1}^{(i)} \ \cdots \ u_{d,0}^{(i)} + u_{d,1}^{(i)})^\intercal \\
\boldsymbol{\alpha}_3^{(i)}|\mathbf{u}_2^{(i)} &= \boldsymbol{\alpha}_2^{(i)}|\mathbf{u}_1^{(i)} + s\mathbf{u}_2^{(i)} - s\mathbf{u}_0^{(i)} = s(u_{1,1}^{(i)} + u_{1,2}^{(i)} \ \cdots \ u_{d,1}^{(i)} + u_{d,2}^{(i)})^\intercal \\
&\vdots \\
\boldsymbol{\alpha}_t^{(i)}|\mathbf{u}_{t-1}^{(i)} &= \boldsymbol{\alpha}_{t-1}^{(i)}|\mathbf{u}_{t-2}^{(i)} + s\mathbf{u}_{t-1}^{(i)} - s\mathbf{u}_{t-3}^{(i)} = s(u_{1,t-2}^{(i)} + u_{1,t-1}^{(i)} \ \cdots \ u_{d,t-2}^{(i)} + u_{d,t-1}^{(i)})^\intercal.
\end{aligned} \tag{6.23}$$

The general case for $K \in \{1, \cdots, t-1\}$ is then

$$\begin{aligned}
&\mathbf{u}_t^{(i)}|\mathbf{u}_{t-1}^{(i)} \sim \mathcal{D}(\boldsymbol{\alpha}_t^{(i)}|\mathbf{u}_{t-1}^{(i)}), \ \text{where} \\
&\boldsymbol{\alpha}_t^{(i)}|\mathbf{u}_{t-1}^{(i)} = \begin{cases} \boldsymbol{\alpha}_{t-1}^{(i)}|\mathbf{u}_{t-2}^{(i)} + s\mathbf{u}_{t-1}^{(i)}, & \text{if } t < K, \\ \boldsymbol{\alpha}_{t-1}^{(i)}|\mathbf{u}_{t-2}^{(i)} + s\mathbf{u}_{t-1}^{(i)} - \boldsymbol{\alpha}_0^{(i)}, & \text{if } t = K, \\ \boldsymbol{\alpha}_{t-1}^{(i)}|\mathbf{u}_{t-2}^{(i)} + s\mathbf{u}_{t-1}^{(i)} - s\mathbf{u}_{t-K-1}^{(i)}, & \text{otherwise.} \end{cases}
\end{aligned} \tag{6.24}$$

Note that, for $t > K$,

$$\boldsymbol{\alpha}_{t-1}^{(i)}|\mathbf{u}_{t-2}^{(i)} + s\mathbf{u}_{t-1}^{(i)} - s\mathbf{u}_{t-K-1}^{(i)} = s\left(\sum_{k=1}^{K} u_{1,t-k}^{(i)} \ \cdots \ \sum_{k=1}^{K} u_{d,t-k}^{(i)}\right)^\intercal. \tag{6.25}$$

We denote the term $\dot{\boldsymbol{\alpha}}_t^{(i)} = s\mathbf{u}_{t-1}^{(i)} - \boldsymbol{\alpha}_{t-K}^{(i)}|\mathbf{u}_{t-K-1}^{(i)}$ as the *velocity* (or rate of change) with which the decision vectors distributions are moving throughout the $(d-1)$-simplex (depicted as the arrow vectors in Fig. 6.2 (b)) per each K units of time. The velocities thus encode the underlying dynamical model (similarly to the KF velocities, see section 6.1.1). Also, for $K=1$, note how the distribution of $\mathbf{u}_t^{(i)}$ depends only on the previous observation, $s\mathbf{u}_{t-1}^{(i)}$. The scale parameter s controls the dispersion of the DD process over time and must be set in advance. With this strategy, the concentration of $\mathbf{u}_t^{(i)}|\mathbf{u}_{t-1}^{(i)}$ goes through a K-fold increase from $t=0$ to $t=K$ and from then on remains constant (for $t > K$), as historical data is removed to give

[4]Differently from the objective space case, we emphasize the decision vectors are actually deterministic.

place to the most recent observations (recalling $\mathbf{u}_t^{(i)} \in S^{d-1}$):

$$\begin{aligned}
\alpha_{C,0}^{(i)} &= \sum_{l=1}^{d} \frac{s}{d} = s, \\
\alpha_{C,1}^{(i)} &= \sum_{l=1}^{d} \frac{s}{d} + s\sum_{l=1}^{d} u_{l,0}^{(i)} = 2s, \\
\alpha_{C,2}^{(i)} &= \sum_{l=1}^{d} \frac{s}{d} + s\sum_{l=1}^{d} u_{l,0}^{(i)} + s\sum_{l=1}^{d} u_{l,1}^{(i)} = 3s, \\
&\vdots \\
\alpha_{C,K}^{(i)} &= \underbrace{s\sum_{l=1}^{d} u_{l,0}^{(i)} + \cdots + s\sum_{l=1}^{d} u_{l,K}^{(i)}}_{K \text{ terms}} = Ks, \\
&\vdots \\
\alpha_{C,t}^{(i)} &= \underbrace{s\sum_{l=1}^{d} u_{l,t-K}^{(i)} + \cdots + s\sum_{l=1}^{d} u_{l,t}^{(i)}}_{K \text{ terms}} = Ks.
\end{aligned} \tag{6.26}$$

Because of the increase in concentration, it is then evident from Eq. (6.26) that the proposed posterior update model of Eq. (6.24) assumes the uncertainty to decrease after K decision periods and to remain constant for $t \geq K$.

Remark: It follows from Eq. (6.24) that the expected value of $\mathbf{u}_t^{(i)}|\mathbf{u}_{t-1}^{(i)}$ is

$$\begin{aligned}
\mathbb{E}\left[\mathbf{u}_t^{(i)}|\mathbf{u}_{t-1}^{(i)}\right] &= \left(\mathbb{E}\left[u_{1,t}^{(i)}|u_{1,t-1}^{(i)}\right] \;\cdots\; \mathbb{E}\left[u_{d,t}^{(i)}|u_{d,t-1}^{(i)}\right]\right)^{\intercal}, \text{ where} \\
\mathbb{E}\left[u_{j,t}^{(i)}|u_{j,t-1}^{(i)}\right] &= m_{u_{j,t}^{(i)}}|u_{j,t-1}^{(i)} = \frac{\sum_{k=1}^{K} u_{j,t-k}^{(i)}}{s\sum_{l=1}^{d}\sum_{k=1}^{K} u_{l,t-k}^{(i)}}.
\end{aligned} \tag{6.27}$$

6.2.4 Predicting DD Mean Decision Vectors

It is then of interest to the proposed anticipatory learning methods the estimation of the predictive mean vectors $\hat{\mathbf{m}}_{\mathbf{u}_{t+h}^{(i)}}|\mathbf{m}_{\mathbf{u}_{t-1}^{(i)}}$, for $h = 1, \cdots, H$. As shown in **chapter 2**, section 2.3.5, Bertuccelli and How [30] proposed a Maximum A Posteriori (MAP) recursive point estimation to track the evolving DD means and variances over time (see Eqs. (2.42) and (2.43)), but the dynamical model assumed the constant predictor $\hat{\mathbf{m}}_{\mathbf{u}_{t+h}^{(i)}}|\mathbf{m}_{\mathbf{u}_{t-1}^{(i)}} = \mathbf{m}_{\mathbf{u}_{t-1}^{(i)}}$. Without loss of generality, sticking to the simpler case of $K = 1$, defining the velocities for the DD mean vectors as $\dot{\mathbf{m}}_{\mathbf{u}_t^{(i)}} = \mathbf{m}_{\mathbf{u}_t^{(i)}} - \mathbf{m}_{\mathbf{u}_{t-1}^{(i)}}$, and assuming a one step ahead prediction ($h = 1$), we propose a linear

Input: The prediction horizon H
Input: The concentration scaling parameter s
Input: A current trade-off solution $\mathbf{u}_t$ not necessarily in $\mathcal{U}_t^N$
Input: A historical sequence of trade-off solutions $\mathcal{U}_{t-K:t-1} = \{\mathbf{u}_{t-K}, \cdots, \mathbf{u}_{t-1}\}$
procedure DDPREDICTION(H, s, $\mathbf{u}_t$, $\mathcal{U}_{t-K:t-1}$)
  for $k = K$ **to** 0 **do** ▷ Historical DD MAP mean tracking
    Estimate the belief coefficient υ_{t-k} using Eq. (6.30)
    Predict $\hat{\mathbf{m}}_{\mathbf{u}_{t-k}}$ using Eq. (6.31)
    Update $\hat{\mathbf{m}}_{\mathbf{u}_{t-k}}|\mathbf{m}_{\mathbf{u}_{t-k}}$ using Eq. (6.33)
  end for
  for $h = 1$ **to** H **do** ▷ H steps ahead prediction
    Estimate the belief coefficient υ_{t+h} using Eq. (6.30)
    Predict $\hat{\mathbf{m}}_{\mathbf{u}_{t+h}}$ using Eq. (6.31)
  end for
  return $\hat{\mathbf{u}}_{t+H}|\mathbf{u}_{t-1} \sim \mathcal{D}(s\sum_{k=1}^{K} \hat{\mathbf{m}}_{\mathbf{u}_{t+H-k}})$ (Eq. (6.24))
end procedure

Pseudocode 6: Dirichlet MAP Tracking and Prediction

dynamical model that assumes the velocity to remain constant at $t+1$, i.e., $\dot{\mathbf{m}}_{\mathbf{u}_t^{(i)}} = \dot{\mathbf{m}}_{\mathbf{u}_{t-1}^{(i)}}$:

$$\begin{aligned}
\hat{\mathbf{m}}_{\mathbf{u}_{t+1}^{(i)}} &= \mathbf{m}_{\mathbf{u}_t^{(i)}} + \hat{\dot{\mathbf{m}}}_{\mathbf{u}_t^{(i)}} = \mathbf{m}_{\mathbf{u}_t^{(i)}} + \dot{\mathbf{m}}_{\mathbf{u}_{t-1}^{(i)}} \\
&= \mathbf{m}_{\mathbf{u}_t^{(i)}} + \mathbf{m}_{\mathbf{u}_t^{(i)}} - \mathbf{m}_{\mathbf{u}_{t-1}^{(i)}} = 2\mathbf{m}_{\mathbf{u}_t^{(i)}} - \mathbf{m}_{\mathbf{u}_{t-1}^{(i)}} \\
&= 2\mathbf{m}_{\mathbf{u}_t^{(i)}} + (\mathbf{m}_{\mathbf{u}_{t-1}^{(i)}} - 2\mathbf{m}_{\mathbf{u}_{t-1}^{(i)}}) = \mathbf{m}_{\mathbf{u}_{t-1}^{(i)}} + 2\mathbf{m}_{\mathbf{u}_t^{(i)}} - 2\mathbf{m}_{\mathbf{u}_{t-1}^{(i)}} \\
&= \mathbf{m}_{\mathbf{u}_{t-1}^{(i)}} + 2(\mathbf{m}_{\mathbf{u}_t^{(i)}} - \mathbf{m}_{\mathbf{u}_{t-1}^{(i)}}) \\
&= \mathbf{m}_{\mathbf{u}_{t-1}^{(i)}} + 2\dot{\mathbf{m}}_{\mathbf{u}_{t-1}^{(i)}}.
\end{aligned} \tag{6.28}$$

However, extrapolating a linear trend is prone to error if the trend changes abruptly and, thus, we define $\upsilon_{t+1} \in \left[\frac{1}{2}, 1\right]$ as the *belief coefficient* which controls the *confidence* that the predictive decision vector distribution will keep its motion according to the same velocity as measured at the preceding decision period. Hence, the confidence is maximal for $\upsilon_{t+1} = 1$, whereas for $\upsilon_{t+1} = \frac{1}{2}$, the prediction will be conservative and just assume the decision vector distribution will not change between subsequent periods, being equivalent to Bertuccelli and How's [30] static model (see Fig. 6.2(b) for an example with $h = 1$), i.e., $\hat{\mathbf{m}}_{\mathbf{u}_{t+1}^{(i)}} = \mathbf{m}_{\mathbf{u}_t^{(i)}}$:

$$\hat{\mathbf{m}}_{\mathbf{u}_{t+1}^{(i)}} = \mathbf{m}_{\mathbf{u}_{t-1}^{(i)}} + \upsilon_{t+1} 2\dot{\mathbf{m}}_{\mathbf{u}_{t-1}^{(i)}}. \tag{6.29}$$

We propose self-adjusting υ_{t+1} for each candidate solution as the complement of the binary entropy function measured over the TIP between $\mathbf{u}_t^{(i)}$ and $\mathbf{u}_{t-1}^{(i)}$, for which case we make use of the KF estimation $\hat{\mathbf{z}}_t^{(i)}|\mathbf{z}_{t-1}^{(i)}$ and of $\mathbf{z}_{t-1}^{(i)}$, where $\mathbf{z}_{t-1}^{(i)} = \mathbf{f}(\mathbf{u}_{t-1}^{(i)})$:

$$\upsilon_{t+1} = 1 - \frac{1}{2}\mathcal{H}(p_{t-1,t}), \tag{6.30}$$

where $p_{t-1,t}$ is the estimated TIP between $\hat{\mathbf{z}}_t^{(i)}|\mathbf{z}_{t-1}^{(i)}$ and $\mathbf{z}_{t-1}^{(i)}$. With this strategy, we propagate the current uncertainty about whether the current i-th ranked candidate decision vector $\mathbf{u}_t^{(i)}$ at t is mutually non-dominated with respect to the preceding i-th decision vector at $t-1$, $\mathbf{u}_{t-1}^{(i)}$.

Input: The anticipation horizon H
Input: The concentration scaling parameter s
Input: A candidate $\hat{\mathbf{u}}_t \in \hat{\mathcal{U}}_t$ and the $\hat{\mathcal{U}}_t$ set
Input: The sets $\hat{\mathcal{U}}^N_{t-K:t}$ and $\mathcal{X}_{t-K:t-1} = \{\mathbf{x}_{t-K}, \cdots, \mathbf{x}_{t-1}\}$
Output: The estimated anticipatory distribution for $\hat{\mathbf{u}}_t$

procedure AnticipatoryDistribution
  Compute the rank i of $\hat{\mathbf{u}}_t$ w.r.t. $\mathcal{U}_t$, yielding $\hat{\mathbf{u}}_t^{(i)}$
  Set $\hat{\mathcal{Z}}^{(i)}_{t:t+H-1} \leftarrow \emptyset$
  if (TLF Regime) **then**
    for $h = 0$ **to** $H-1$ **do**
      $\hat{\mathbf{z}}^{(i)}_{t+h} \leftarrow$ KFPrediction(h, $\mathbf{u}_t^{(i)}$, $\mathcal{X}_{t-K:t-1}$)
      $\hat{\mathcal{Z}}^{(i)}_{t:t+H-1} \leftarrow \hat{\mathcal{Z}}^{(i)}_{t:t+H-1} \cup \{\hat{\mathbf{z}}^{(i)}_{t+h}\}$
    end for
  else if (TL Regime) **then**
    Retrieve rank i decisions in $\hat{\mathcal{U}}^{(i)}_{t-K:t} \subset \hat{\mathcal{U}}_{t-K:t}$
    for $h = 0$ **to** $H-1$ **do**
      Predict $\hat{\mathbf{u}}^{(i)}_{t+h}$ using $\hat{\mathcal{U}}^{(i)}_{t-K:t}$ with Eqs. 6.31–6.33
      Predict $\hat{\mathbf{z}}^{(i)}_{t+h}$ using $\hat{\mathcal{Z}}^{(i)}_{t-K:t}$ with the KF (Fig. 6.1)
      $\hat{\mathbf{z}}^{(i)}_{t+h} \leftarrow \hat{\mathbf{z}}^{(i)}_{t+h} + \mathbf{h}(\mathbf{m}_{\hat{\mathbf{u}}^{(i)}_{t+h}}, \mathbf{m}_{\hat{\mathbf{u}}^{(i)}_{t+h-1}})$
      $\hat{\mathcal{Z}}^{(i)}_{t:t+H-1} \leftarrow \hat{\mathcal{Z}}^{(i)}_{t:t+H-1} \cup \{\hat{\mathbf{z}}^{(i)}_{t+h}\}$
    end for
  end if
  Apply OAL over $\hat{\mathcal{Z}}^{(i)}_{t:t+H-1}$ using Eq. (6.10)
  return The anticipatory distribution $\hat{\mathbf{z}}_t^{(i)} | \hat{\mathbf{z}}^{(i)}_{t+1:t+H-1}$
end procedure

Pseudocode 7: Anticipatory Distribution Estimation

The prediction of $\hat{\mathbf{m}}_{\mathbf{u}^{(i)\star}_{t+1}}$ from all historical estimated non-inferior mean candidate solutions $\hat{\mathbf{m}}_{\mathbf{u}^{(i)\star}_{t-1}} \in \mathbb{E}\left[\hat{\mathcal{U}}^{N\star}_{t-1}\right]$ and $\hat{\mathbf{m}}_{\mathbf{u}^{(i)\star}_{t-2}} \in \mathbb{E}\left[\hat{\mathcal{U}}^{N\star}_{t-2}\right]$ (for $i = 1, \cdots, N$) is thus performed as:

$$\hat{\mathbf{m}}_{\mathbf{u}^{(i)\star}_{t+1}} = \hat{\mathbf{m}}_{\mathbf{u}^{(i)\star}_{t-1}} + \upsilon_{t+1} 2\dot{\mathbf{m}}_{\mathbf{u}^{(i)\star}_{t-1}}, \tag{6.31}$$

where $\dot{\mathbf{m}}_{\mathbf{u}^{(i)\star}_{t-1}} = \hat{\mathbf{m}}_{\mathbf{u}^{(i)\star}_{t-1}} - \mathbf{m}_{\mathbf{u}^{(i)\star}_{t-2}}$.

Therefore, our surrogate estimate for $\mathcal{U}^{N\star}_{t+1}$ is obtained as:

$$\mathbb{E}\left[\hat{\mathcal{U}}^{N\star}_{t+1}\right] \approx \left\{\hat{\mathbf{m}}_{\mathbf{u}^{(1)\star}_{t+1}}, \cdots, \hat{\mathbf{m}}_{\mathbf{u}^{(N)\star}_{t+1}}\right\}, \tag{6.32}$$

6.2.5 Maximum A Posteriori Correction for DD Mean Decision Vectors

The MAP recursive update step for correcting $\hat{m}_{u^{(i)\star}_{l,t+1}}$ upon the observation of the d candidate mean decision variables $m_{u^{(i)}_{l,t+1}}$ at $t+1$ (for $l = 1, \cdots, d$) is then [30] (see Eq. (2.42)):

$$\hat{m}_{u^{(i)\star}_{l,t+1}} | u^{(i)}_{l,t+1} = \hat{m}_{u^{(i)\star}_{l,t+1}} + \frac{\mathrm{Var}\left[\hat{u}^{(i)\star}_{l,t+1}\right]}{\hat{m}_{u^{(i)\star}_{l,t+1}}(1 - \hat{m}_{u^{(i)\star}_{l,t+1}})}(m_{u^{(i)}_{l,t+1}} - \hat{m}_{u^{(i)\star}_{l,t+1}}), \tag{6.33}$$

Note from Eq. (6.33) that, whenever each component $u^{(i)}_{l,t+1}$ of a decision vector is observed, i.e., when the optimal flexible Pareto-optimal decision vector of rank i is finally observed (computed) at time $t+1$, the prediction made at time t, $\hat{m}_{u^{(i)\star}_{l,t+1}}$, is *corrected* towards the direction of the actual observed DD mean vector, $m_{u^{(i)}_{l,t+1}}$. Such update is weighted by a *correction factor* which is estimated from the underlying uncertainty of the DD model regarding the prediction $\hat{m}_{u^{(i)\star}_{l,t+1}}$. Such factor is expressed as the ratio between the variance of the predicted DD distribution, $\mathrm{Var}\left[\hat{u}^{(i)\star}_{l,t+1}\right]$, and the product between the predicted DD components and its complement, $\hat{m}_{u^{(i)\star}_{l,t+1}}(1 - \hat{m}_{u^{(i)\star}_{l,t+1}})$.

The Pseudocode 6 describes the steps to track and predict the trajectory of a given proportion vector given the last K observations. In line 5, note that the MAP update $\hat{\mathbf{m}}_{\mathbf{u}_{t-k}} | \mathbf{m}_{\mathbf{u}_{t-k}}$ requires the computation of the variance of the corresponding DD, which depends on the concentration scaling input parameter s (see Eq. (6.33)).

The Pseudocode 7 describes the full procedure for computing the *anticipatory distributions* $\hat{\mathbf{z}}^{(i)}_t | \hat{\mathbf{z}}^{(i)}_{t+1:t+H-1}$ for each $\mathbf{u}^{(i)}_t \in \mathcal{U}^N_t$, utilizing tracking and prediction in both the objective and search spaces with the KF (Pseudocode 5), for both TLF and TL regimes, and the Dirichlet procedure, for the TL regime (Pseudocode 6). Tracking and prediction begin in the search space, where the trajectories of each candidate Anticipated Maximal Flexible Choice are predicted according to the previously implemented choices, after which the KF is applied for the resulting objective vectors, which are then adjusted with the predicted cost component $\mathbf{h}$. The anticipation rates $\lambda_{t+h}, \cdots, \lambda_{t+H-1}$ are then determined and used to combine the resulting predictive random objective vectors to finally yield the anticipatory distributions.

6.3 Computing the Expected Anticipatory Hypervolume Contributions

Given a partition set (class) of objective distributions, $\mathcal{C}$, whose mean vectors are mutually (Pareto) non-dominated, we recall Eq. (3.23) at **chapter 3** which defined the S-Metric contribution of the objective vector associated with a candidate decision to its non-dominance class ($\mathbf{z} \in \mathcal{C}$) as:

$$\Delta_{\mathcal{S}}(\mathbf{z}, \mathcal{C}) = \mathcal{S}(\mathcal{C}) - \mathcal{S}(\mathcal{C} \setminus \{\mathbf{z}\}). \tag{6.34}$$

The expected value of $\Delta_{\mathcal{S}}$ for a given anticipatory random objective vector $\hat{\mathbf{z}}^{(i)}_t | \hat{\mathbf{z}}^{(i)}_{t+h}$ (Eq. 6.10) is essential for solving AS-MOO problems because it provides valuable information for heuristically

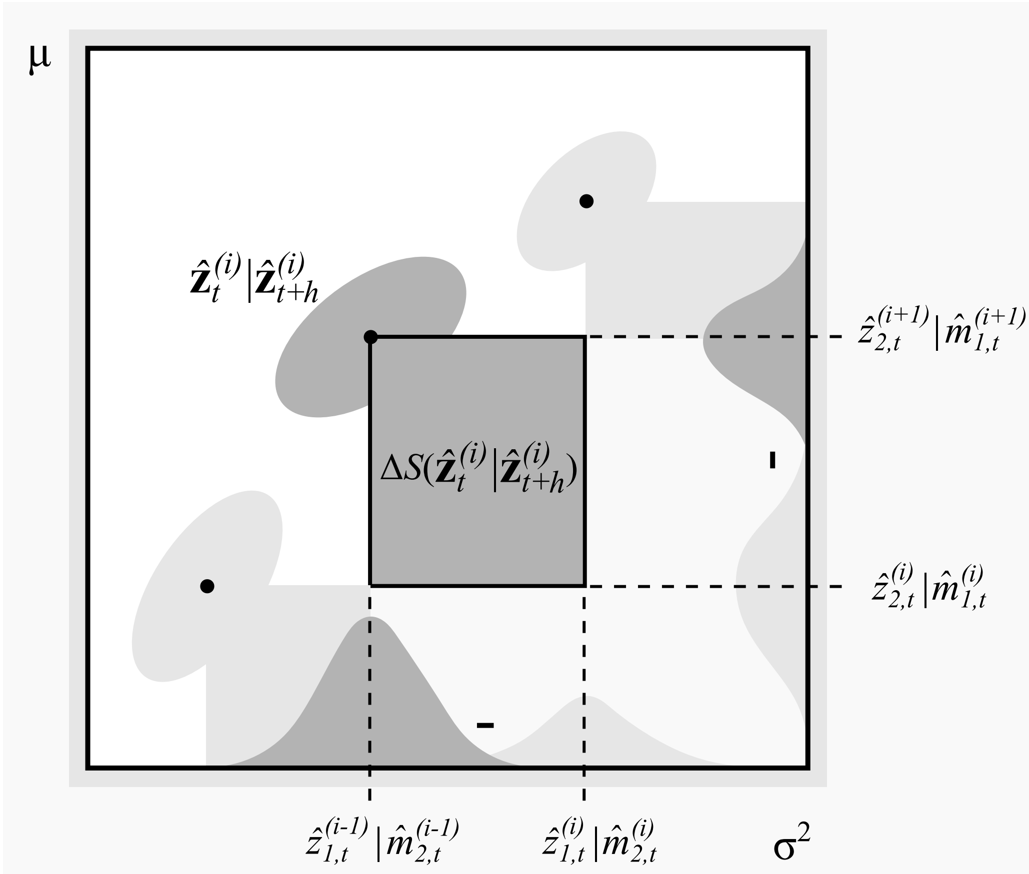


Figure 6.3: Computing $\mathbb{E}[\Delta_{\mathcal{S}}]$ in the objective space by conditioning the decision criteria at their mean values, assuming max return and min risk.

improving the expected overall S-Metric value of a given candidate trade-off set (see section 6.5). Without loss of generality, let $\mathbf{z}_t^{(i)} \in \mathcal{C}$ be a non-extremal[5] bivariate random objective vector ($|\mathcal{C}| \geq 3$). Then, assuming minimization of g_1 (risk) and maximization of g_2 (return), and given two *neighboring* objective vectors w.r.t. their mean objective values ($\mathbf{z}_t^{(i-1)}, \mathbf{z}_t^{(i+1)} \in \mathcal{C}$), the $\mathcal{S}$ contribution of $\mathbf{z}_t^{(i)}$ to $\mathcal{C}$ is expressed as

$$\Delta_{\mathcal{S}}(\mathbf{z}_t^{(i)}) = \left[z_{1,t}^{(i)} - z_{1,t}^{(i-1)}\right]\left[z_{2,t}^{(i+1)} - z_{2,t}^{(i-1)}\right]. \tag{6.35}$$

Remark: The conditional distribution of a multivariate Gaussian is also Gaussian and, thus, δ_1, δ_2 are Gaussian random variables.

Because the KF estimation assumes g_1 and g_2 to be jointly Gaussian, it follows that the marginal distributions for each $\hat{z}_{j,t}^{(i)}$ ($j \in \{1, 2\}$) are univariate Gaussian[6]. In Eq. (6.35), however, we make

[5]The extrema objective vectors within a given non-dominance class $\mathcal{C}$ are those containing the minimum mean value w.r.t. one of the objective functions among the other vectors in $\mathcal{C}$. The $\mathcal{S}$-Metric contributions of such extrema vectors are computed in terms of a reference point $\mathbf{z}^{\text{ref}}$, which bounds the preferred feasible region in the performance space.

[6]In order to simplify the exposition, from this point on, we write $\hat{\mathbf{z}}_t^{(i)} \equiv \hat{\mathbf{z}}_t^{(i)}|\hat{\mathbf{z}}_{t+h}^{(i)}$ and $\hat{z}_{j,t}^{(i)} \equiv \hat{z}_{j,t}^{(i)}|\hat{z}_{j,t+h}^{(i)}$.

use of conditional distributions instead of the marginals, which can also shown to be univariate Gaussian. The reason for using the conditional distributions in the S-Metric contribution estimation is due to our Stochastic Pareto Frontier (SPF) definition, that is given in terms of the expected value of the random objective vectors. Therefore, when partitioning the candidate decisions into non-dominance classes in the proposed AS-MOO solver (see section 6.5), the Pareto Dominance is applied over the *mean objective vectors*. It is mentioning that the conditional distributions would be *truncated* Gaussians if the conditioning were done so as to require $\Delta_{\mathcal{S}} > 0$, but conditioning on the means has the advantage of preserving the original partitioning obtained by the non-dominated sorting procedure over the mean vectors.

In simple terms, we compute $\Delta_{\mathcal{S}}$ for an anticipatory random objective vector $\hat{\mathbf{z}}_t^{(i)}$ using the conditional distribution of the j-th objective value given that the k-th objective (for $k, j \in \{1, 2\}$ and $k \neq j$) is fixed at its mean value (m_k) (see Fig. 6.3):

$$\Delta_{\mathcal{S}}(\hat{\mathbf{z}}_t^{(i)}) = (\hat{z}_{1,t}^{(i)}|\hat{m}_{2,t}^{(i)} - \hat{z}_{1,t}^{(i-1)}|\hat{m}_{2,t}^{(i-1)})(\hat{z}_{2,t}^{(i+1)}|\hat{m}_{1,t}^{(i+1)} - \hat{z}_{2,t}^{(i)}|\hat{m}_{1,t}^{(i)}). \tag{6.36}$$

The product between two Gaussians is a product-normal distribution, though. Nevertheless, for the sake of approximately solving AS-MOO problems, we are only interested in estimating the expected value of $\Delta_{\mathcal{S}}(\hat{\mathbf{z}}_t^{(i)})$. In order to do so, we describe in the following a few useful facts.

Remark: The *conditional expected value* for a bivariate Gaussian, $\mathbf{x} = (x_1\ x_2)^\intercal \sim \mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$, is expressed as

$$\mathbb{E}\left[x_1|(x_2 = c)\right] = m_1 + \frac{\sigma_1}{\sigma_2}\rho(c - m_2), \tag{6.37}$$

where ρ is the correlation coefficient between x_1 and x_2.

If the conditioning is done over the mean value of the second variable (i.e., $c = m_2$), it is then easy to verify that:

$$\mathbb{E}\left[x_1|(x_2 = m_2)\right] = m_1. \tag{6.38}$$

Remark: The expected value of the product of any two random variables with arbitrary probability densities can be expressed as the sum of the product of the individual expectancies and the covariance:

$$\begin{aligned}
\mathrm{Cov}(x_1, x_2) &= \mathbb{E}\left[(x_1 - \mathbb{E}\left[x_1\right])(x_2 - \mathbb{E}\left[x_2\right])\right] \\
&= \mathbb{E}\left[x_1x_2 - x_1\mathbb{E}\left[x_2\right] - x_2\mathbb{E}\left[x_1\right] + \mathbb{E}\left[x_1\right]\mathbb{E}\left[x_2\right]\right] \\
&= \mathbb{E}\left[x_1x_2\right] - \mathbb{E}\left[x_1\right]\mathbb{E}\left[x_2\right] \\
\implies \mathbb{E}\left[x_1x_2\right] &= \mathbb{E}\left[x_1\right]\mathbb{E}\left[x_2\right] + \mathrm{Cov}(x_1, x_2).
\end{aligned} \tag{6.39}$$

Remark: Another useful property about covariances is the following identity: let a, b, c, and d be constant values in the real line. Then,

$$\begin{aligned}
\mathrm{Cov}(ax_1 + bx_2, cx_3 + dx_4) = {} & ac\mathrm{Cov}(x_1, x_3) + ad\mathrm{Cov}(x_1, x_4) \\
& + bc\mathrm{Cov}(x_2, x_3) + bd\mathrm{Cov}(x_2, x_4).
\end{aligned} \tag{6.40}$$

Hence, the expected value of $\Delta_{\mathcal{S}}(\hat{\mathbf{z}}_t^{(i)})$ can be expressed as in Theorem 6.3.1:

Theorem 6.3.1. *The expected value of* $\Delta_{\mathcal{S}}(\hat{\mathbf{z}}_t^{(i)})$ *is*

$$\begin{aligned}\mathbb{E}\left[\Delta_{\mathcal{S}}(\hat{\mathbf{z}}_t^{(i)})\right] = &\left(\hat{m}_{1,t}^{(i)} - \hat{m}_{1,t}^{(i-1)}\right)\left(\hat{m}_{2,t}^{(i+1)} - \hat{m}_{2,t}^{(i)}\right)\\ &+ Cov(\hat{z}_{1,t}^{(i)}|\hat{m}_{2,t}^{(i)},\quad \hat{z}_{2,t}^{(i+1)}|\hat{m}_{1,t}^{(i+1)})\\ &- Cov(\hat{z}_{1,t}^{(i)}|\hat{m}_{2,t}^{(i)},\quad \hat{z}_{2,t}^{(i)}|\hat{m}_{1,t}^{(i)})\\ &- Cov(\hat{z}_{1,t}^{(i-1)}|\hat{m}_{2,t}^{(i-1)},\quad \hat{z}_{2,t}^{(i+1)}|\hat{m}_{1,t}^{(i+1)})\\ &+ Cov(\hat{z}_{1,t}^{(i-1)}|\hat{m}_{2,t}^{(i-1)},\quad \hat{z}_{2,t}^{(i)}|\hat{m}_{1,t}^{(i)}).\end{aligned} \tag{6.41}$$

Proof. Using the identities presented in the last four remarks, after simple algebraic rearrangement using the aforementioned properties, we have:

$$\begin{aligned}\mathbb{E}\left[\Delta_{\mathcal{S}}(\hat{\mathbf{z}}_t^{(i)})\right] &= \mathbb{E}\left[(\hat{z}_{1,t}^{(i)}|\hat{m}_{2,t}^{(i)} - \hat{z}_{1,t}^{(i-1)}|\hat{m}_{2,t}^{(i-1)})(\hat{z}_{2,t}^{(i+1)}|\hat{m}_{1,t}^{(i+1)} - \hat{z}_{2,t}^{(i)}|\hat{m}_{1,t}^{(i)})\right]\\ &= \mathbb{E}\left[\hat{z}_{1,t}^{(i)}|\hat{m}_{2,t}^{(i)} - \hat{z}_{1,t}^{(i-1)}|\hat{m}_{2,t}^{(i-1)}\right] \times \mathbb{E}\left[\hat{z}_{2,t}^{(i+1)}|\hat{m}_{1,t}^{(i+1)} - \hat{z}_{2,t}^{(i)}|\hat{m}_{1,t}^{(i)}\right]\\ &\quad + \mathrm{Cov}(\hat{z}_{1,t}^{(i)}|\hat{m}_{2,t}^{(i)} - \hat{z}_{1,t}^{(i-1)}|\hat{m}_{2,t}^{(i-1)},\quad \hat{z}_{2,t}^{(i+1)}|\hat{m}_{1,t}^{(i+1)} - \hat{z}_{2,t}^{(i)}|\hat{m}_{1,t}^{(i)})\\ &= \left(\mathbb{E}\left[\hat{z}_{1,t}^{(i)}|\hat{m}_{2,t}^{(i)}\right] - \mathbb{E}\left[\hat{z}_{1,t}^{(i-1)}|\hat{m}_{2,t}^{(i-1)}\right]\right)\left(\mathbb{E}\left[\hat{z}_{2,t}^{(i+1)}|\hat{m}_{1,t}^{(i+1)}\right] - \mathbb{E}\left[\hat{z}_{2,t}^{(i)}|\hat{m}_{1,t}^{(i)}\right]\right)\\ &\quad + \mathrm{Cov}(\hat{z}_{1,t}^{(i)}|\hat{m}_{2,t}^{(i)} - \hat{z}_{1,t}^{(i-1)}|\hat{m}_{2,t}^{(i-1)},\quad \hat{z}_{2,t}^{(i+1)}|\hat{m}_{1,t}^{(i+1)} - \hat{z}_{2,t}^{(i)}|\hat{m}_{1,t}^{(i)})\\ &= \left(\hat{m}_{1,t}^{(i)} - \hat{m}_{1,t}^{(i-1)}\right)\left(\hat{m}_{2,t}^{(i+1)} - \hat{m}_{2,t}^{(i)}\right)\\ &\quad + \mathrm{Cov}(\hat{z}_{1,t}^{(i)}|\hat{m}_{2,t}^{(i)},\quad \hat{z}_{2,t}^{(i+1)}|\hat{m}_{1,t}^{(i+1)}) - \mathrm{Cov}(\hat{z}_{1,t}^{(i)}|\hat{m}_{2,t}^{(i)},\quad \hat{z}_{2,t}^{(i)}|\hat{m}_{1,t}^{(i)})\\ &\quad - \mathrm{Cov}(\hat{z}_{1,t}^{(i-1)}|\hat{m}_{2,t}^{(i-1)},\quad \hat{z}_{2,t}^{(i+1)}|\hat{m}_{1,t}^{(i+1)}) + \mathrm{Cov}(\hat{z}_{1,t}^{(i-1)}|\hat{m}_{2,t}^{(i-1)},\quad \hat{z}_{2,t}^{(i)}|\hat{m}_{1,t}^{(i)}).\end{aligned}$$

□

Note how $\mathbb{E}\left[\Delta_{\mathcal{S}}(\hat{\mathbf{z}}_t^{(i)}|\hat{\mathbf{z}}_{t+h}^{(i)})\right]$ depends not only on the mean vectors of $\hat{\mathbf{z}}_t^{(i)}|\hat{\mathbf{z}}_{t+h}^{(i)}$ and of its non-dominance class neighbors in the objective space ($\hat{\mathbf{z}}_t^{(i-1)}|\hat{\mathbf{z}}_{t+h}^{(i-1)}$ and $\hat{\mathbf{z}}_t^{(i+1)}|\hat{\mathbf{z}}_{t+h}^{(i+1)}$), but also on the covariances of the conditional (marginal) distributions. Therefore, the learned *predictive covariance* structure among the objective functions is naturally used to provide uncertainty awareness in the maximization in AS-MOO. That is to say, the expected $\Delta_{\mathcal{S}}$ value of a given anticipatory objective distribution $\hat{\mathbf{z}}_t^{(i)}|\hat{\mathbf{z}}_{t+h}^{(i)}$ also accounts for second order statistical information between neighboring anticipatory distributions.

Remark: Although online anticipatory learning is applied *independently* for each candidate solution, $\mathbb{E}\left[\Delta_{\mathcal{S}}(\hat{\mathbf{z}}_t^{(i)}|\hat{\mathbf{z}}_{t+h}^{(i)})\right]$ provides a means to *combine* the learned covariance from neighboring anticipatory distributions, with the goal of guiding the AS-MOO process towards the stochastic Pareto frontier.

Remark: The learned covariances among the conflicting noisy objective-functions evaluations encode valuable statistical trade-off information in the vicinity of each anticipatory distribution.

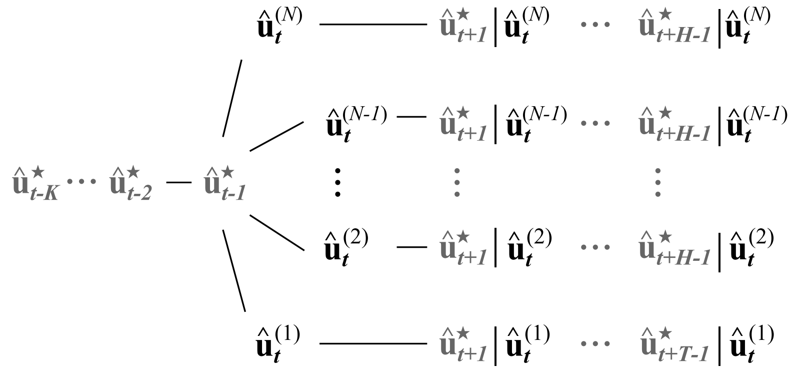


Figure 6.4: The N possible predictive trajectories for the sequence of future $(H-1)N$ Estimated Maximal Flexible Choices (EMFCs) over time, conditioned on each available candidate Pareto-efficient decision in the estimated essential set $\hat{\mathcal{U}}_t^{N\star}$.

Remark: The pairwise covariance terms appearing in Theorem 6.3.1 merge trade-off information between multiple neighboring anticipatory distributions in the objective space. The expected anticipatory $\mathcal{S}$-metric contribution thus allows for the *mutual interference* and *exchange of trade-off information* between parallel neighboring anticipatory paths.

6.4 Estimating the Anticipated Maximal Flexible Choice

In the TL regime, the future anticipatory Pareto Sets attaining maximal future hypervolume depends on the identification of the Anticipated Maximal Flexible Choices (AMFCs) in early decision periods. We recall that such dependency is captured with the notation $\mathcal{U}_{t+h}^{N\star}|\mathbf{u}_t^\star$ in **chapter 5**, Eq. (5.11). The identification of which AMFC leads to maximal hypervolume Pareto Sets in the future is thus required. In order to solve this problem with minimal approximation error[7], one would need to fully traverse the *predicted* decision tree depicted in Fig. 5.4, what would require an exponential $O(N^H)$ number of recursive re-optimizations, where N is the number of mutually non-dominated AMFCs specified in the AS-MOO problem and H is the number of decision periods. This computational burden is infeasible for any practical means when N and H are large.

In reinforcement learning methods arising in Approximate Dynamic Programming (ADP [168], or neural-dynamic programming [28]), Monte Carlo (MC) simulation is often used to explore the decision tree so to approximate the value of *policies* – which correspond to input/output temporal mappings linking states at specific decision periods to optimal actions/decisions. Those methods can be regarded as *scenario-based* approaches for solving online optimization problems. Although, in ADP, the MC methods can greatly reduce the exponential cost of exploring a simulated graph of state transitions (known as a Markov Decision Process *transition graph*), the goal of approximately learning a complete optimal policy is more ambitious (and therefore considerably more computational intensive) than to produce decisions on-the-fly.

[7]This problem cannot be solved "exactly" (i.e., with minimum error), unless one has access to perfect information about the future trajectory of the state vector $\mathbf{x}_t$.

We therefore propose in this thesis an online *prediction-based* approach for assessing the decision-tree of Fig. 5.4, requiring only bilinear complexity $O(NH)$ for which, in general, $H < N$. Besides, because our anticipatory learning methods rely on such surrogate future estimatives, no re-optimizations are needed. The goal here is to take advantage of our problem-solving internal predictive models with minimum computational effort in terms of repeated optimizations and of objective function evaluations.

The reduction in computational complexity is achieved by independently tracking and predicting (1) each element of the sets of candidate alternatives $\hat{\mathcal{U}}^{N\star}_{t+h}$ (for $1 \leq h \leq H$); (2) the corresponding objective vectors $\mathbf{z}^{(i)}_{t+h}$; and (3) the sequence of implemented and estimated AMFCs $\mathbf{u}^\star_1, \cdots, \mathbf{u}^\star_{t+h}$ (for $0 \leq h \leq H$):

Step 1 Predict the future sets $\hat{\mathcal{U}}^N_{t+1}, \cdots, \hat{\mathcal{U}}^N_{t+H-1}$ element-wise;

Step 2 Predict the objective component $\mathbf{g}$, without the cost component $\mathbf{h}$ (see **chapter 5**, Eq. (5.4)): $\{\hat{\mathbf{z}}^{(i)}_{t+h}\}^{H-1}_{h=1}$;

Step 3 Estimate $(H-1)N$ AMFCs (see Fig. 6.4) conditioned at $\hat{\mathbf{u}}^{(i)}_t \in \hat{\mathcal{U}}^{N\star}_t$: $\left\{\hat{\mathbf{u}}^\star_{t+h}|\hat{\mathbf{u}}^{(i)}_t\right\}^{H-1}_{h=1}$;

Step 4 Predict the $N \times N$ cost-adjusted objective distributions, *given* $\hat{\mathbf{u}}^{(j)}_t$ ($i, j \in \{1, \cdots, N\}$): $\{\hat{\mathbf{z}}^{(i)}_{t+h}|\hat{\mathbf{u}}^{(j)}_t\}^{H-1}_{h=1}$;

Step 5 Apply OAL and identify the AMFC using E Monte Carlo simulated scenarios (see **chapter 7**, section 7.2.2):

$$\hat{\mathbf{u}}^\star_t = \arg \max_{\hat{\mathbf{u}}_t \in \hat{\mathcal{U}}^{N\star}_t} \frac{1}{E} \sum_{e=1}^{E} \mathcal{S}\left(\sum_{h=1}^{H-1} \lambda_{t+h} \left\{\hat{\mathbf{z}}^{(i)}_{e,t+h}|\hat{\mathbf{u}}_t\right\}^N_{i=1}\right). \tag{6.42}$$

With this approach, no re-optimizations are required because the predicted optimal AMFC sets are taken into account instead. Stinis [198], for instance, has approached global stochastic (single-objective) global optimization as a filtering (tracking) problem in the search space using particle filters [75]. Our idea is somewhat related to Stinis' [198], the difference being that we aim to track multiple moving (Pareto) optima across decision periods, whereas Stinis [198] treated the tracking of a single, static global best solution across the search process iterations.

6.5 An Anticipatory Metaheuristic for AS-MOO

We propose in this section a framework for MOO evolutionary-inspired metaheuristics that are not only able to anticipate trade-off solutions attaining good performance in future environments, but also to estimate the future $\mathcal{S}$-metric contributions, which are indispensable features for optimizing the future hypervolume in AS-MOO problems. The proposed metaheuristics rely on the estimation of the anticipatory distributions in the objective space, and on the tracking and estimation of Anticipated Maximal Flexible Choices (AMFCs) in the search space (see section 6.1). Although there are several ways to design MOO metaheuristics for approximating the

Pareto Frontier, this thesis advocates the so called *non-dominated sorting* algorithms that operate by means of partitioning the population of candidate objective vectors into non-dominance classes, such as the NSGA-II [66] (see **chapter 3**, section 3.4.1).

The non-dominated sorting MOO algorithms convey valuable properties that can be easily explored for heuristically solving SMOO problems. For instance, given a set $\mathcal{P}$ of candidate solutions, Azevedo and Araújo [13] showed that the solutions attaining maximal Pareto Non-Dominance Probability (PNDP[8]) within a given non-dominance class $\mathcal{C}_i \subset \mathcal{P}$ are always assigned greater or equal PNDP than that assigned to the maximal PNDP solutions belonging to inferior classes $\mathcal{C}_j$ (with $j > i$). This result, summarized in the Monotonicity of Maximal PNDPs Theorem [13], strongly motivates the design of metaheuristics biased towards solutions belonging to superior non-dominance classes, while progressively discarding those in inferior classes.

6.5.1 Sorting the Population of Random Objective Vectors

One significant challenge to overcome when designing non-dominated sorting metaheuristics for SMOO problems, however, is how to sort and partition random objective vectors into non-dominance classes. A natural approach is to consider *multivariate stochastic dominance* techniques to determine whether $\mathbf{z}^{(i)}$ (stochastically) dominates $\mathbf{z}^{(j)}$ (with $i \neq j$). For instance, there are known sufficient conditions for the Gaussian multivariate case when the DM has e.g. an increasing convex utility function:

Theorem 6.5.1. *The random objective vector $\mathbf{z}^{(i)}$ is said to stochastically dominate $\mathbf{z}^{(j)}$ if (a) $\mathbf{m}_{\mathbf{z}^{(i)}} \preceq \mathbf{m}_{\mathbf{z}^{(j)}}$; and (b) $\mathbf{\Sigma}_{\mathbf{z}^{(j)}} - \mathbf{\Sigma}_{\mathbf{z}^{(i)}}$ is positive semidefinite.*

Proof. See Müller [158]. □

Note that the theorem relies on the deterministic Pareto order ($\preceq$) over the expected objective values. In addition, it assumes that the (multiattribute) utility function of the DM is chosen from a family of increasing convex utility functions mapping $\mathcal{R}^m$ into $\mathcal{R}$ (m is the number of objective functions), and that $\mathbf{z}^{(i)} \sim \mathcal{N}(\mathbf{m}_{\mathbf{z}^{(i)}}, \mathbf{\Sigma}_{\mathbf{z}^{(i)}})$ and $\mathbf{z}^{(j)} \sim \mathcal{N}(\mathbf{m}_{\mathbf{z}^{(j)}}, \mathbf{\Sigma}_{\mathbf{z}^{(j)}})$. Nevertheless, finding two random objective vectors simultaneously satisfying both conditions (a) and (b) of Theorem 6.5.1 can be challenging, undermining the selective pressure of a non-dominated sorting metaheuristic designed for SMOO problems.

Another approach for sorting and partitioning random objective vectors is to assess the probability of $\mathbf{z}^{(i)} \preceq \mathbf{z}^{(j)}$, which can be easily computed for multivariate Gaussians but requires the determination of thresholds. For instance, the two objective vectors can be classified as mutually non-dominated if neither one dominates the other with a probability higher than a certain critical level. Nevertheless, we argue that, since we are already able to handle the estimated uncertainty (see section 6.1) and to combine the learned predictive correlation into the computation of $\mathbb{E}\left[\Delta_{\mathcal{S}}\right]$ (see section 6.3), there is little need to incorporate uncertainty awareness *directly* into the dominance relation. Therefore, our proposed non-dominated sorting procedures are executed in terms of the deterministic Pareto Dominance over the estimated means of the

[8]The probability that a solution $\mathbf{u} \in \mathcal{P}$ is not dominated by any other solution $\mathbf{u}' \in \mathcal{P}$. It is equivalent to the probability of $\mathbf{u} \in \mathcal{C}_1 \subset \mathcal{P}$, where $\mathcal{C}_1$ is the class of non-inferior solutions, see **chapter 3**, section 3.4.1.

Input: The anticipation horizon H
Input: $\hat{\mathcal{U}}_{t-K:t-1}^{N}$ and the implemented EMFCs $\left\{\mathbf{u}_{t-h}^{\star}\right\}_{h=1}^{K}$
Input: Historical $\mathcal{X}_{t-K:t-1} = \{\mathbf{x}_{t-K}, \cdots, \mathbf{x}_{t-1}\}$
Output: A maximal Hypv set $\hat{\mathcal{U}}_{t}^{N\star}$ for H-1 steps ahead
Output: An EMFC $\hat{\mathbf{u}}_{t}^{\star}$ satisfying **chapter 5** Eq. (5.15)
procedure ASMS-EMOA
  Set the generation count as $g \leftarrow 0$
  Set the current population as $\hat{\mathcal{U}}_{t}^{N} \leftarrow \hat{\mathcal{U}}_{t-1}^{N}$
  for $i = 1$ **to** N **do**
    Compute the anticipatory distribution $\hat{\mathbf{z}}_{t}^{(i)}|\hat{\mathbf{z}}_{t+1:t+H-1}^{(i)}$ using OAL ▷ Pseudocode 7
  end for
  repeat
    Compute $\{\mathcal{C}_j\}_{j=1}^{C}$ from $\mathbb{E}\left[\{\hat{\mathbf{z}}_{t}^{(i)}|\hat{\mathbf{z}}_{t+1:t+H-1}^{(i)}\}_{i=1}^{N}\right]$
    Compute $\mathbb{E}\left[\Delta_{\mathcal{S}}(\hat{\mathbf{z}}_{t}^{(i)}|\hat{\mathbf{z}}_{t+1:t+H-1}^{(i)}, \mathcal{C}_j)\right]$, for each $\hat{\mathbf{z}}_{t}^{(i)}|\hat{\mathbf{z}}_{t+1:t+H-1}^{(i)} \in \mathcal{C}_j$ and $1 \leq j \leq C$
    Generate a new decision $\hat{\mathbf{u}}_{t}'$ from $\hat{\mathcal{U}}_{t}^{N}$
    Augment $\hat{\mathcal{U}}_{t}^{N+1} \leftarrow \hat{\mathcal{U}}_{t}^{N} \cup \{\hat{\mathbf{u}}_{t}'\}$
    Compute the anticipatory distribution $\hat{\mathbf{z}}_{t}'|\mathbf{z}_{t+1:t+H-1}'$ using OAL ▷ Pseudocode 7
    Reduce $\hat{\mathcal{U}}_{t}^{N} \leftarrow$ ANTICIPATORY REDUCE$(\hat{\mathcal{U}}_{t}^{N+1})$
    Advance to the next generation $g \leftarrow g + 1$
  until Stopping criteria are not met
  Identify the Anticipated Maximal Flexible Choice $\hat{\mathbf{u}}_{t}^{\star}$ using Eq. (6.42) ▷ Section 6.4
  return $\left\{\hat{\mathcal{U}}_{t}^{N\star}, \hat{\mathbf{u}}_{t}^{\star}\right\}$
end procedure

Pseudocode 8: The Anticipatory $\mathcal{S}$-Metric Selection EMOA

Input: An augmented candidate set $\hat{\mathcal{U}}_{t}^{N+1}$
Output: A reduced candidate set $\hat{\mathcal{U}}_{t}^{N}$
procedure REDUCE
  Compute $\{\mathcal{C}_j\}_{j=1}^{C}$ from $\mathbb{E}\left[\{\hat{\mathbf{z}}_{t}^{(i)}|\hat{\mathbf{z}}_{t+1:t+H-1}^{(i)}\}_{i=1}^{N+1}\right]$
  Select $\hat{\mathbf{u}}^{*} = \arg\min_{\hat{\mathbf{u}} \in \mathcal{C}_C} \mathbb{E}\left[\Delta_{\mathcal{S}}(\hat{\mathbf{u}}, \mathcal{C}_C)\right]$
  return $\hat{\mathcal{U}}_{t}^{N+1} \setminus \{\hat{\mathbf{u}}^{*}\}$
end procedure

Pseudocode 9: Anticipatory Reduce $\mathcal{U}_{t}^{N+1}$

random objective vectors, what is in accordance with our stochastic Pareto frontier definition (see Fig. 6.5 and **chapter 5**, definition 5.3).

6.5.2 The Anticipatory S-Metric Selection EMOA

We recall from **chapter 3** that the S-Metric Selection EMOA (SMS-EMOA) [32] attempts to explicitly maximize $\mathcal{S}$ by progressively improving $\Delta_{\mathcal{S}}$ within the computed non-dominance classes. The SMS-EMOA takes advantage of the Pareto dominance relations among (deterministic) objective vectors by partitioning the population of candidate decisions into C classes

$\mathcal{C}_1, \cdots, \mathcal{C}_C$, so that the mutually non-dominated solutions composing the first class are never replaced by the dominated ones in $\bigcup_{n=2}^{C} \mathcal{C}_n$, whereas the least $\Delta_{\mathcal{S}}$ contributing decisions belonging to $\mathcal{C}_C$ are progressively replaced by heuristically generated candidate trade-off solutions. The $\Delta_{\mathcal{S}}$ contribution therefore determines the selective pressure within each non-dominance class. By progressively replacing the least contributing solutions, the hopes are that the maximization of $\mathcal{S}$ can be sustained along the optimization process in order to attain $\mathcal{C}_1 \subseteq \mathcal{U}^N \subset \mathcal{U}^{N\star}$, what is ultimately the goal of all set-based MOO solvers [229].

In this thesis, we propose tailoring the SMS-EMOA to approximately solve AS-MOO problems, leading to the Anticipatory SMS-EMOA (ASMS-EMOA)[9]. Not only the ASMS-EMOA improves over its myopic and deterministic counterpart (SMS-EMOA) by providing internal predictive and estimation Bayesian models to track time-varying decision and objective vectors (sections 6.2 and 6.1), but also it provides the means for coping with stochastic uncertainty by implementing the proposed anticipatory learning rules and by computing the expected anticipatory $\mathcal{S}$-Metric contributions $\mathbb{E}\left[\Delta_{\mathcal{S}}\right]$ over the estimated anticipatory distributions.

The ASMS-EMOA thus utilizes the KF estimation and prediction in the objective space and the proposed Dirichlet Bayesian MAP tracking and prediction in the search space. Moreover, in the stochastic case, the candidate trade-off solutions and their respective random objective vectors are partitioned into non-dominance classes (see section 6.5.1 and also Eq. (3.22) in **chapter 3**) using the standard Pareto dominance *over the estimated mean objective values*, so as to progressively improve over the partitioned decisions for approximating the stochastic Pareto frontier.

The Pseudocodes 8 and 9 describe the proposed anticipatory MOO metaheuristic, whose outputs are (1) the set of mutually non-dominated (w.r.t. the mean vectors) candidate solutions set ($\mathcal{U}_t^{N\star}$), for both the TL and TLF regimes, and the anticipatory decision ($\mathbf{u}_t^\star \in \mathcal{U}_t^{N\star}$) satisfying AS-MOO Eq. (5.15), **chapter 5**, in the TL regime.

6.6 Summary of the Contributions

This chapter's contributions to the thesis are as follows:

1. Two novel approaches for tracking and predicting time varying quantities in the objective and search spaces subject to stochastic uncertainty have been devised;

2. It proposed the principle of *online anticipatory learning* for combining stochastic predictive knowledge, with different heuristics for automatically determine discount rates in terms of the estimated stochastic uncertainty as well as of the observed residuals;

3. It showed a way to incorporate uncertainty awareness in the computation of the expected anticipatory $\mathcal{S}$-Metric contributions;

4. A procedure for identifying the candidate anticipatory decision maximizing the future $\mathcal{S}$-Metric in the time-linkage regime has been proposed; and

[9] Source codes at http://www.researchgate.net/profile/Carlos_Azevedo2.

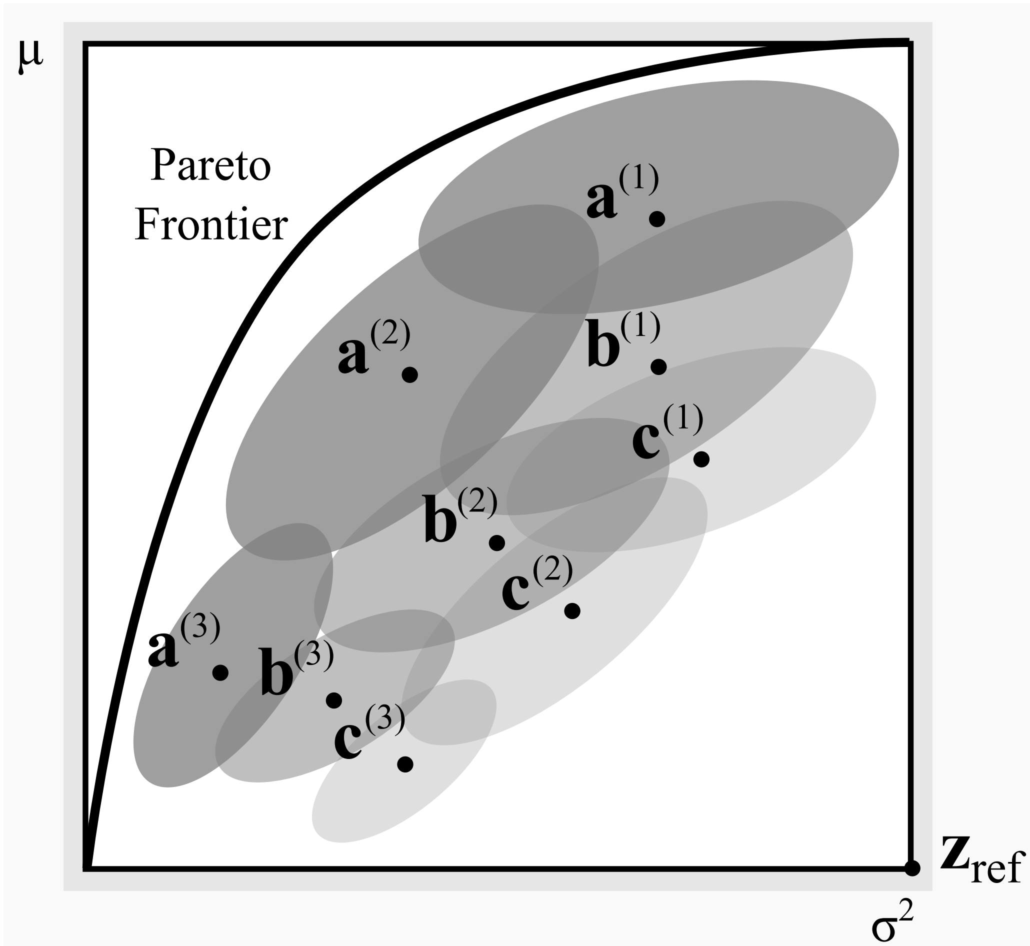


Figure 6.5: Non-dominance sorting procedure over the mean objective vectors. The first class is composed from alternatives $\mathcal{C}_1 = \{\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \mathbf{a}^{(3)}\}$, whose mean vectors are not dominated by any other mean vector of other alternatives in the population. Classes two and three are composed of alternatives $\mathcal{C}_2 = \{\mathbf{b}^{(1)}, \mathbf{b}^{(2)}, \mathbf{b}^{(3)}\}$ and $\mathcal{C}_3 = \{\mathbf{c}^{(1)}, \mathbf{c}^{(2)}, \mathbf{c}^{(3)}\}$, respectively. Notice that if alternatives from class $\mathcal{C}_1$ were removed from the population, then the mean vectors of the alternatives from $\mathcal{C}_2$ would not been dominated by any other remaining alternative mean vector. Moreover, notice that, in this sorting procedure, alternatives from lower classes still have a chance (although with low probability) of dominating the alternatives from higher classes.

5. Finally, an anticipatory $\mathcal{S}$-Metric selection algorithm has been proposed for approximately solving anticipatory SMOO problems under two different temporal regimes.

In the next chapter, we provide the experimental results of the application of the flexible anticipatory multi-objective machine learning and optimization models and algorithms devised in this chapter for bi-objective, noisy, sequential, and cost-adjusted portfolio selection.

Chapter 7

Case Study: Online Anticipatory Learning for Flexible Multi-Objective Portfolio Selection

The hardest hit, as everywhere, are those who have no choice.

– Theodor Adorno

Freedom is not worth having if it does not include the freedom to make mistakes.

–Mahatma Gandhi

This chapter presents the application of the AS-MOO/MCDM methodology for approximating sets of efficient portfolios lying in the Stochastic Pareto Frontier (SPF) that can reveal the trade-offs between expected return maximization and expected risk minimization within a non-stationary and stochastic investment environment. The trade-off portfolio predicted to lead to future Pareto sets with maximal expected hypervolume is taken as a basis for rebalancing the current implemented portfolio. Thus, the evolved portfolios are deployed in a time-varying, post-optimization unseen environment. Experiments on three real-world financial stock markets as well as on simulated benchmark data show that the ASMS-EMOA approach is able to obtain future SPF approximation sets conveying higher average hypervolume for future out-of-sample investment rounds when compared to the myopic SMS-EMOA [32]. Moreover, implementing the Estimated Maximal Flexible Choices (EMFCs) was confirmed as a better strategy than to randomly take a portfolio from the evolved approximated Pareto set.

7.1 Experimental Design

The aim of the experimental design is to investigate the effects of the following two factors on the sample average Hypervolume (Hypv) computed over future out-of-sample investment data:

Factor 1 (a) tracking and prediction in the decision and objective spaces to evolve an anticipatory SPF, compared to (b) the assumption of a static SPF of a myopic optimizer;

Factor 2 (a) identifying and implementing EMFCs over time, compared to (b) randomly selecting portfolios from the Pareto-flexible set.

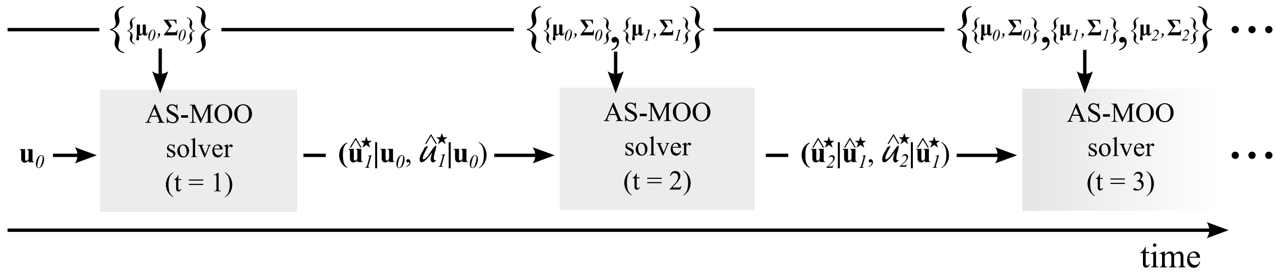


Figure 7.1: Online AS-MOO methodology for anticipating Estimated Maximal Flexible Choices (EMFCs), as a stream of state parameters is accumulated over time. For portfolio selection, $\mathbf{u}_0$ is an initial investment allocation and $\{\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t\}_{t=1}^{T}$ is a stream of parameters of the dynamic joint return distribution, which is assumed to be a Gaussian process. The historical time-dependent Pareto Sets and EMFCs also become available to the AS-MOO solvers at subsequent periods.

We name those the (1) *anticipation* and the (2) *DM* factors. The formulated hypotheses are thus that the Hypv computed over the evolved SPF approximation for future unseen environments is higher when utilizing strategies (F1.a) and (F2.a) as opposed to utilizing (F1.b) and (F2.b).

7.1.1 Investigated Algorithmic Variants: Indiffirence vs. Flexibility

As stated in **chapter 1**, the proposed anticipatory methodology does not find parallel in the literature. That is to say, no previous studies that we are aware of experimented with evolving stochastic Pareto frontiers over time, anticipating changes in multiple noisy objective functions, while adjusting for costs in a time-linkage setting. The EMOO/MCDM literature thus lacks competent algorithms that could met our methodology main requirement: outputting a sequence of Pareto-efficient solutions predicted to improve future diversity of choice (see Fig. 7.1.1).

In light of the shortage of methods for AS-MOO, we follow a similar approach as that of Farina et al. [84], where the first (and now widespread) set of benchmark Dynamic MOO (DMOO) problems was defined. Because DMOO algorithms were virtually nonexistent, they investigated a baseline algorithm, but did not attempt to build over the static EMOO algorithms that were available at the time. In this thesis, we not only adapt a state-of-the-art EMOO algorithm as a *myopic baseline* for AS-MOO, but we build over it to compose the proposed anticipatory algorithm (**chapter 6**, section 6.5). The SMS-EMOA [32] baseline is thus instantiated when (i) the window size used for KF and DD tracking is zero; and (ii) portfolios are randomly chosen from the evolved Pareto-flexible set.

Investigated Factors and Algorithms

The anticipation factor is controlled by four levels of window size (K): $K \in \{0, 1, 2, 3\}$. For $K = 0$, we have the myopic baseline SMS-EMOA (SMS, in short) for which constant predictions are made, i.e., $\hat{\mathbf{u}}_{t+h} \sim \hat{\mathbf{u}}_t$ and $\hat{\mathbf{z}}_{t+h} \sim \hat{\mathbf{z}}_t$. The stationarity assumption implies the estimated Temporal Incomparability Probabilities (TIPs, see **chapter 6**, Eq. (6.4)) equal 1/2 and, hence, all anticipation rates are self-adjusted to $\lambda_t^{(i)} = 1$ $(i = 1, \cdots, N)$. This case is equivalent to a DM betting on a static market. For $K > 0$, the past solutions and objective distributions obtained in the latest K periods serve as input to the KF and DD tracking methods, in which case anticipation is enabled.

The DM factor determines whether (a) predicted maximal flexible choices or (b) randomly drawn portfolios from the Pareto-flexible Set are implemented. The goal is to contrast an *indifferent* DM (Random DM, RDM) – for whom two or more paths may be equally satisfying – from a *flexible* DM, whose preferences cannot be reliably elicited and are thus left open: the flexible DM (maximal Hypv, mHDM). For the mHDM, the strategy becomes postponing preferences specification by implementing flexible choices. Therefore, four variants are investigated for assessing the anticipation and DM effects on future diversity of choice: (F1.a) + (F2.a) = ASMS/mHDM; (F1.a) + (F2.b) = ASMS/RDM; (F1.b) + (F2.a) = SMS/mHDM; and (F1.b) + (F2.b) = SMS/RDM:

ASMS/mHDM Bayesian tracking and prediction are performed in both decision and objective space ($K > 0$, Factor 1.a). Predicted flexible maximal choices foreseen to yield to maximal expected hypervolume trade-off sets are implemented (Factor 2.a);

ASMS/RDM Bayesian tracking and prediction are performed in both decision and objective space ($K > 0$, Factor 1.a). Randomly selected portfolios from the evolved approximation to the Pareto-flexible set are implemented (Factor 2.b);

SMS/mHDM The stationary assumption is adopted and constant predictors are integrated into a myopic algorithm ($K = 0$, Factor 1.b). Predicted flexible maximal choices foreseen to yield to maximal expected hypervolume trade-off sets are implemented (Factor 2.a);

SMS/RDM The stationary assumption is adopted and constant predictors are integrated into a myopic algorithm ($K = 0$, Factor 1.b). Randomly selected portfolios from the evolved approximation to the Pareto-flexible set are implemented (Factor 2.b).

7.2 Solving Portfolio Selection with AS-MOO

We validate the AS-MOO methodology on three real-world stock indices: London's Financial Times Stock Exchange (FTSE–100), Dow Jones Index (DJI), and Hong Kong's Hang Seng Index (HSI). Also, we design an artificial instances generator, controlling for periodicity and for severity of change (see Sec. 7.2.1), thus providing different regime switching patterns.

Let $\mathbf{r} \in \mathcal{R}^N$ be the return vector of N risky assets, in which $\boldsymbol{\mu}_\mathbf{r}$ and $\boldsymbol{\Sigma}_\mathbf{r}$ are its mean and covariance, and let $\mathbf{u} \in S^{N-1}$ denote the proportions of wealth to be invested. Then, a classical myopic formulation for risk minimization [151] is: $\mathbf{u}^\star = \arg\min_\mathbf{u} \{\mathbf{u}^\intercal \boldsymbol{\Sigma}_\mathbf{r} \mathbf{u} : \boldsymbol{\mu}_\mathbf{r}^\intercal \mathbf{u} \geq r_0\}$, in which r_0 is the minimum acceptable return. The application of AS-MOO for approximating portfolios lying in the stochastic Pareto frontier can reveal the trade-offs between expected return maximization and expected risk minimization within a nonstationary and probabilistic investment environment. In AS-MOO, the investment decision leading to future Pareto-flexible Sets with maximal expected Hypv is chosen as a basis for rebalancing the current implemented portfolio. The state parameters $\mathcal{X}_t = \left\{\hat{\boldsymbol{\mu}}_{\mathbf{r},t}, \hat{\boldsymbol{\Sigma}}_{\mathbf{r},t}\right\}$ must be estimated from the available data.

We then propose the sequential formulation

$$\min_{\mathbf{u}_t} \quad \mathbf{u}_t^\intercal \hat{\mathbf{\Sigma}}_{\mathbf{r},t} \mathbf{u}_t \tag{7.1}$$

$$\max_{\mathbf{u}_t} \quad \hat{\boldsymbol{\mu}}_{\mathbf{r},t}^\intercal \mathbf{u}_t - h(\mathbf{u}_t, \mathbf{u}_{t-1}^\star) \tag{7.2}$$

$$\text{s.t.} \quad c_l \leq c(\mathbf{u}_t) \leq c_u, \tag{7.3}$$

where $c(\mathbf{u}_t)$ computes the number of assets in $\mathbf{u}_t$ with non-zero weight ($u_t > 0$). Following the AS-MOO notation, $\mathbf{z}_t|\mathbf{u}_{t-1}^\star = \mathbf{g}(\mathbf{u}_t, \mathcal{X}_t) + \mathbf{h}(\mathbf{u}_t, \mathbf{u}_{t-1}^\star)$, where:

$$\mathbf{g}(\mathbf{u}_t, \mathcal{X}_t) = \left(\mathbf{u}_t^\intercal \hat{\mathbf{\Sigma}}_{\mathbf{r},t} \mathbf{u}_t \quad \hat{\boldsymbol{\mu}}_{\mathbf{r},t}^\intercal \mathbf{u}_t\right)^\intercal, \text{ and} \tag{7.4}$$

$$\mathbf{h}(\mathbf{u}_t, \mathbf{u}_{t-1}^\star) = \left(0 \quad h(\mathbf{u}_t, \mathbf{u}_{t-1}^\star)\right)^\intercal, \tag{7.5}$$

in which h is a cost function representing all incurring transaction fees and commissions. We consider the risk of $\mathbf{u}_t$ to not be affected by the previous portfolio.

7.2.1 Simulated Benchmark Data Generator

Notwithstanding the hard quest of simultaneously handling various sources of uncertainty, many problems could be more realistic tackled if modeled as dynamic stochastic multi-objective ones, what could dramatically expand EMOO applications. Still, most works end up only handling isolated aspects, partly due to the complexity of assessing the performance of EMOO algorithms in properly designed benchmark scenarios (see **chapter 3**). For a survey of benchmark problems in the case of deterministic dynamic MOO, we refer to [114].

Therefore, for comparing the behavior of myopic and anticipatory optimizers under controlled conditions, we devised a random sequential portfolios selection instance generator, with two factors: (1) the periodicity; and (2) the severity with which disruptive changes affect the parameters of the joint assets return distribution. The modeled disruption is an additive noise process, in which the parameters are perturbed as $\mathcal{X}_{\text{new}} = (1-\eta)\mathcal{X}_{\text{old}} + \eta\mathcal{P}_{\text{random}}$, where $\mathcal{P}_{\text{random}}$ is a set of random parameters – i.e., a random mean vector and a random covariance matrix – and η controls the *severity* with which the random parameters are absorbed into the update of $\mathcal{X}_{\text{new}}$.

Let $\mathcal{X}_1 = \{\boldsymbol{\mu}_1, \mathbf{\Sigma}_1\}$ be the initial set of parameters, where $\boldsymbol{\mu}_1$ is the mean vector and $\mathbf{\Sigma}_1$ is the covariance matrix. Let t_0 be the period when the *first disruptive change occurs*. The set of parameters at period t_0 is described as a convex combination between the parameters at the starting period in $t = 1$, $\mathcal{X}_1$, and the set of random parameters generated at $t = t_0, \mathcal{P}_{t_0}$:

$$\mathcal{X}_{t_0} = (1-\eta)\mathcal{X}_1 + \eta\mathcal{P}_{t_0}, \tag{7.6}$$

in which the sequence of parameters in $h \in (1, t_0)$, i.e., $(\mathcal{X}_1, \mathcal{X}_2, \cdots, \mathcal{X}_{t_0})$ are update by progressively following a sequence of convex combinations, whose strength is indexed by time:

$$\mathcal{X}_t = \left(1 - \frac{t}{t_0}\right)\mathcal{X}_1 + \frac{t}{t_0}\mathcal{X}_{t_0}. \tag{7.7}$$

In the general case, the k-th disruptive event will yield the following update:

$$\mathcal{X}_{t_{k-1}} = (1-\eta)\mathcal{X}_{t_{k-1}} + \eta\mathcal{P}_{t_k}, \tag{7.8}$$

$$\mathcal{X}_t = \left(1 - \frac{t}{t_{k-1}}\right)\mathcal{X}_{t_{k-1}} + \frac{t}{t_k}\mathcal{P}_{t_k}. \tag{7.9}$$

Any distribution over $\boldsymbol{\mu}_{t_k}$ can be used for generating $\mathcal{P}_{t_k}$. In this thesis, we use a uniform distribution over the hyperbox $[\mu_{\text{lb}}, \mu_{\text{ub}}]^N$, where μ_{lb} and μ_{ub} are the lower and upper bounds of return for all assets[1]. For generating *random covariance matrices*, we use a Gram Matrix-based method [117], which consists of uniformly sampling unit vectors in the unit hypersphere ($\mathcal{G}^{N-1}$), i.e., $\frac{\mathbf{x}}{||\mathbf{x}||} \sim U(\mathcal{G}^{N-1})$, with $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, and taking $\rho_{\mathcal{P}} = \mathbf{x}\mathbf{x}^\intercal$. The matrix is then built as $\boldsymbol{\Sigma}_{\mathcal{P}}(i,j) = \{\rho_{\mathcal{P}}(i,j)\sigma_i\sigma_j\}$, where σ_i is the standard deviation of the i-th asset, uniformly sampled from $[\sigma_{\text{lb}}, \sigma_{\text{ub}}]$.

7.2.2 Experimental Methodology

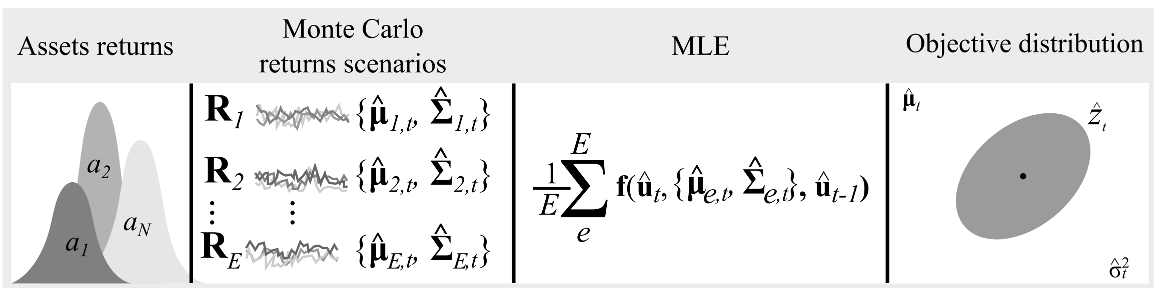


Figure 7.2: Estimating objective distributions by sampling from the joint assets return distribution given for each investment environment, from which E daily returns scenarios are sampled. The return sample mean vectors and covariance matrices are estimated for each scenario. Finally, a Maximum-Likelihood Estimation (MLE) is performed to fit a Gaussian distribution over the sampled risk (variance) and return (mean) of a given portfolio.

Let $\mathcal{W}_t$ be the initial wealth to be invested in a probabilistic market and set $t \leftarrow K$. Given a stream of T known Gaussian returns parameters $\{\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t\}_{t=1}^T$ (T is the number of periods), the following procedures are repeated in all ASMS variants at each period (see Fig. 7.2):

Step 1 Sample a total of $S \times K$ return vectors from each available $\mathcal{N}\left(\boldsymbol{\mu}_{t-K}, \boldsymbol{\Sigma}_{t-K}\right), \cdots, \mathcal{N}\left(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t\right)$, where E is the sample size;

Step 2 Estimate $\left\{\hat{\mathcal{X}}_{t-k}\right\}_{k=1}^K = \left\{\hat{\boldsymbol{\mu}}_{t-k}, \hat{\boldsymbol{\Sigma}}_{t-k}\right\}_{k=1}^K$ from the set of simulated returns;

Step 3 Use the parameter stream $\left\{\hat{\mathcal{X}}_{t-k}\right\}_{k=1}^K$ for solving AS-MOO at the t-th period to obtain $\hat{\mathcal{U}}_t^{N\star}$ and $\hat{\mathbf{u}}_t^\star$;

[1]The source code for the benchmark generator is publicly available at ...

Table 7.1: Brokerage fees utilized in the Brazilian stock market.

Traded value	Proportional Fee	Fixed Fee
< 135.07	0.0%	2.70
≥ 135.08 and < 498.62	2.0%	0.00
≥ 498.63 and $< 1,514.69$	1.5%	2.49
$\geq 1,514.70$ and $< 3,029.38$	1.0%	10.06
$\geq 3,029.39$	0.5%	25.21

Step 4 Use the out-of-sample future test state parameters[2] $\mathcal{X}_{t+1}$ to estimate the Stochastic Pareto Frontier (SPF) approximation

$$\hat{\mathcal{Z}}_{t+1} = \mathbf{f}\left(\hat{\mathcal{U}}_t^{N\star}, \mathcal{X}_{t+1}, \hat{\mathbf{m}}_{\mathbf{u}_t^\star}\right), \tag{7.10}$$

and collect the estimated sample average future Hypv:

$$\hat{\mathcal{S}}_{t+1} = \frac{1}{E}\sum_{e=1}^{E} \mathcal{S}\left(\hat{\mathcal{Z}}_{e,t+1}, \mathbf{z}_{\mathrm{ref}}\right); \tag{7.11}$$

Step 5 Implement the selected mean portfolio $\hat{\mathbf{m}}_{\mathbf{u}_t^\star}$ from the evolved SPF, advance to $t \leftarrow t+1$, and update the current investor wealth as:

$$\begin{aligned} \mathcal{W}_t \leftarrow \mathcal{W}_{t-1} + & \text{ mean observed return for } \hat{\mathbf{u}}_t^\star \\ - & \text{ transaction costs of transforming} \\ & \hat{\mathbf{m}}_{\mathbf{u}_{t-1}^\star} \text{ into } \hat{\mathbf{m}}_{\mathbf{u}_t^\star}. \end{aligned} \tag{7.12}$$

We set the initial wealth to $\mathcal{W}_0 = 10,000.00$ monetary units, and construct the cost function $\mathbf{h}$ following a suggestion by the Securities Commission of Brazil which is widespread, practiced by the major brokers operating in the São Paulo BOVESPA stock market. The brokerage costs depend on the value of trade, combining fixed and proportional fees (Tab. 7.1).

In order to assess the statistical significance of the anticipation and DM factors, we use a two-way Analysis of Variance (ANOVA). Pairwise significance is assessed with the non-parametric Mann–Whitney U test, which does not require a normality assumption.

Percentage Of Change In Direction, POCID

The POCID is a commonly used measure in time series forecasting applications [155] that captures the proportion of correct predictions w.r.t. the directions toward which the j-th objective function change between subsequent periods:

$$\mathrm{POCID}(j) = 100 \times \frac{1}{N \times (T-1)} \sum_{i=1}^{N}\sum_{t=1}^{T-1} D(j,i,t), \tag{7.13}$$

$$D(j,i,t) = \begin{cases} 1, & \text{if } (z_{j,t+1}^{(i)\star} - z_{j,t}^{(i)\star})(\hat{m}_{\hat{z}_{j,t+1}^{(i)\star}} - \hat{m}_{\hat{z}_{j,t}^{(i)\star}}) \geq 0, \\ 0, & \text{otherwise,} \end{cases}$$

[2]Those sets of parameters are only used for a post-hoc assessment of the obtained results and do not interfere in the optimization problem.

where T is the number of decision periods and N is the number of portfolio objective distributions in the population. In this thesis, we consider an overall POCID measure over risk and expected return of a fixed portfolio: POCID $= \frac{1}{2} \times$ (POCID(return) + POCID(risk)). In Eq. (7.13), $z^{(i)\star}_{j,t+1}$ and $z^{(i)\star}_{j,t}$ are two objective values observed for the i-th ranked portfolio in subsequent periods w.r.t. the j-th objective function ($j = 1$ for risk and $j = 2$ for mean return), whereas $\hat{m}_{\hat{z}^{(i)\star}_{j,t+1}}$ and $\hat{m}_{\hat{z}^{(i)\star}_{j,t}}$ are the corresponding predicted mean values obtained via KF tracking. Note that a 100% POCID means that the predictor can perfectly capture the changing patterns of direction on the quantity of interest.

Similarity in the Decision Space

In order to assess the dynamics over time of the portfolios compositions in the $(d-1)$–Simplex, we propose a *coherence* measure over the Pareto-flexible sets inspired in [106]. It consists of summing over the cosine similarities of each portfolio to the population centroid, $\mathbf{u}_t^C$:

$$\text{Coherence}\left(\hat{\mathcal{U}}_t^{N\star}\right) = \sum_{i=1}^{N} \frac{\hat{\mathbf{m}}_{\mathbf{u}_t^{(i)\star}} \cdot \mathbf{u}_t^C}{||\hat{\mathbf{m}}_{\mathbf{u}_t^{(i)\star}}||_2 ||\mathbf{u}_t^C||_2}, \tag{7.14}$$

where the operator $\cdot$ denotes the dot product, and

$$\mathbf{u}_t^C = \frac{1}{N} \sum_{i=1}^{N} \hat{\mathbf{m}}_{\mathbf{u}_t^{(i)\star}}. \tag{7.15}$$

In this way, the overall similarity of a set of N candidate mean DD portfolios can be assessed. Note that normalization is needed because the centroid will not necessarily have unit norm.

7.2.3 Parameters Settings and Datasets

We now describe the data collection and experimental setup.

Artificial and Real-World Datasets

A total of 10 investment scenarios are assessed: (a) seven artificial; and (b) three real-world stock markets: London FTSE 100, Dow Jones Index (DJI), and Hong Kong's Hang Seng Index (HSI). Each scenario encode different combinations of *severity* and *periodicity* of disruptive change. All benchmarks provide $d = 30$ simulated assets for composing the portfolios, whereas for the real-world instances we have $d = 87$ for FTSE; $d = 30$ for DJI; and $d = 49$ for HSI[3], which are represented in a $(d-1)$–Simplex space.

The total number of periods[4] for all instances is $T = 25$. The real-world scenarios are composed of daily adjusted close prices collected between $20/11/2006 - 31/12/2012$, from which 50 days lagged returns were computed: $r_{n,k} = \frac{\mathcal{V}_{n,k+50} - \mathcal{V}_{n,k}}{\mathcal{V}_{n,k+50}}$, where $\mathcal{V}_{n,k+50}$ is the value of the n-th asset at day $k+50$. Each investment period comprises one and a half year worth of daily returns data. Between each period, the 50 oldest lagged returns are discarded and the 50 latest ones are

[3]The data used in our experiments is publicly available at ...

[4]Except for FTSE with $T = 24$ due to data availability issues at the time the experiments were executed.

included in the sample from which the parameters $\left\{\hat{\boldsymbol{\mu}}_t, \hat{\boldsymbol{\Sigma}}_t\right\}_{t=1}^{T}$ are estimated, where 50 business days roughly corresponds to two and a half months. Thus, the period $t = 1$ in FTSE comprises data from 20/11/2006 – 20/05/2008, $t = 2$ to 01/02/2007 – 30/07/2008, and so forth.

As for the artificial benchmarks, we investigate severities of $\eta \in \{0.5, 1.0\}$ and periodicities of disruptive change of $\tau \in \{2, 4, 8\}$. We refer to the six instances as PO(η,τ). The seventh instance was generated to simulate an environment in which the covariance $\boldsymbol{\Sigma}_t$ varies at every four periods, but that is *static* only w.r.t. the *mean returns* $\boldsymbol{\mu}_t$. We refer to this instance as the "static mean PO", despite the dynamic volatility of each asset.

Bayesian Tracking Parameters

Tracking with KF and the DD MAP requires setting (a) the covariance matrices representing the initial uncertainty and the measurement noise, for the KF; and (b) the scale factor s representing the initial dispersion, for the DD tracking. In the former case, the initial covariance was set directly from the data, using the sample covariance of a given objective distribution evaluated at $t = 0$, whereas the noise covariance was set to $\frac{1}{E} \times \Sigma_t$, which corresponds to the covariances of the sample means. For each simulation scenario involving the stochastic evaluation of the objective-functions (i.e., average return and average risk), $E = 1000$, i.e., the risk and return of each candidate portfolio is evaluated from 1000 simulated daily return values. The scale factor in DD tracking is set to $s = 1$, after finding that its influence is negligible in preliminary tests.

Online Anticipatory Learning (OAL) was used to self-adjust the anticipation rates using Eqs. (6.6) and (6.9) at **chapter 6**. Hence, we have set

$$\lambda_{t+h} = \frac{1}{2}\left(\lambda_{t+h}^{(\mathcal{H})} + \lambda_{t+h}^{(\mathcal{K})}\right), \tag{7.16}$$

for which the anticipation horizon is $H = 2$ (one-step-ahead prediction). The anticipation rate of each portfolio is thus determined not only by the estimated temporal incomparability probability between the current and the predictive objective distribution ($\lambda_{t+h}^{(\mathcal{H})}$), but also by the observed normalized sum of KF squared residuals ($\lambda_{t+h}^{(\mathcal{K})}$).

The motivation for taking the average (the $\frac{1}{2}$ factor) of the two aforementioned self-adjustment strategies for setting λ_{t+h} can be explained by the intuition of providing a *balanced tension* in the OAL rule between:

- The desire of *trusting* in decision paths that *were shown* to lead to higher predictable consequences *in the past* ($\lambda_{t+h}^{(\mathcal{K})}$ in Eq (6.9)); and
- The desire of *trusting* in decision paths that *are estimated* to lead to higher predictable consequences *in the future* ($\lambda_{t+h}^{(\mathcal{H})}$ in Eq (6.6)).

Constraint Handling

We considered minimum and maximum cardinality of $c_l = 5$ and $c_u = 15$ assets. Moreover, the feasible objective space includes all portfolios attaining mean return of at least 0% and at most 20% risk. The feasibility probability was set to 99% ($\epsilon = 0.99$, see Def. 5.5). For two anticipatory distributions $\hat{\mathbf{z}}_t^{(a)}$ and $\hat{\mathbf{z}}_t^{(b)}$, constraint handling in ASMS is done similarly as in [69]:

- If both $\hat{\mathbf{z}}_t^{(a)}$ and $\hat{\mathbf{z}}_t^{(b)}$ are ϵ-feasible, then PD is applied over the mean vectors as in Def. 5.3;
- Else if $p(\mathbf{z}_t^{(a)})_\epsilon = \Pr\left\{\hat{\mathbf{z}}_t^{(a)} \in \mathcal{Z}_t\right\} \geq \epsilon$ but $p(\mathbf{z}_t^{(b)})_\epsilon = \Pr\left\{\hat{\mathbf{z}}_t^{(b)} \in \mathcal{Z}_t\right\} < \epsilon$, then $\hat{\mathbf{z}}_t^{(a)} \preceq \hat{\mathbf{z}}_t^{(b)}$ (and vice-versa);
- Else, if both are not ϵ-feasible, then PD is applied to compare the vectors $\left(p(\hat{\mathbf{z}}_t^{(a)})_{1,\epsilon}, p(\hat{\mathbf{z}}_t^{(a)})_{2,\epsilon}\right)^\intercal$ and $\left(p(\hat{\mathbf{z}}_t^{(b)})_{1,\epsilon}, p(\hat{\mathbf{z}}_t^{(b)})_{2,\epsilon}\right)^\intercal$, where $p(\hat{\mathbf{z}})_{j,\epsilon}$ is the marginal ϵ-feasibility probability w.r.t. the j-th objective function.

ASMS Parameters

Population size in all variants was set to $N = 20$ and 30 generations were executed for each investment period. Seeding was utilized so that the previous portfolios served as the new starting points in subsequent periods. Mutation rate was set to 0.3; crossover probability to 0.2; and binary tournament selection was utilized using the Pareto Dominance over the sample mean vectors, with expected Hypv contribution (Eq. (6.35)) serving as a tiebreaker. Finally, the reference point for computing Hypv was set to $\mathbf{z}^{\text{ref}} = (0.2, 0.0)^\intercal$ in terms of risk and return, coinciding with the objective space feasibility boundaries (maximum risk of 20% and minimum mean return of 0%).

Search Operators

We utilized uniform crossover over the mean DD vectors. For mutation, we randomly choose between (1) modifying the non-zero weights; or (2) adding/removing assets. If operator (1) is selected, then, with probability 1/2, we either increase or decrease the investment on a randomly chosen asset by a uniformly drawn factor from 10 to 50%. If (2) is selected, then, with probability 1/2, we either add or remove a randomly chosen asset. If it is removed, we simply set its weight to zero. If it is added, we randomly set its weight within a $\pm 10\%$ range from an equally-balanced allocation. All modifyied DD vectors are renormalized.

7.3 Results and Discussion

All experiments consider AS-MOO instances for which Pareto-efficient solutions are evolved according to a stream of historical 50-lagged daily return data. The implemented solution affects the objective function evaluations at subsequent periods, thus modifying the admissible future Pareto-flexible sets.

The differences in the estimated anticipatory distributions within the evolved sets of Pareto-flexible portfolios are illustrated in Fig. 7.3 (a)–(d), for four evolved SPFs in one run of ASMS/mHDM at four arbitrarily chosen periods of the PO(8,1.0) instance. Note that high return portfolios were associated with high risk, and that there may exist discontinuities in the theoretical SPFs, what is expected when cardinality constraints are considered (see [83]).

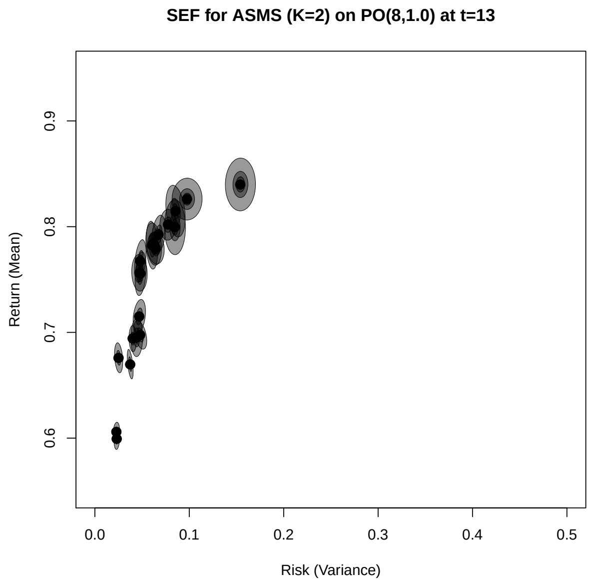
Moreover, the spreads of the anticipatory distributions are larger for portfolios attaining higher risk/return when compared to the lower risk/return ones. This result is consistent with the intuition that risky portfolios tend to be less stable over time [15] (see **chapter 4**). The KF tracking seems to have correctly captured this uncertainty. Finally, the rotation of the bivariate Gaussian objective distributions varied with the position along the evolved SPFs. This result suggests that there might be different correlation patterns between risk minimization and return maximization, depending on the composition of each evolved portfolio.

7.3.1 Estimated Confidence Over the Stochastic Pareto Frontiers (SPFs)

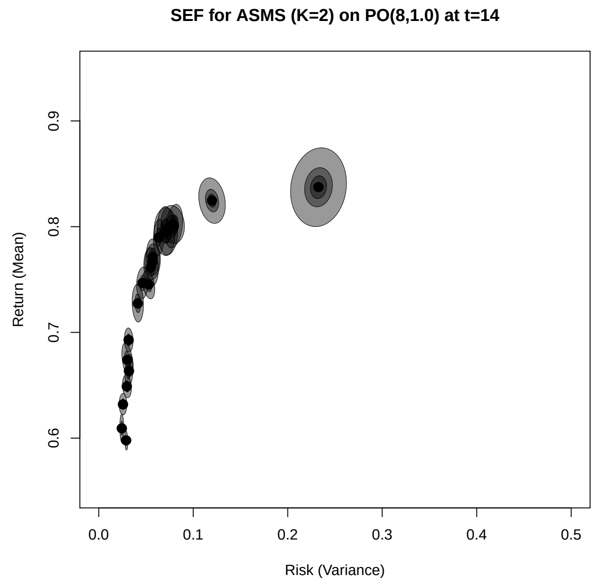
We recall that, according to the OAL rules proposed in **chapter 6**, the available KF predictive objective distributions of portfolios associated with higher predictive confidence are more intensely incorporated into the resulting anticipatory distributions, which are effectively considered for guiding the search towards the anticipatory SPF. Moreover, recall that rank 1 portfolios correspond to the highest risk/return, whereas a rank 20 one corresponds to the lowest risk/return in the population.

The barplots shown in Fig. 7.4 (a)–(f) for FTSE-100 depict the predictability degrees (or confidence, which is the complement of the binary entropy of the TIP) for the evolved portfolios, expressed in terms of Eq. (6.6) in **chapter 6**. For the smaller window sizes of $K = 1$ and $K = 2$, the confidence distribution was slightly multimodal, with peaks around portfolio ranks of 5 and 15, whereas for the largest window size ($K = 3$), a decreasing pattern can be observed, where higher ranked portfolios (higher return ones) tended to be associated with higher predictability in terms of the Pareto dominance relation between current and predicted performance. The same pattern for $K = 3$ was also observed for the HSI market (Fig. 7.6 (c)–(d)).

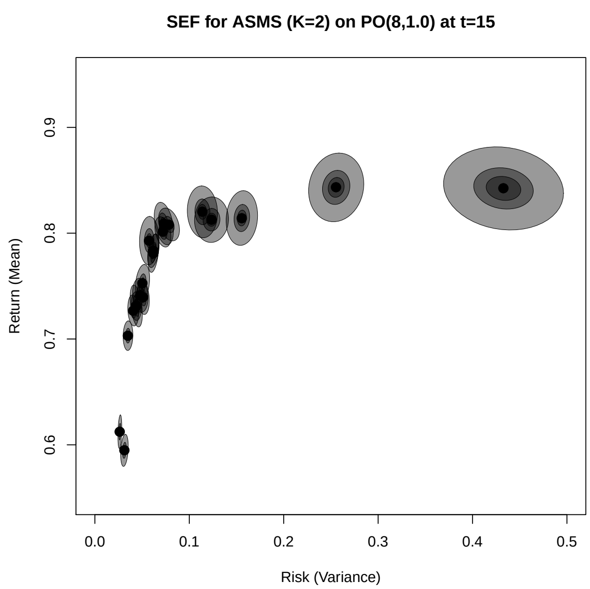
Different yet orderly confidence distributions were observed for the remaining instances. For instance, the reverse tendency was observed for the DJI market with $K = 3$ (Fig. 7.5 (c)–(d)), when compared to that observed for the FTSE-100 market, where lower ranked portfolios (less risky ones) were associated with higher predictive confidence. The emergence of such



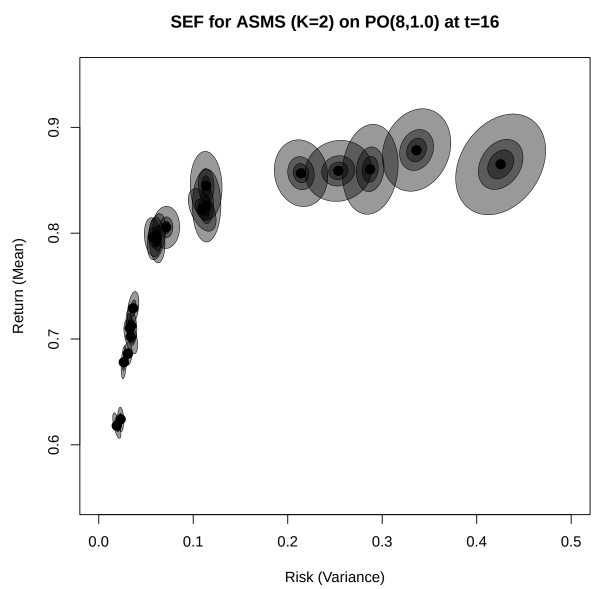
(a) SPF approximation at t=13



(b) SPF approximation at t=14



(c) SPF approximation at t=15

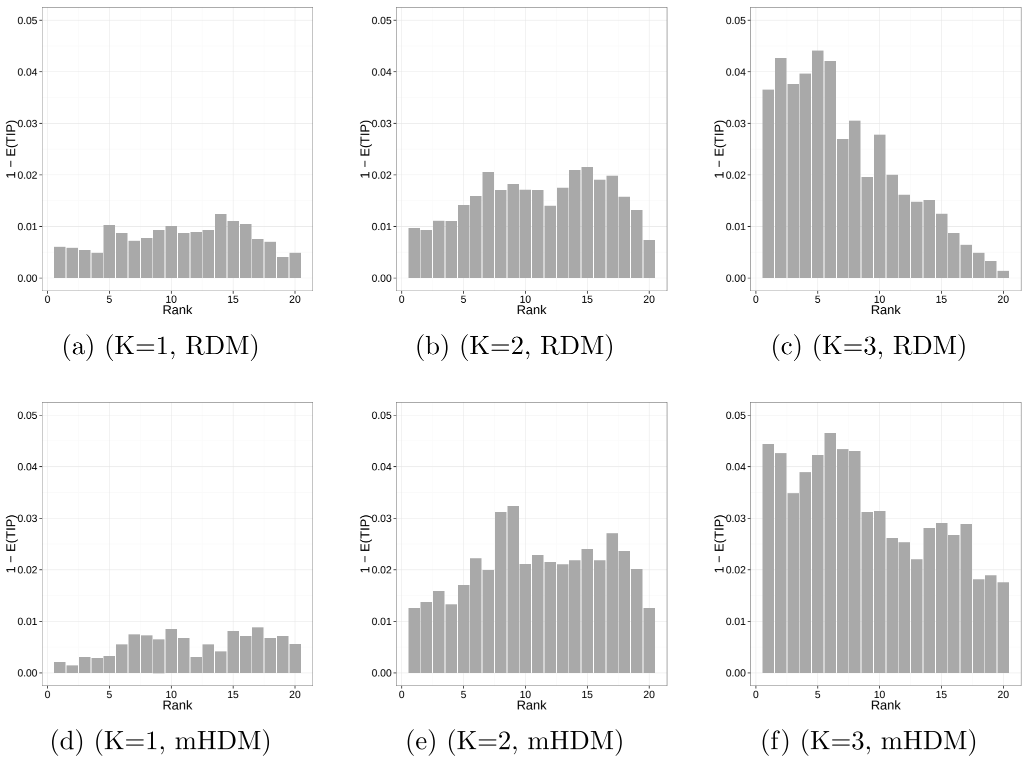


(d) SPF approximation at t=16

Figure 7.3: Example of anticipatory SPFs evolved with ASMS/mHDM.

tendencies suggests that OAL effectively self-adapted the anticipation rates, according to the market dynamics and the chosen window size.

A general observation from the predictive confidence distributions for the remaining instances (Figs. 7.4–7.13) is that, although the patterns were very similar, the mHDM variants (parts (d)–(f) of those figures) have lead to slightly more conservative SPF approximations, when compared to the RDM ones (parts (a)–(b)), in terms of the extent to which predictive knowledge was incorporated into the evolved anticipatory distributions. This result may indicate a potential

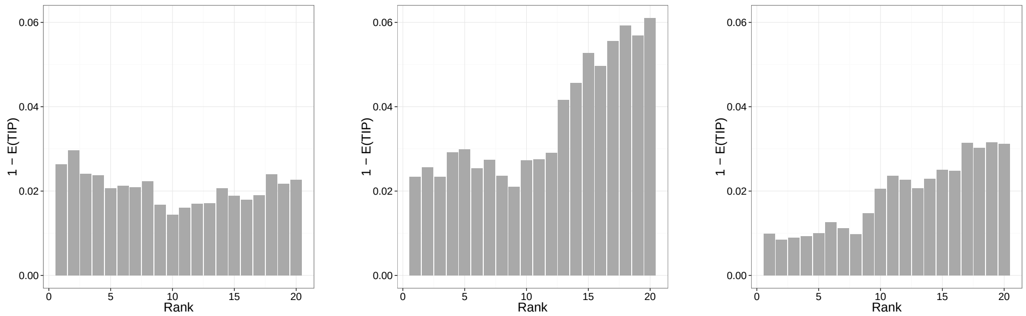


(a) (K=1, RDM) (b) (K=2, RDM) (c) (K=3, RDM)

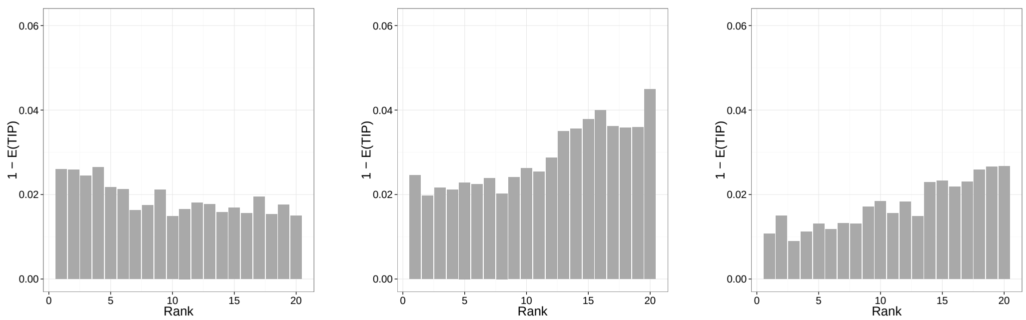
(d) (K=1, mHDM) (e) (K=2, mHDM) (f) (K=3, mHDM)

Figure 7.4: Confidence distributions over the SPFs, averaged over all periods for FTSE-100.

trade-off between expected future Hypv maximization and predictive confidence.

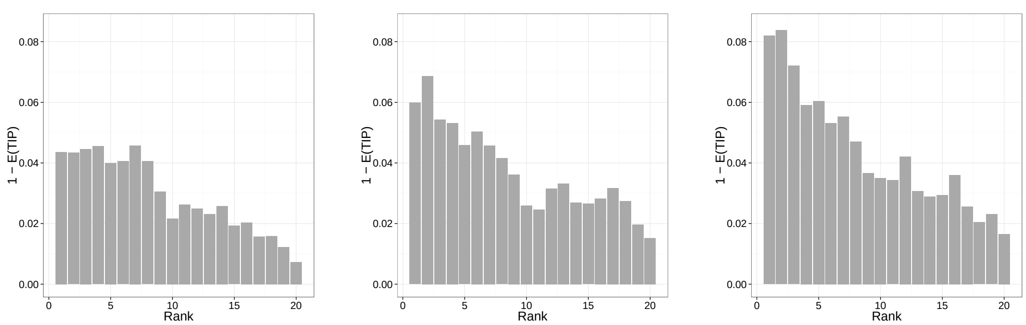


(a) (K=1, RDM) (b) (K=2, RDM) (c) (K=3, RDM)

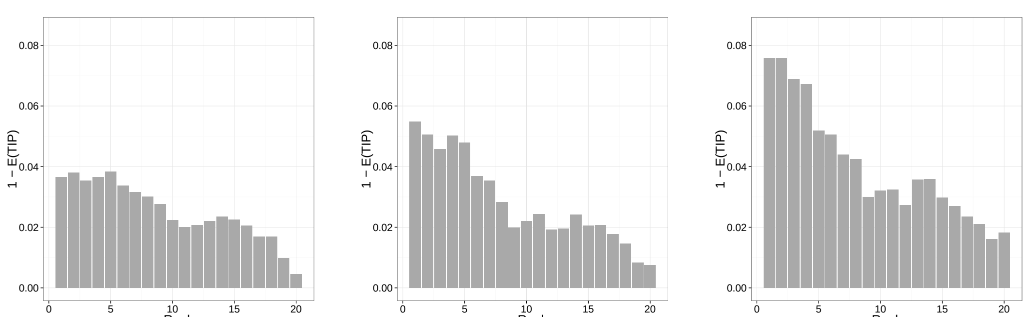


(d) (K=1, mHDM) (e) (K=2, mHDM) (f) (K=3, mHDM)

Figure 7.5: Confidence distributions over the SPFs, averaged over all periods for DJI.

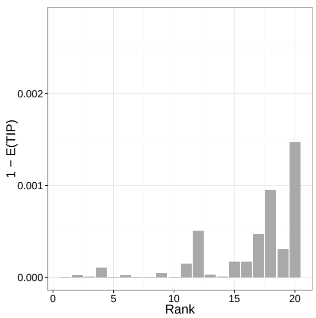


(a) (K=1, RDM) (b) (K=2, RDM) (c) (K=3, RDM)

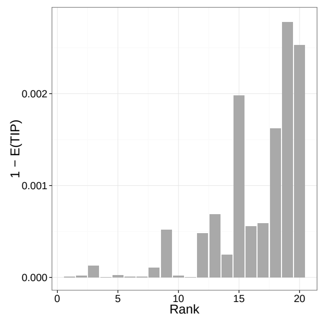


(d) (K=1, mHDM) (e) (K=2, mHDM) (f) (K=3, mHDM)

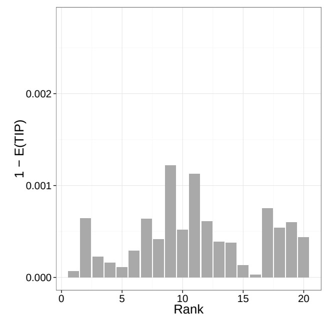
Figure 7.6: Confidence distributions over the SPFs, averaged over all periods for HSI.



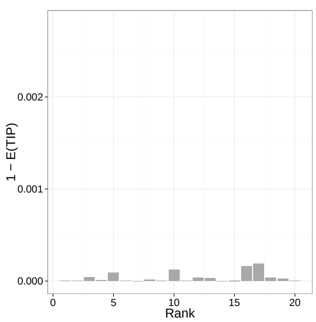
(a) (K=1, RDM)



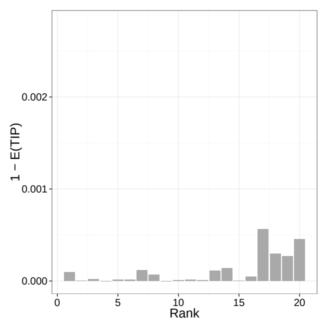
(b) (K=2, RDM)



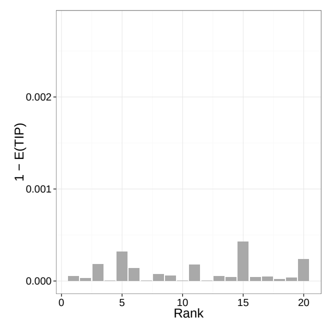
(c) (K=3, RDM)



(d) (K=1, mHDM)

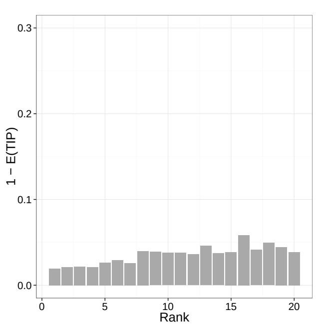


(e) (K=2, mHDM)

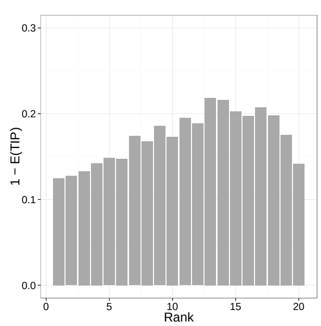


(f) (K=3, mHDM)

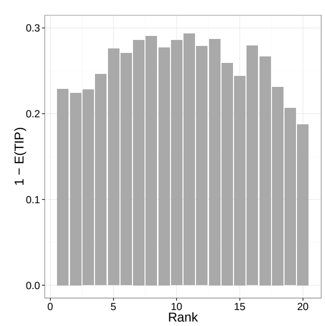
Figure 7.7: Confidence distributions over the SPFs, averaged over all periods for static mean PO.



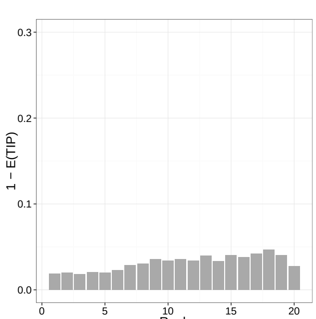
(a) (K=1, RDM)



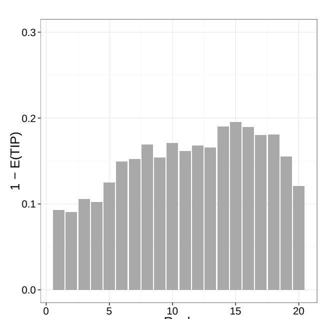
(b) (K=2, RDM)



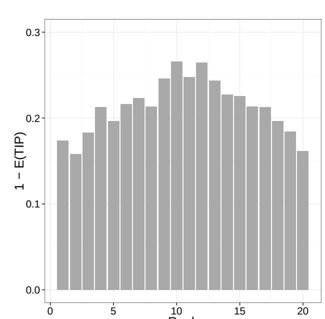
(c) (K=3, RDM)



(d) (K=1, mHDM)

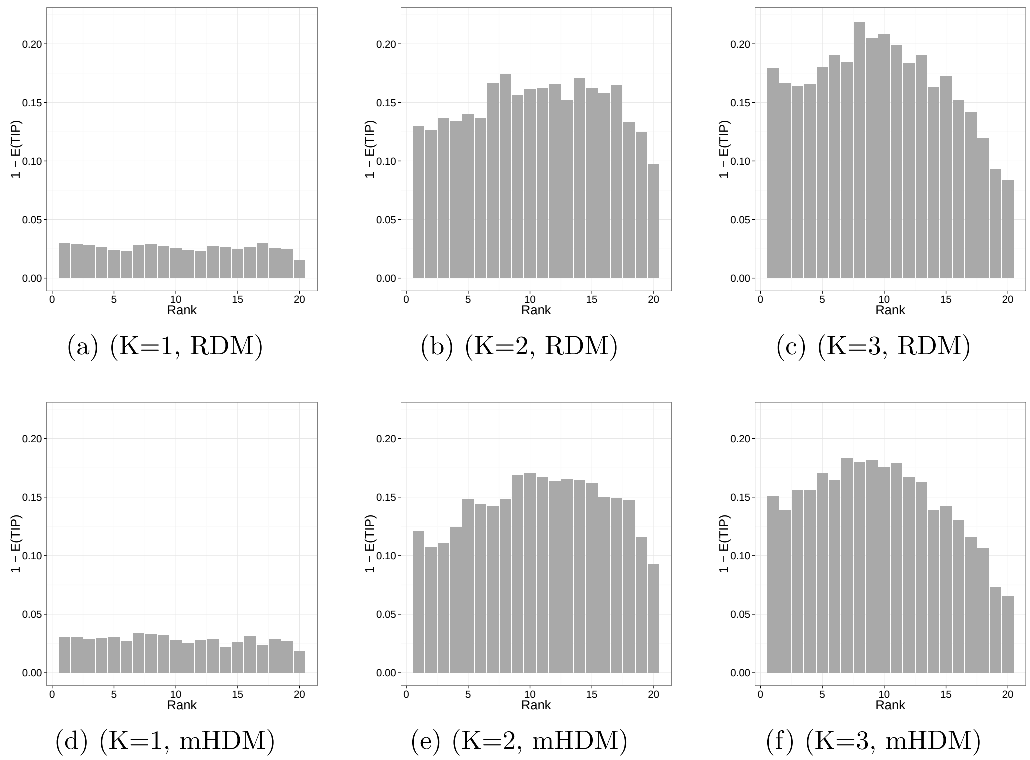


(e) (K=2, mHDM)



(f) (K=3, mHDM)

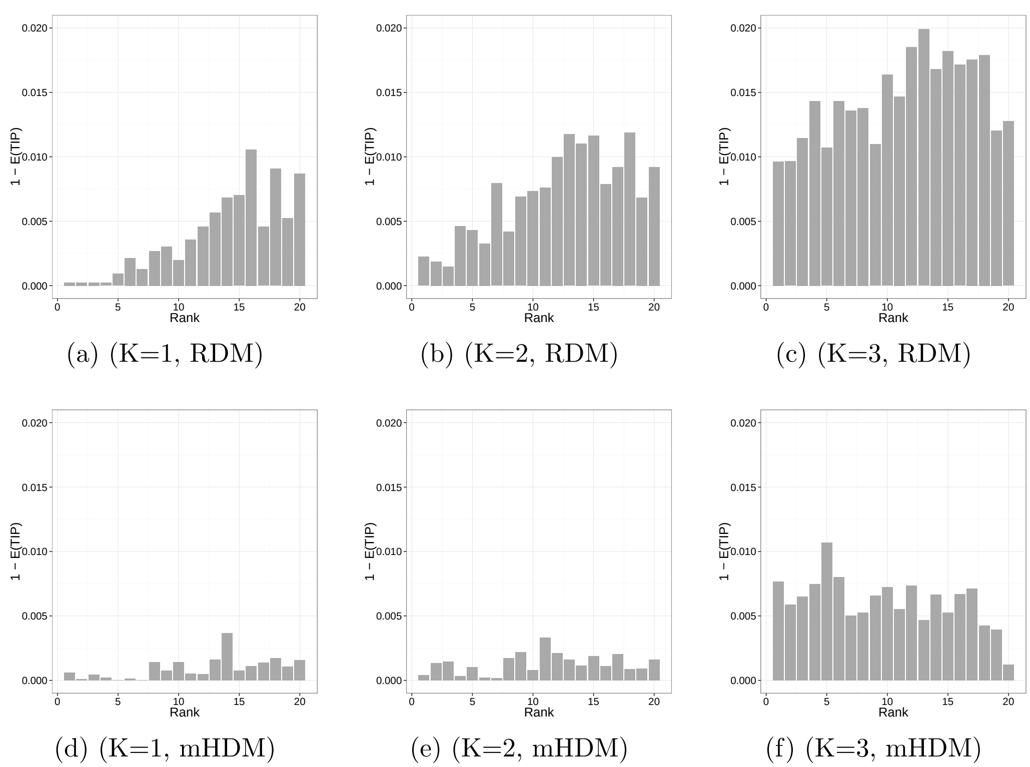
Figure 7.8: Confidence distributions over the SPFs, averaged over all periods for PO(2,0.5).



(a) (K=1, RDM) (b) (K=2, RDM) (c) (K=3, RDM)

(d) (K=1, mHDM) (e) (K=2, mHDM) (f) (K=3, mHDM)

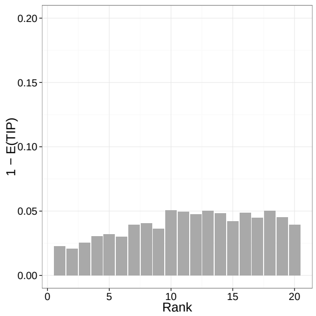
Figure 7.9: Confidence distributions over the SPFs, averaged over all periods for PO(2,1.0).



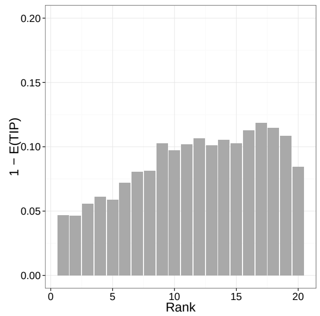
(a) (K=1, RDM) (b) (K=2, RDM) (c) (K=3, RDM)

(d) (K=1, mHDM) (e) (K=2, mHDM) (f) (K=3, mHDM)

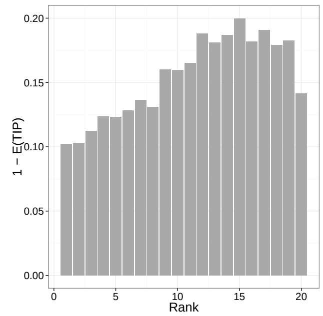
Figure 7.10: Confidence distributions over the SPFs, averaged over all periods for PO(4,0.5).



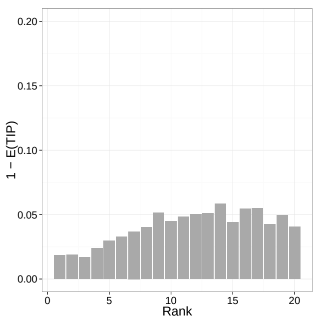
(a) (K=1, RDM)



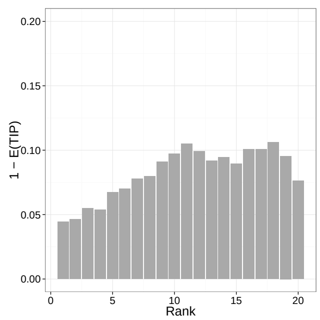
(b) (K=2, RDM)



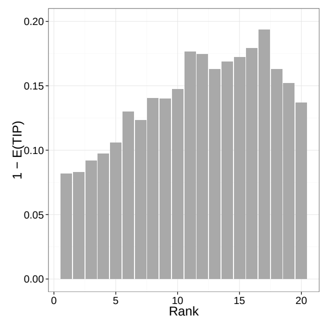
(c) (K=3, RDM)



(d) (K=1, mHDM)

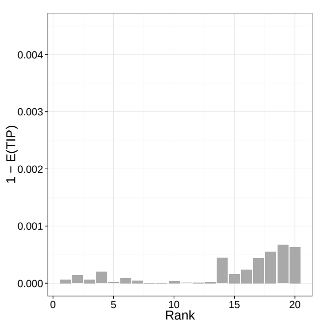


(e) (K=2, mHDM)

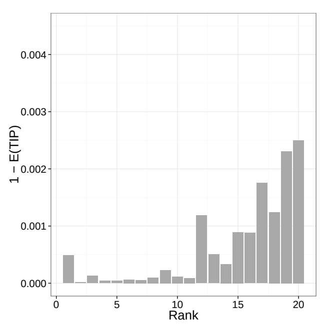


(f) (K=3, mHDM)

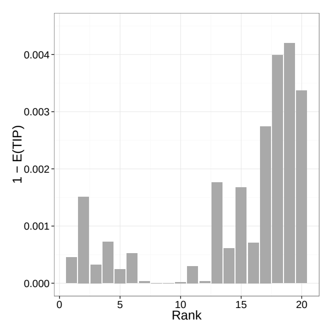
Figure 7.11: Confidence distributions over the SPFs, averaged over all periods for PO(4,1.0).



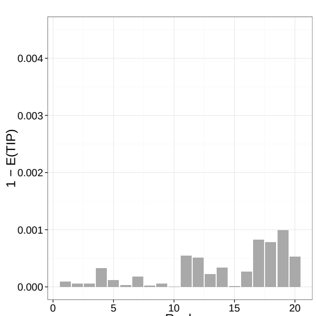
(a) (K=1, RDM)



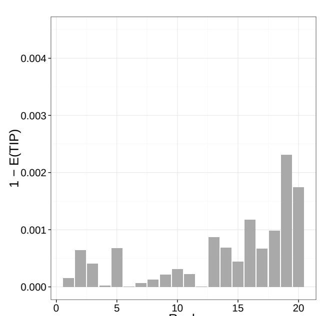
(b) (K=2, RDM)



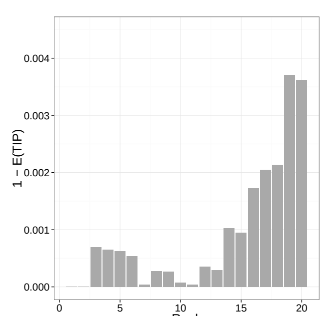
(c) (K=3, RDM)



(d) (K=1, mHDM)

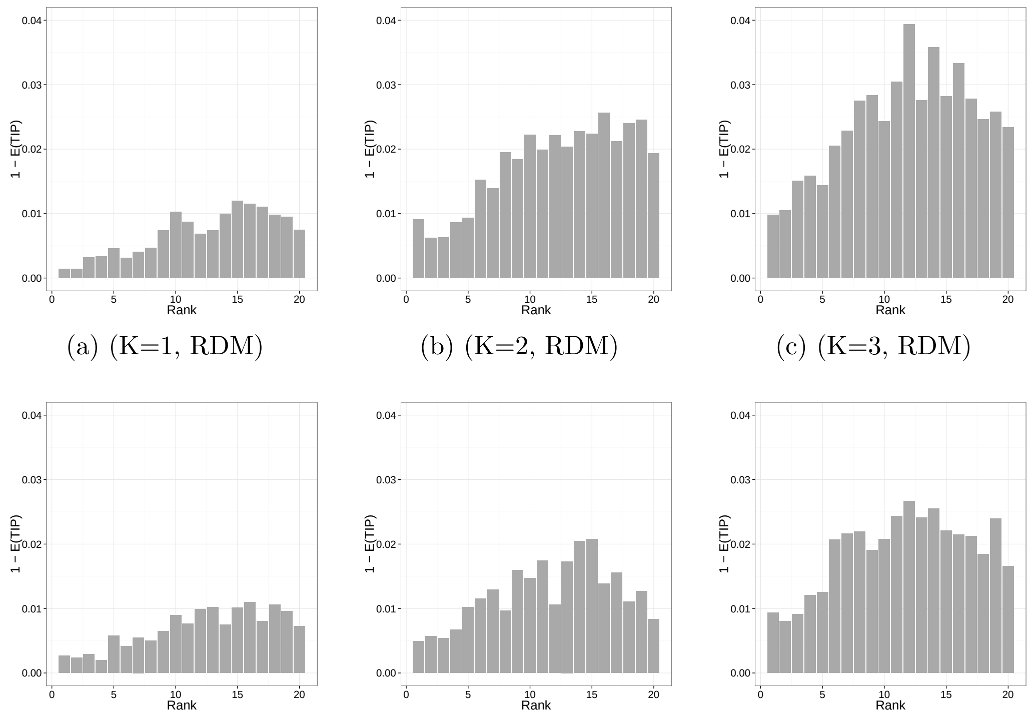


(e) (K=2, mHDM)



(f) (K=3, mHDM)

Figure 7.12: Confidence distributions over the SPFs, averaged over all periods for PO(8,0.5).



(a) (K=1, RDM) (b) (K=2, RDM) (c) (K=3, RDM)

(d) (K=1, mHDM) (e) (K=2, mHDM) (f) (K=3, mHDM)

Figure 7.13: Confidence distributions over the SPFs, averaged over all periods for PO(8,1.0).

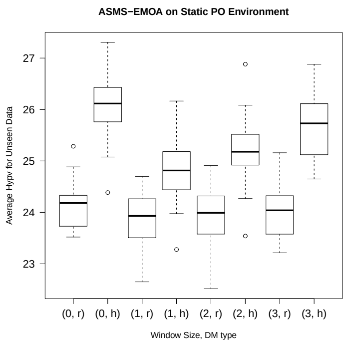
7.3.2 Bayesian Tracking Influence on Percentage On Change In Direction

The effectiveness of Bayesian tracking in ASMS can be analyzed in terms of POCID (Eq. (7.13)). The results for the PO benchmarks are shown in Fig. 7.14 (a)–(g). As expected for the static mean returns scenario with varying volatility (see Sec. 7.2.3) the best median POCID results ($\approx 85\%$) were obtained with the myopic SMS ($K = 0$, see Fig. 7.14 (a)). It is likely that, for $K > 0$, tracking may have been severely affected by the sampling noise. As for the nonstationary mean returns instances, prediction was more effective when disruptive changes occurred at low frequency (every $\tau = 8$ periods), with moderate severity ($\eta = 0.5$). The highest median POCID in this case ($\approx 72\%$) was observed with the largest window size, $K = 3$ (see Fig. 7.14 (f)), whereas, for the faster-paced environments ($\tau = 2, 4$), the shortest window size ($K = 1$) allowed for the best performance.

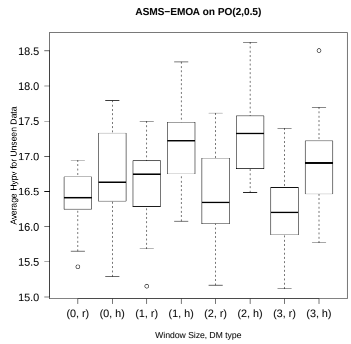
The performance decrease in average POCID observed for large window sizes in faster-paced artificial PO environments was expected. In those cases, outdated historical data is kept by the KF and the DD MAP tracking tools, even though the joint assets return distribution have already abruptly changed. The inability to detect such disruptive changes leaves the estimation of the portfolios objective distribution velocities subject to significant inaccuracies, what end up degrading the ability of the Bayesian tracking tools to correctly identify the novel directions of change. This latter result is consistent with the intuition that fast changing distributions require responsive tracking capable of quickly discarding uncorrelated past data.

Furthermore, the results for the real-world datasets shown in Fig. 7.14 (h)–(k) reveal that tracking was significantly more effective with $K = 2$, for FTSE, and with $K = 1$ for DJI and HSI, suggesting that the 50 days lagged returns distribution dynamics of those markets may be better approximated with more responsiveness tracking. We point out that the correct identification of a proper dynamical model can significantly affect the predictive ability of the implemented tracking tools. Although the system identification step was left to the user in this thesis, i.e., although no effort has been made to automatically learn a dynamical model, the relatively high average POCID results obtained for certain window size values (K) in all artificial and real-world PO instances indicate that the Bayesian tracking tools utilized in our experiments are adequate for the investigated portfolio selection application.

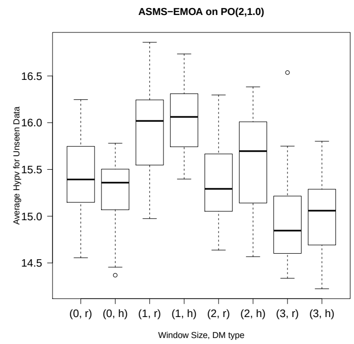
Finally, we observe that the stationary assumption (for $K = 0$) was significantly outperformed in terms of POCID for at least one value of $K > 0$ in all 9 nonstationary instances, including the real-world markets, what suggests that OAL has succeeded on learning portfolios attaining good predictability, as captured by POCID. Another noteworthy effect in POCID was that of the DM factor for PO(4,0.5), PO(8,0.5), and DJI, what suggests that rebalancing the implemented portfolios according to the mHDM strategy may lead to more predictable Pareto-flexible sets, when compared to the sets of portfolios evolved after randomly selecting Pareto-incomparable investment decisions with the RDM strategy. Moreover, the observed similarity of the POCID patterns obtained for the real-world DJI and HSI datasets to those observed for the artificial PO(4,0.5) suggests that the changing dynamics of those markets may have been roughly reproduced with our benchmark generator.



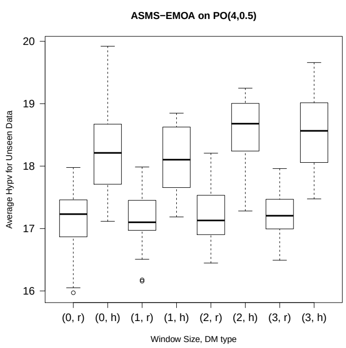
(a) Static mean PO.



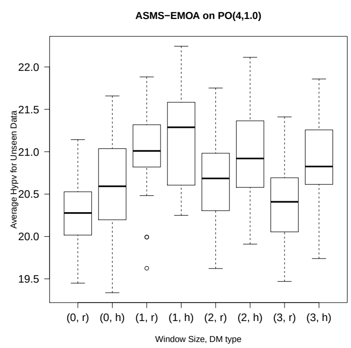
(b) PO(2,0.5)



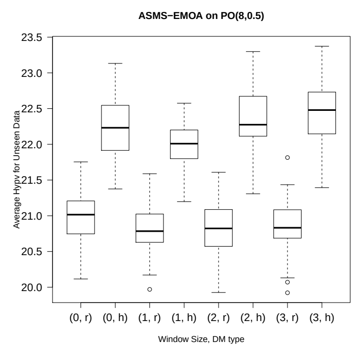
(c) PO(2,1.0)



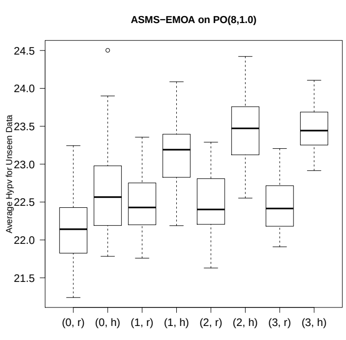
(d) PO(4,0.5)



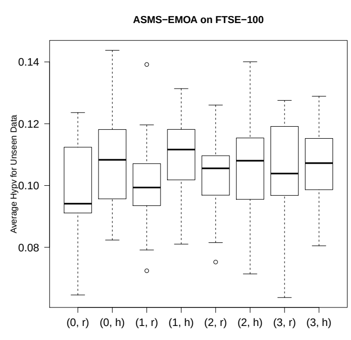
(e) PO(4,1.0)



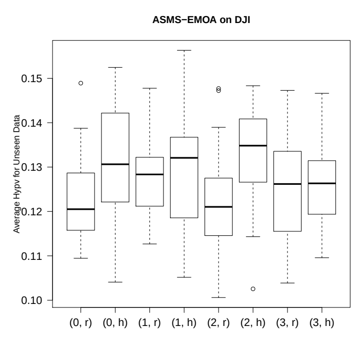
(f) PO(8,0.5)



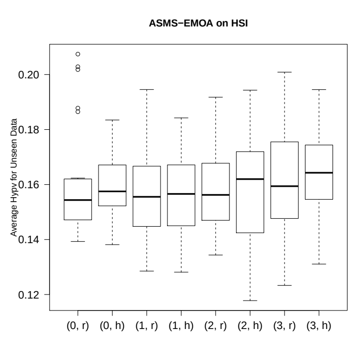
(g) PO(8,1.0)



(h) FTSE



(i) DJI



(j) HSI

Figure 7.15: Boxplots of Out-Of-Sample Future Average Hypervolume. In the legends, h stands for mHDM and r for RDM within the notation (K, DM).

7.3.3 Effects on the Out-Of-Sample Future Average Hypervolume

In this section, we draw attention to the observation in **chapter 3**, section 3.1.4 that maximizing future Hypv for upcoming periods is consistent with axiomatic notions of preference for flexibility in decision theory and is related to intuitive notions of future diversity of choice maximization. For achieving this goal, we argued in **chapter 5** that taking the Estimated Maximal Flexible Choice is what enables the attainment of maximal future expected Hypv Pareto-flexible sets in the AS-MOO model.

Under that interpretation, we assessed whether the proposed methodology has succeed in attaining its *main goal* (see Fig. 1.2 in **chapter 1**): handling undefined preferences by achieving better future sets of options. The *figure of merit* for this investigation is the out-of-sample future average Hypv ($\hat{\mathcal{S}}_{t+1}$, see Eq. (7.11)) computed from the obtained Pareto-flexible sets at each period. The assessed hypothesis is that both the anticipatory and the maximal Hypv DM variants achieve higher $\hat{\mathcal{S}}_{t+1}$ values than those achieved by the myopic and the RDM ones.

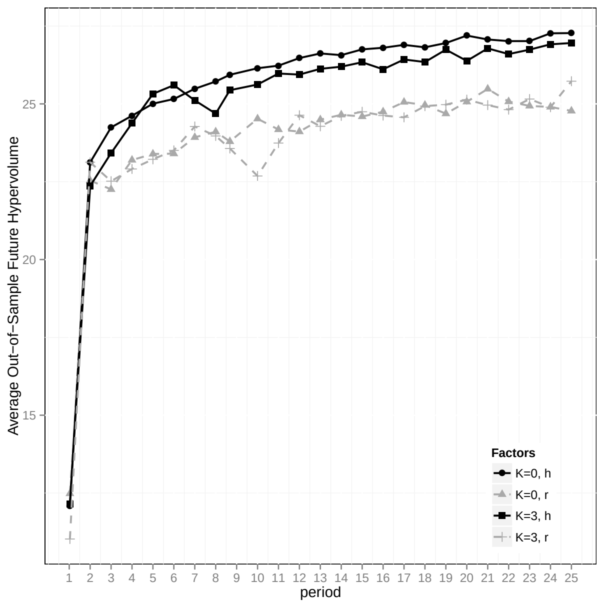
Table 6.2 shows the two-way ANOVA results for the anticipation (WS) and DM effects on $\hat{\mathcal{S}}_{t+1}$, besides their interaction described in the "WS:DM" rows. The *DM factor* had a statistically significant effect on future Hypv in six out of seven artificial scenarios (85%). The exception was the PO(2,1.0) instance, for which disruptive changes happen with the highest frequency and severity within the test suite. Inspecting the $\hat{\mathcal{S}}_{t+1}$ boxplot results in Fig. 7.15, the superiority of the mHDM over the RDM becomes clear, for the artificial scenarios. Moreover, from the real-world ANOVA results, the DM factor was deemed statistically significant on the DJI dataset. Besides, the boxplots results in Fig. 7.15 reveal that the mHDM also achieved significantly higher median $\hat{\mathcal{S}}_{t+1}$ values than those achieved by the RDM on 4 out of 12 real-world test cases, considering FTSE, DJI, and HSI and the four window size levels, being statistically equivalent to RDM for the remaining 8 cases.

As for the *anticipation factor*, significant effects on future Hypv were detected for five out of seven artificial scenarios (71%). The exceptions were the PO(8,0.5) and PO(4,0.5), although the effects would have been significant at the 10% level. In addition, Fig. 7.15 reveals that in 15 out of 20 scenarios (70%) – considering *all* 10 instances combined with the two DM types – *at least one* anticipatory variant significantly outperformed its myopic counterpart, for some values of $K > 0$. For the remaining five scenarios, the differences in performed were not significant, although we recall that those results take into consideration all the experimental data collected over all 25 investment periods, i.e., there might be some periods in those five scenarios for which significant differences might have been obtained, what can be visually hinted by the $\hat{\mathcal{S}}_{t+1}$ time series depicted in part (a) of Figs. 7.16–7.25. For those figures, the myopic variants SMS/RDM and SMS/mHDM are compared with the best window size value for the anticipatory variants ASMS/RDM and ASMS/mHDM, identifyied from the $\hat{\mathcal{S}}_{t+1}$ boxplot results in Fig. 7.15.

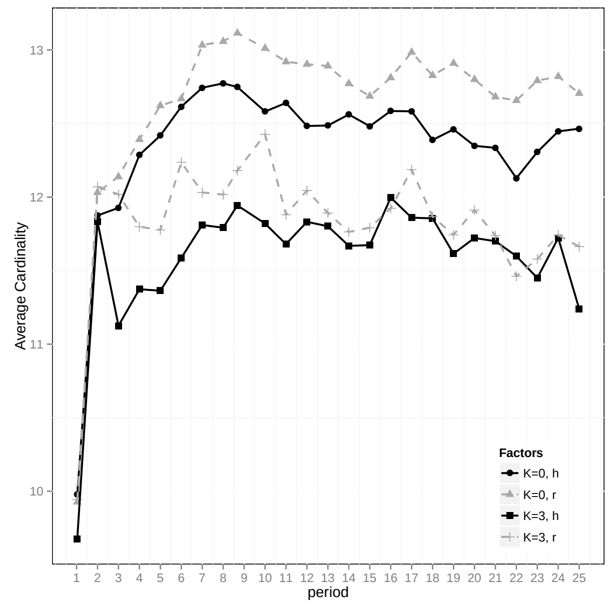
Also, ANOVA revealed significant *interaction* between the anticipation and the DM factors in the static mean PO, PO(2,0.5), and PO(8,1.0) instances, where the effects of window size *depended* on the DM factor: for mHDM, varying K had significant effects on $\hat{\mathcal{S}}_{t+1}$, whereas the same cannot be said for RDM. In the remaining 7 cases (see Fig. 7.15), varying K has lead to significant effects on $\hat{\mathcal{S}}_{t+1}$ in both mHDM and RDM variants, *independent* of the decision-making strategy. This latter result suggests that, in some cases, when incorporating predictive knowledge into the estimated anticipatory distributions, the improvements in future Hypv can be only achieved when jointly implementing the predicted maximal flexible decisions, although this was generally not the case for what improved performance by using prediction was also obtained even when implementing randomly selected Pareto-incomparable portfolios.

Table 7.2: ANOVA for WS and DM factors on $\hat{\mathcal{S}}_{t+1}$ ($^{**}p < 0.01$; $^{*}p < 0.05$).

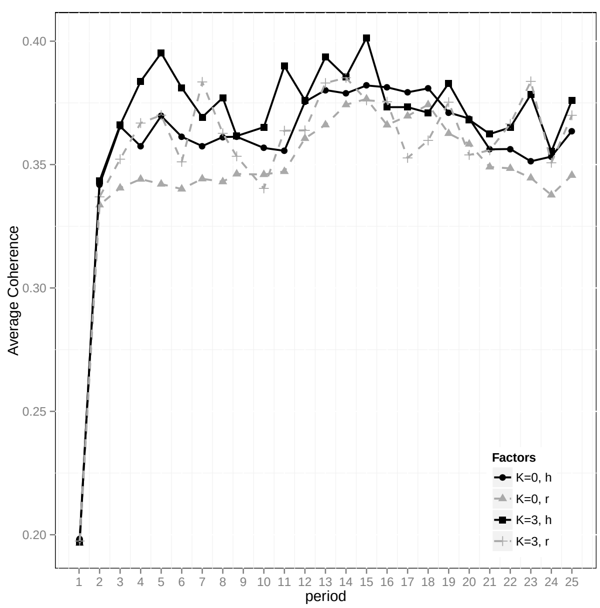
AS-MOO Instance	Factor	Mean Sq	F value	Significance
Static mean PO	WS	143.04	15.15	**
	DM	2504.54	265.26	**
	WS:DM	65.58	6.95	**
PO(2, 0.5)	WS	42.67	6.53	**
	DM	500.83	76.67	**
	WS:DM	18.28	2.80	*
PO(2, 1.0)	WS	354.94	33.50	**
	DM	5.70	0.54	≈
	WS:DM	9.06	0.86	≈
PO(4, 0.5)	WS	21.74	2.42	*
	DM	1626.40	180.95	**
	WS:DM	11.27	1.25	≈
PO(4, 1.0)	WS	79.58	13.19	**
	DM	92.88	15.40	**
	WS:DM	7.52	1.25	≈
PO(8, 0.5)	WS	17.26	2.26	≈
	DM	2268.88	297.36	**
	WS:DM	10.91	1.43	≈
PO(8, 1.0)	WS	82.76	23.22	**
	DM	757.69	212.56	**
	WS:DM	16.85	4.73	**
DJI	WS	0.00	0.19	≈
	DM	0.03	2.21	≈
	WS:DM	0.01	0.66	≈
FTSE	WS	0.00	0.58	≈
	DM	0.03	9.80	**
	WS:DM	0.00	1.46	≈
HSI	WS	0.01	0.34	≈
	DM	0.00	0.12	≈
	WS:DM	0.00	0.02	≈



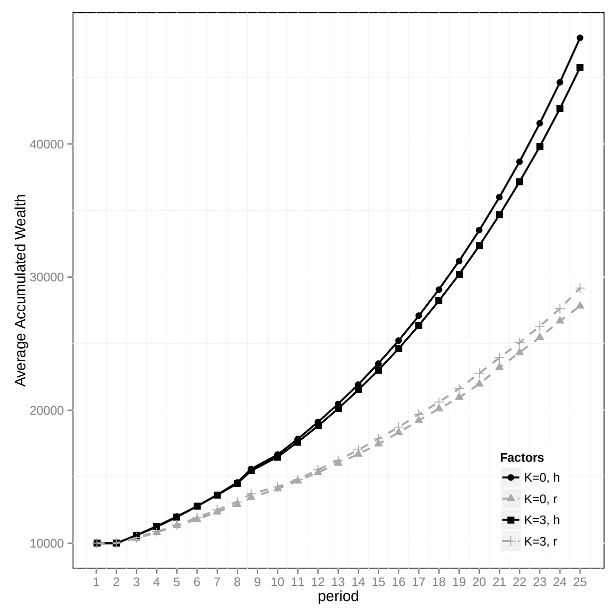
(a) Future Hypv ($\hat{\mathcal{S}}_{t+1}$)



(b) Cardinality

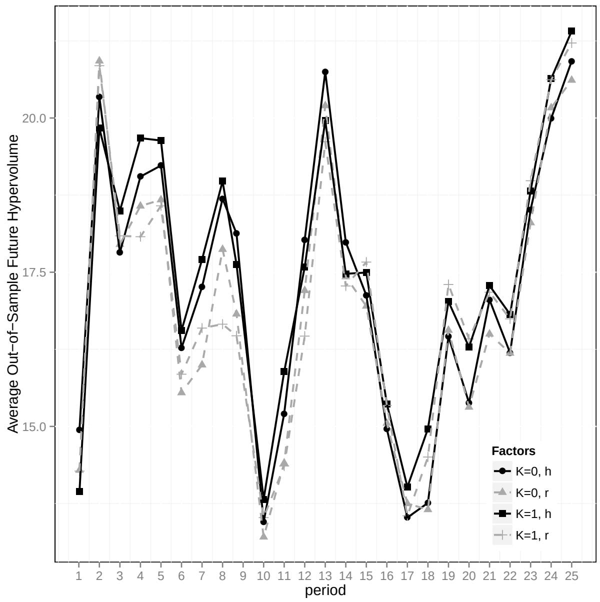


(c) Coherence

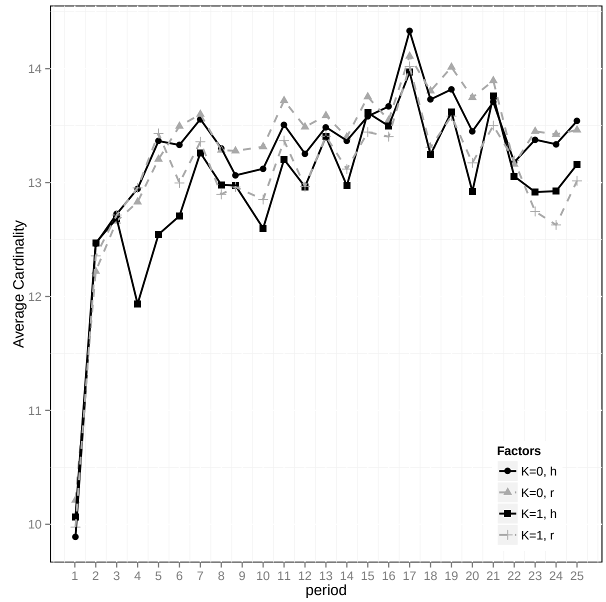


(d) Accumulated Wealth

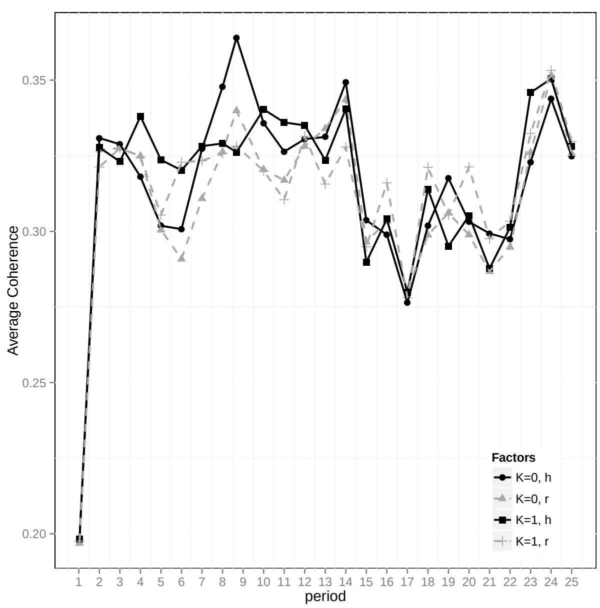
Figure 7.16: Dynamics of the evolved Pareto-flexible sets of portfolios at the end of each investment period for the static mean PO instance in terms of (a) average out-of-sample future Hypv; (b) average cardinality; (c) coherence; and (d) accumulated wealth.



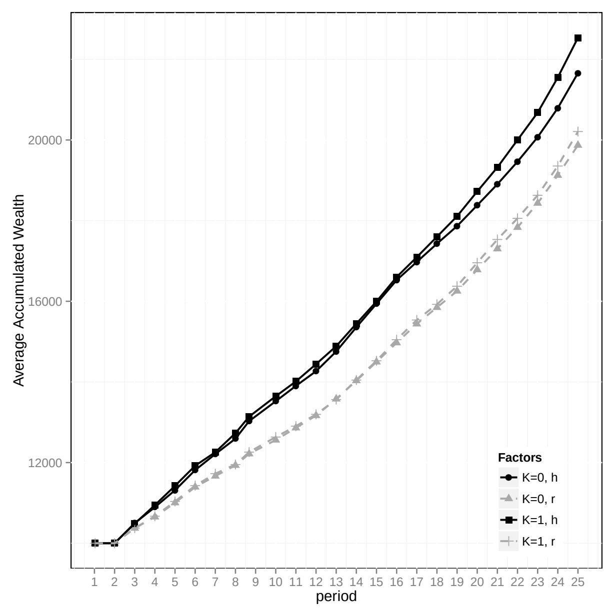
(a) Future Hypv ($\hat{\mathcal{S}}_{t+1}$)



(b) Cardinality

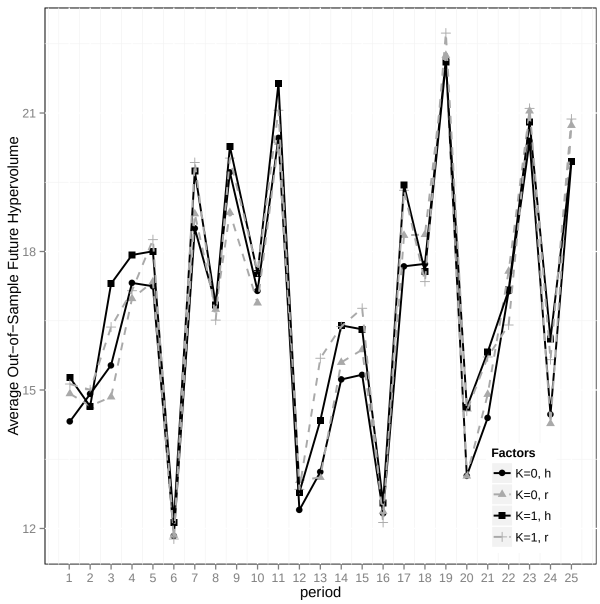


(c) Coherence

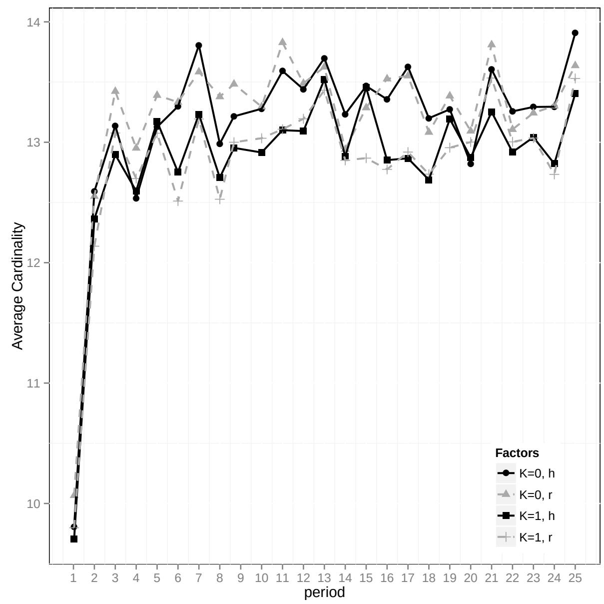


(d) Accumulated Wealth

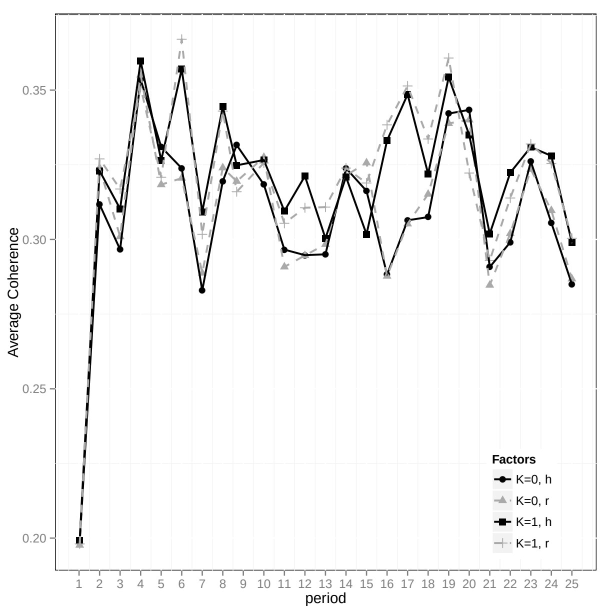
Figure 7.17: Dynamics of the evolved Pareto-flexible sets of portfolios at the end of each investment period for the PO(2,0.5) in terms of (a) average out-of-sample future Hypv; (b) average cardinality; (c) coherence; and (d) accumulated wealth.



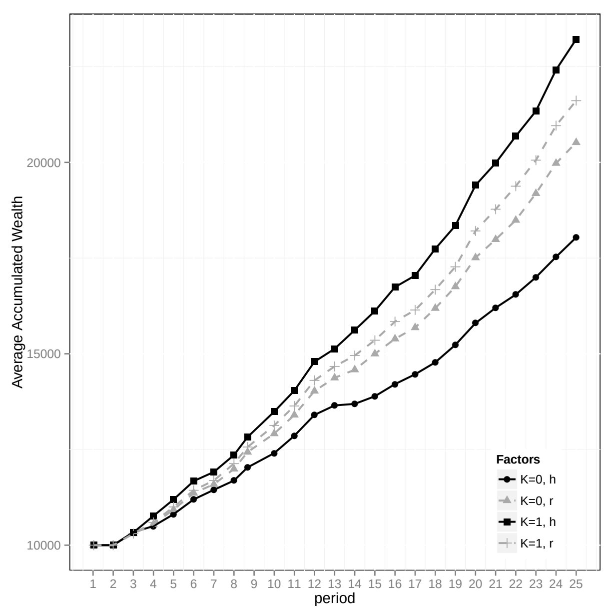
(a) Future Hypv ($\hat{\mathcal{S}}_{t+1}$)



(b) Cardinality

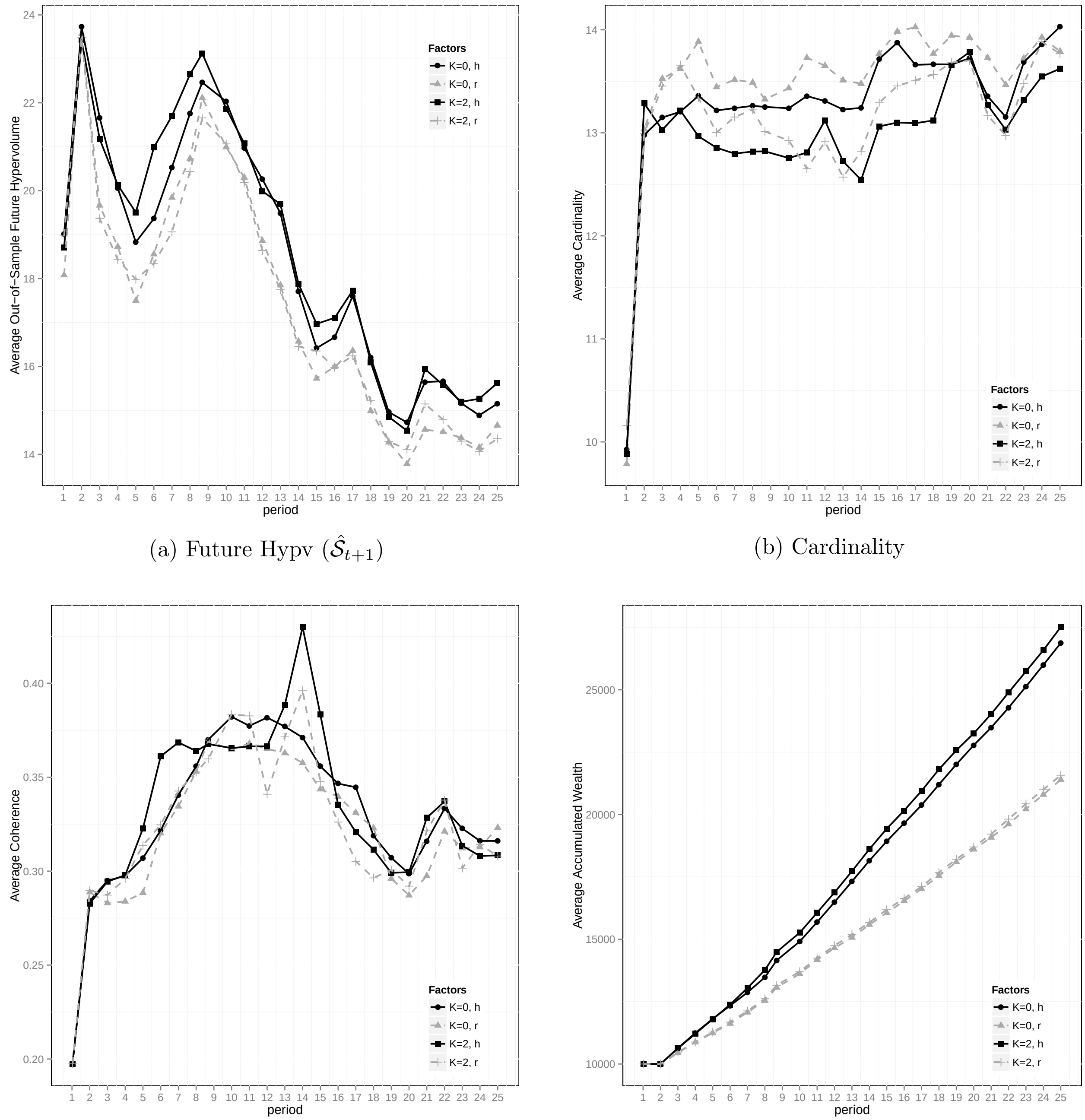


(c) Coherence



(d) Accumulated Wealth

Figure 7.18: Dynamics of the evolved Pareto-flexible sets of portfolios at the end of each investment period for the PO(2,1.0) in terms of (a) average out-of-sample future Hypv; (b) average cardinality; (c) coherence; and (d) accumulated wealth.



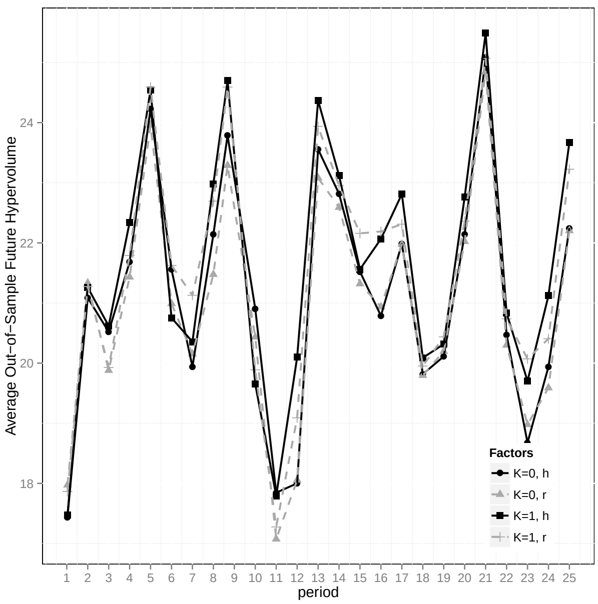
(a) Future Hypv ($\hat{\mathcal{S}}_{t+1}$)

(b) Cardinality

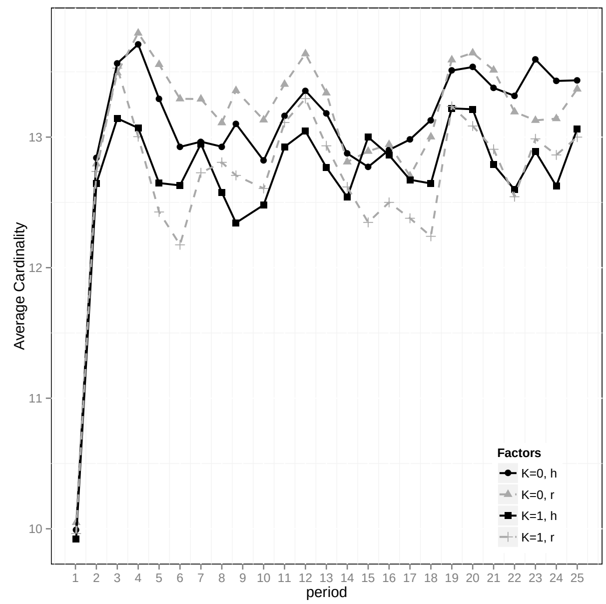
(c) Coherence

(d) Accumulated Wealth

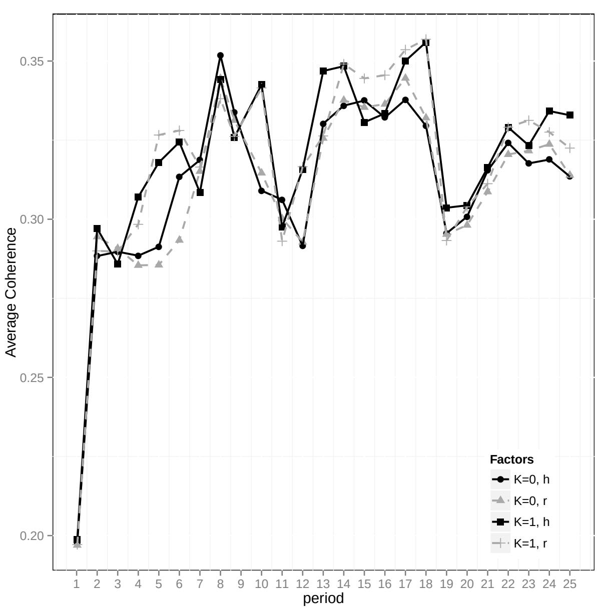
Figure 7.19: Dynamics of the evolved Pareto-flexible sets of portfolios at the end of each investment period for the PO(4,0.5) in terms of (a) average out-of-sample future Hypv; (b) average cardinality; (c) coherence; and (d) accumulated wealth.



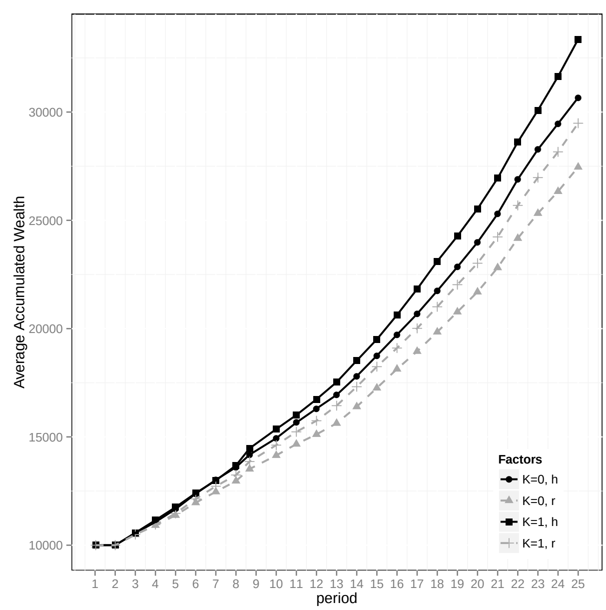
(a) Future Hypv ($\hat{\mathcal{S}}_{t+1}$)



(b) Cardinality

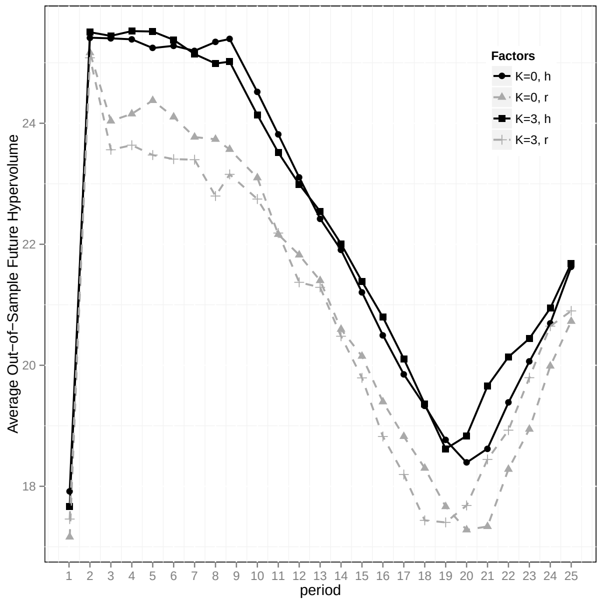


(c) Coherence

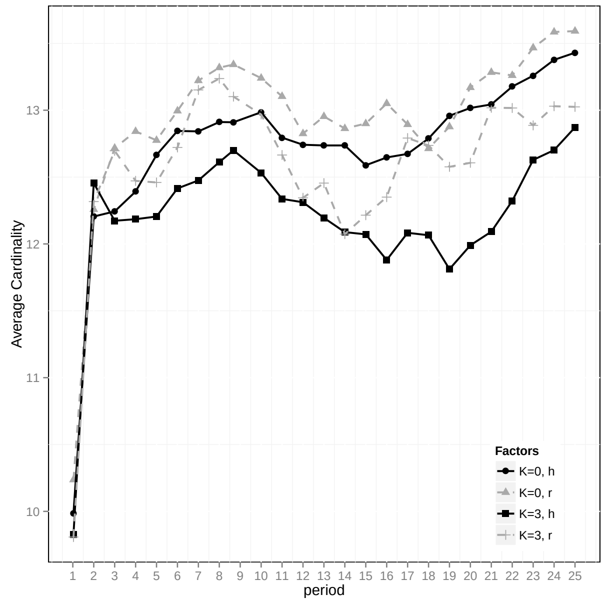


(d) Accumulated Wealth

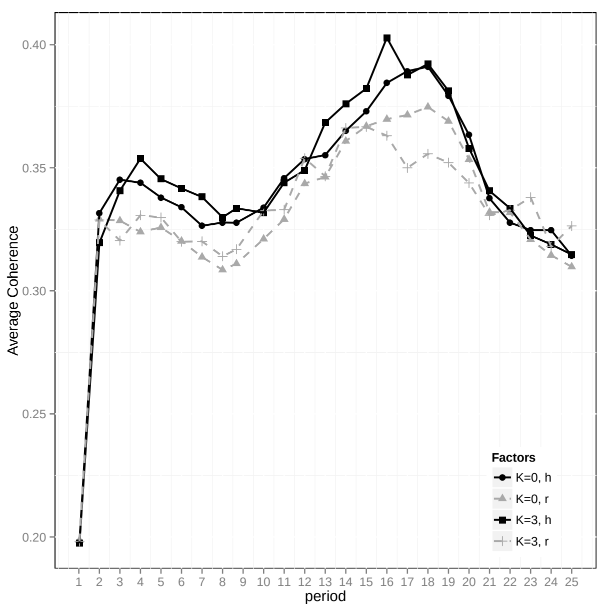
Figure 7.20: Dynamics of the evolved Pareto-flexible sets of portfolios at the end of each investment period for the PO(4,1.0) in terms of (a) average out-of-sample future Hypv; (b) average cardinality; (c) coherence; and (d) accumulated wealth.



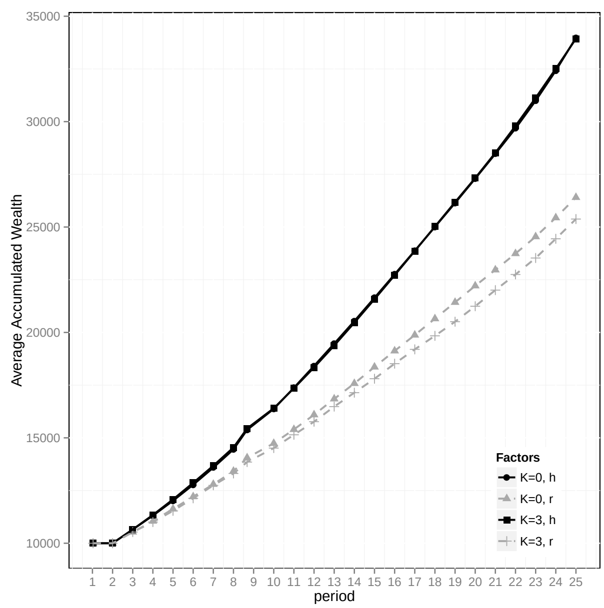
(a) Future Hypv ($\hat{\mathcal{S}}_{t+1}$)



(b) Cardinality

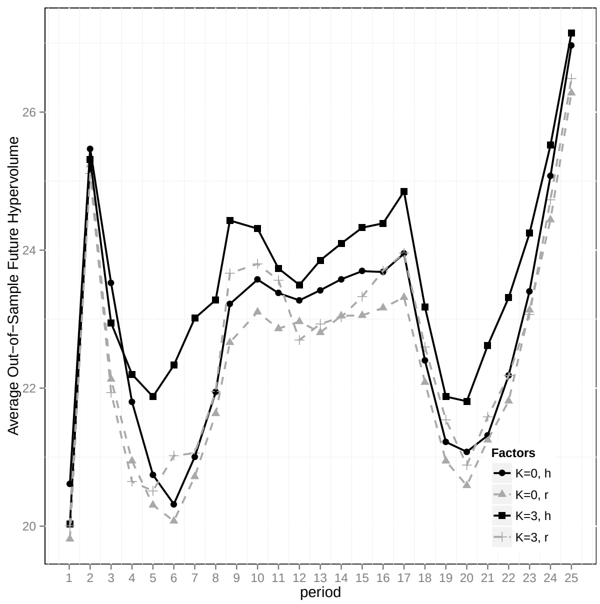


(c) Coherence

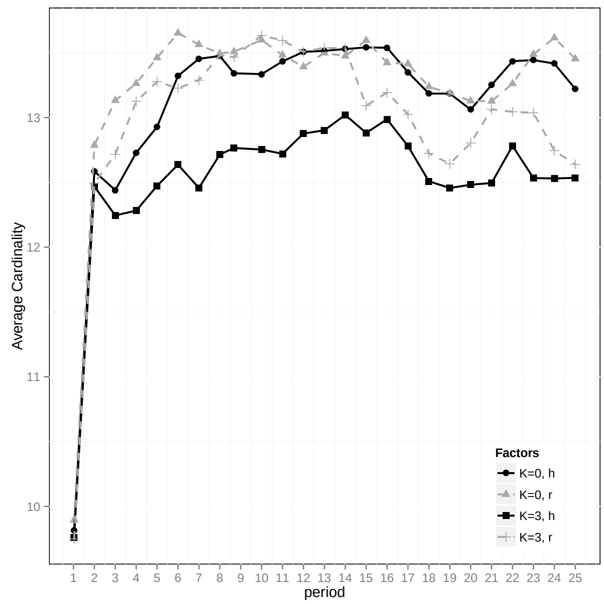


(d) Accumulated Wealth

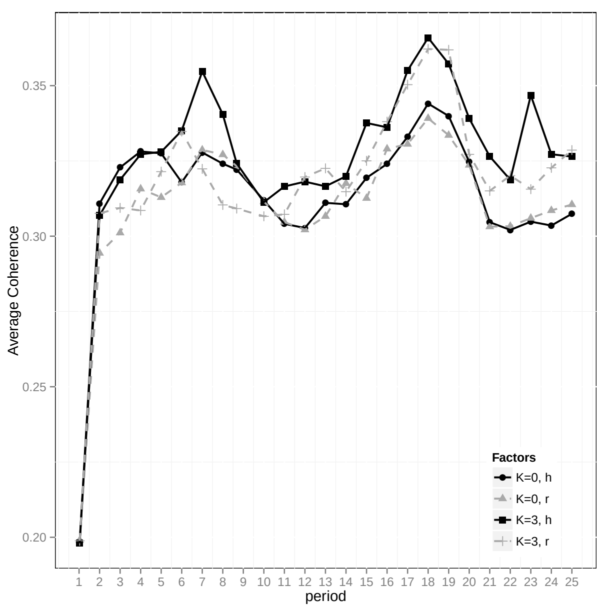
Figure 7.21: Dynamics of the evolved Pareto-flexible sets of portfolios at the end of each investment period for the PO(8,0.5) in terms of (a) average out-of-sample future Hypv; (b) average cardinality; (c) coherence; and (d) accumulated wealth.



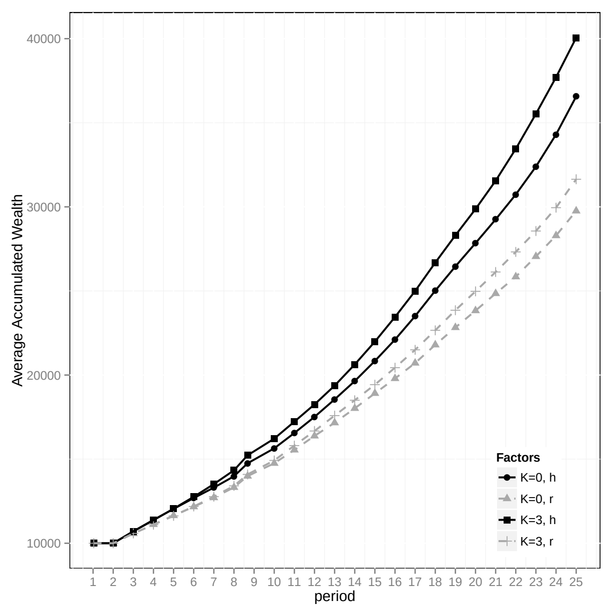
(a) Future Hypv ($\hat{\mathcal{S}}_{t+1}$)



(b) Cardinality

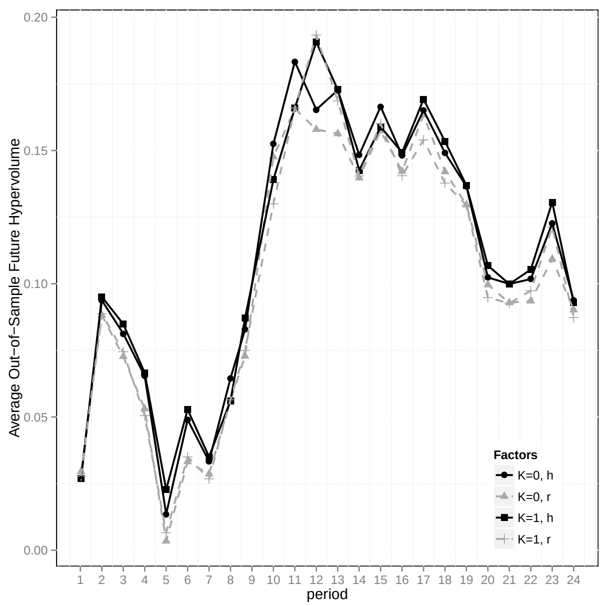


(c) Coherence

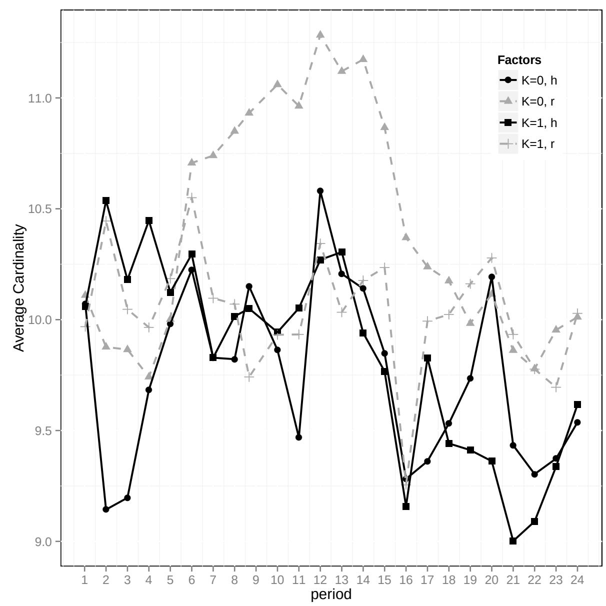


(d) Accumulated Wealth

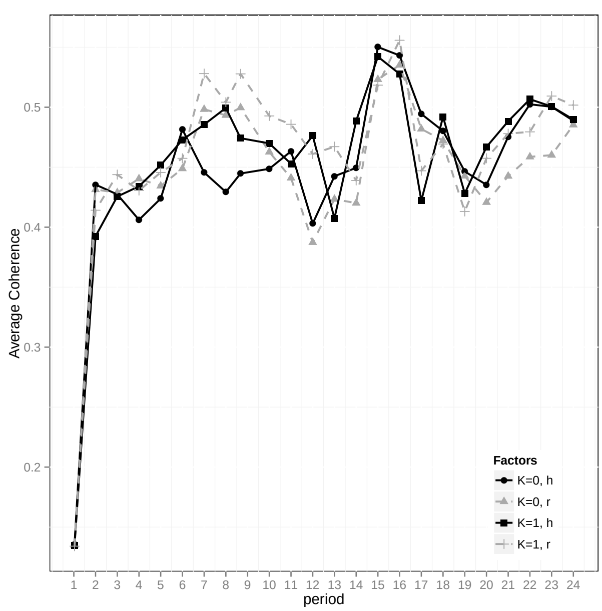
Figure 7.22: Dynamics of the evolved Pareto-flexible sets of portfolios at the end of each investment period for the PO(8,1.0) in terms of (a) average out-of-sample future Hypv; (b) average cardinality; (c) coherence; and (d) accumulated wealth.



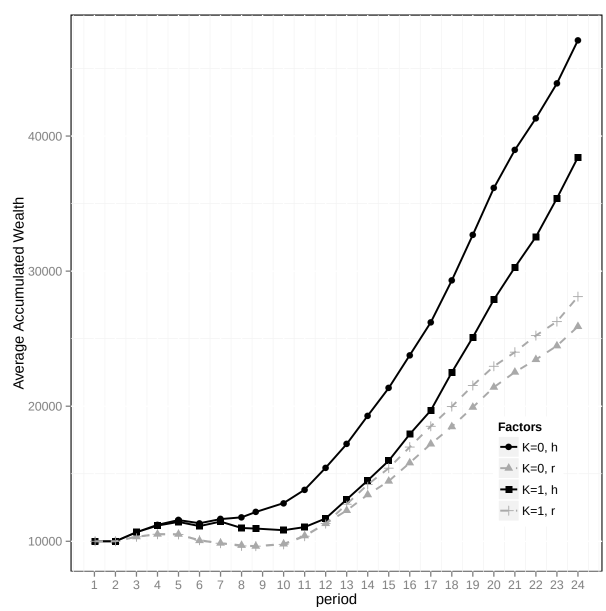
(a) Future Hypv ($\hat{\mathcal{S}}_{t+1}$)



(b) Cardinality

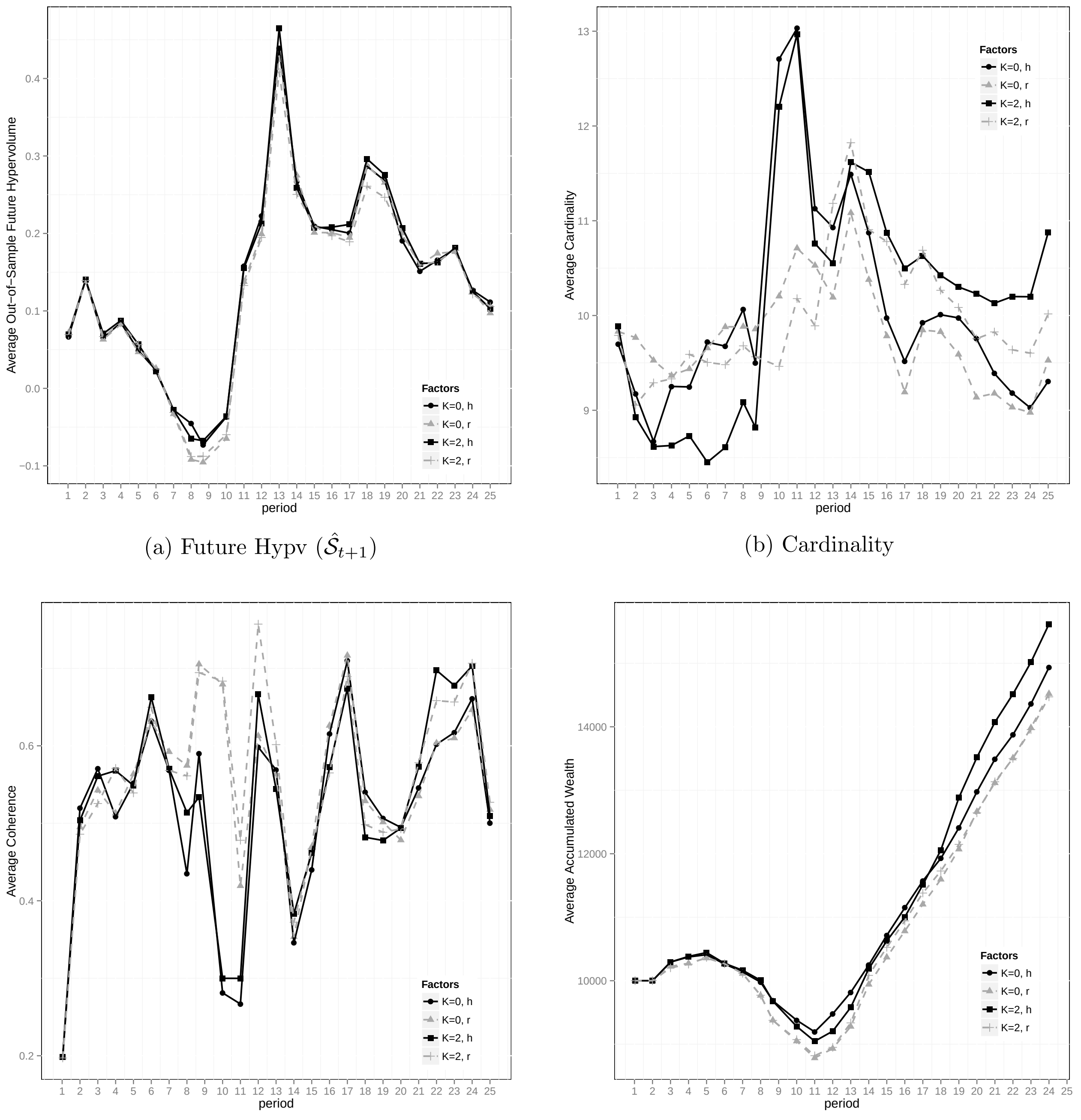


(c) Coherence



(d) Accumulated Wealth

Figure 7.23: Dynamics of the evolved Pareto-flexible sets of portfolios at the end of each investment period for the FTSE-100 in terms of (a) average out-of-sample future Hypv; (b) average cardinality; (c) coherence; and (d) accumulated wealth.



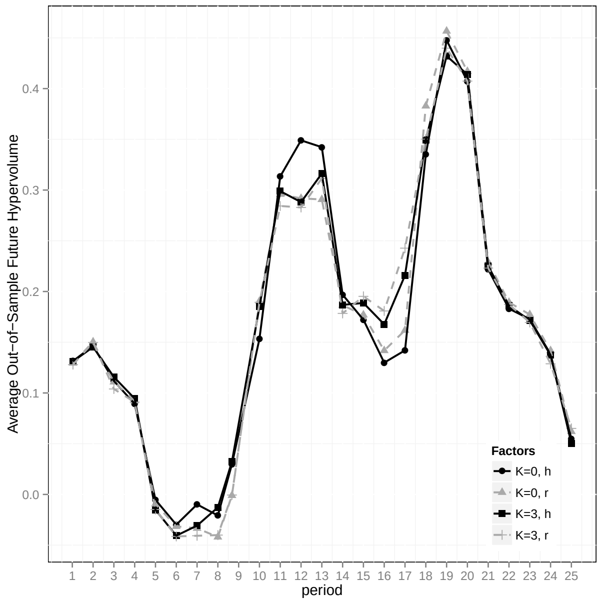
(a) Future Hypv ($\hat{\mathcal{S}}_{t+1}$)

(b) Cardinality

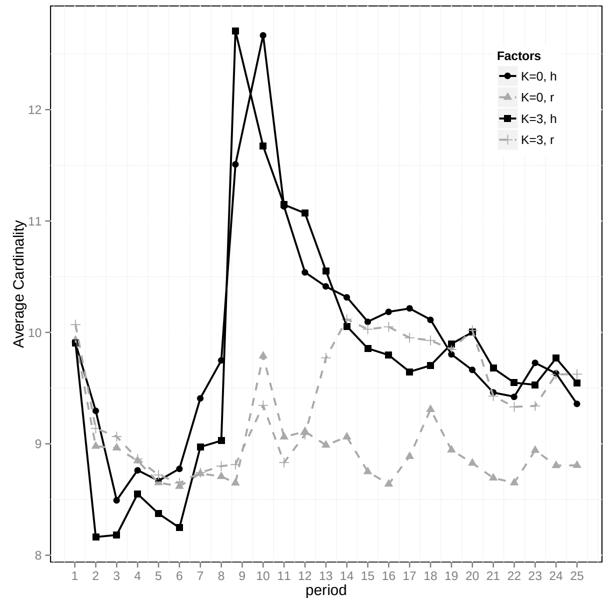
(c) Coherence

(d) Accumulated Wealth

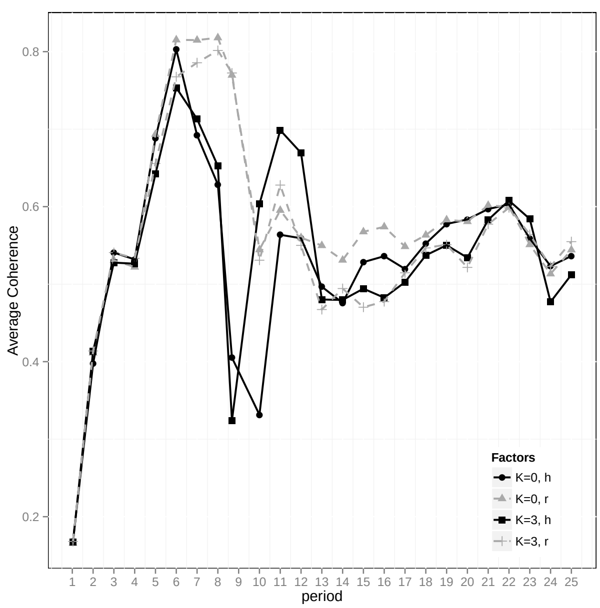
Figure 7.24: Dynamics of the evolved Pareto-flexible sets of portfolios at the end of each investment period for the DJI in terms of (a) average out-of-sample future Hypv; (b) average cardinality; (c) coherence; and (d) accumulated wealth.



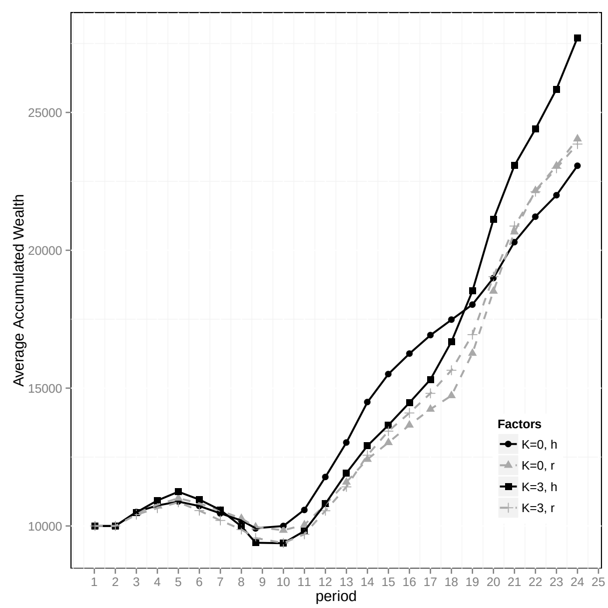
(a) Future Hypv ($\hat{\mathcal{S}}_{t+1}$)



(b) Cardinality



(c) Coherence



(d) Accumulated Wealth

Figure 7.25: Dynamics of the evolved Pareto-flexible sets of portfolios at the end of each investment period for the HSI in terms of (a) average out-of-sample future Hypv; (b) average cardinality; (c) coherence; and (d) accumulated wealth.

7.3.4 Dynamics of Future Average Hypv and of the Portfolio Compositions

The average evolution of out-of-sample future expected Hypv ($\hat{\mathcal{S}}_{t+1}$) over all investment periods can be visualized in part (a) of Figs. 7.16–7.22. For instance, the disruption patterns in the

Table 7.3: Pearson Correlation Coefficients (**$p < 0.01$; *$p < 0.05$)

Instance	Coherence vs. $\hat{\mathcal{S}}_{t+1}$	Cardinality vs. $\hat{\mathcal{S}}_{t+1}$	Coherence vs. Cardinality
Static mean PO	0.70**	0.18**	-0.01
PO(2,0.5)	0.42**	-0.12**	-0.03
PO(2,1.0)	0.19**	-0.05	-0.01
PO(4,0.5)	0.27**	-0.32**	-0.02
PO(4,1.0)	0.38**	-0.08*	-0.03
PO(8,0.5)	0.16**	-0.07	-0.06
PO(8,1.0)	0.30**	-0.09*	0.03
DJI	-0.07	-0.12**	-0.30**
FTSE	0.20**	-0.10**	-0.34**
HSI	-0.24**	0.09*	-0.35**

PO(8,1.0) simulated market (Fig. 7.16 (a)) can be inferred from the maximal $\hat{\mathcal{S}}_{t+1}$ inflection points at every 8 periods, after which performance degrades as prediction fails to respond to abrupt changes until the implemented Bayesian tracking methods manage to collect enough evidence to capture new trends in both the objective and search spaces. The same observation holds for the other benchmarks.

From the parts (b) and (c) of those time series, we noted a tendency – particularly for the static mean PO, PO(4,0.5), PO(8,0.5), PO(8,1.0) – wherein the ASMS/mHDM variant achieved the *highest* average portfolio coherences for most investment periods, while yielding at the same time the *lowest* cardinalities for most decision periods. While this tendency was not consistent for all instances and across every investment period, such result motivated us to perform a statistical correlation study between coherence, cardinality, and future average out-of-sample Hypv. We point out that the experimental data utilized in this study consider all four algorithmic variants (SMS/RDM, SMS/mHDM, ASMS/RDM, and ASMS/mHDM) over all decision periods. That is to say, statistically significant different than zero correlation coefficients suggest that consistent patterns of pairwise interaction between those three measures may exist regardless of the algorithm and of the investment period being analyzed. However, even where non-significant correlation is observed, the analysis can not rule out the possibility that significant correlation might exist for certain algorithmic variants and investment periods.

The Pearson correlation of $\hat{\mathcal{S}}_{t+1}$ with (1) coherence; and with (2) cardinality – and between the latter two measures – was computed over all settings ($K \in \{0, 1, 2, 3\}$ and DM $\in \{\text{mHDM}, \text{RDM}\}$) and can be read from Tab. 7.3.3. Coherence was significantly positively correlated with $\hat{\mathcal{S}}_{t+1}$ for 8 out of the 10 instances. On the other hand, correlation between cardinality and $\hat{\mathcal{S}}_{t+1}$ was significantly negative for 5 out of the 10 instances. Also, coherence was negatively correlated with cardinality on the three real-world market instances.

As illustrated in e.g. Fig. 7.22 (a)–(c) for the PO(8,1.0) simulated market, those results may indicate that portfolios with a smaller cardinality may have allowed for a more efficient search for both the myopic SMS and the ASMS variants. Another hypothesis is that lower cardinality may have lead to more coherent portfolio populations that, by its turn, may have helped mitigating the effects of the transaction costs from the DM choices on the Hypv of future investment decisions. The correlation patterns for cardinality vs. $\hat{\mathcal{S}}_{t+1}$ were not as significant when compared with the correlations between coherence vs. cardinality for the real-world markets and with the correlations between coherence vs. $\hat{\mathcal{S}}_{t+1}$ in general, but we suspect that cardinality may also be directly associated with reduced adaptation costs so to improve $\hat{\mathcal{S}}_{t+1}$.

7.3.5 Dynamics of the Investor Wealth

The average expected wealth of an investor implementing the choices produced by both SMS and the ASMS variants are given in part (d) of Figs. 7.16–7.22. It can be noted that the mHDM variant has outperformed the RDM in 7 out of the 10 (70%) PO instances. The results for the real-world datasets are given in parts (d) of Figs. 7.23–7.25, for which similar patterns of wealth evolution were observed. It can be noted that both modeled investors have faced negative returns between periods 5–11, in all three instances, although the mHDM had a slight edge over the RDM in the FTSE and DJI markets.

Interestingly, the period $t = 11$ corresponds to the starting date of 08/03/2009, the day after the FTSE 100 index closed at a historical six-year 3,700 points lower mark, during the financial crisis. That day later proved to be the turning point for a steady market uphill recovery that still endures to this day. For FTSE, the SMS/mHDM variant achieved the highest long-run returns, whereas for the DJI and HSI markets, the ASMS/mHDM investment decisions performed better. Those results suggest that, although expected wealth was not directly optimized for, implementing flexible decisions may not necessarily conflict with wealth maximization and can actually help the DM on achieving higher returns. This result is consistent with the account of Benjaafar et al. [25] discussed in **chapter 1**, Sec. 1.2.5.

7.3.6 Average Predicted Future Hypv Along the Evolved SPFs

Figures 7.26–7.35 show the predicted future Hypv values, averaged across all decision periods, distributed over the evolved mutually Pareto-incomparable portfolios, from which the patterns of choice of the mHDM variant can be inferred – i.e., higher predicted future Hypv portfolios tended to be more frequently chosen. For some PO instances such as the DJI market, the same tendency observed for the confidence distribution in Fig. 7.5 is present for the predicted Hypv on DJI (Fig. 7.34): when $K = 0$, the portfolios predicted to yield maximal future Hypv tended to be associated with higher return/risk, whereas the trend was reversed for $K > 1$. Thus, in that case, the *myopic* SMS/mHDM ($K = 0$) tended to produce more risky portfolio sequences, whereas ASMS/mHDM ($K > 0$) tended to choose flexible portfolios with less expected risk.

This behavioral shift may be associated with the improved predictability of the less risky portfolios (comparing $K > 0$ to $K = 0$), as captured by the confidence patterns in Fig. 7.5, but, in the general case, this can not be inferred from every other instance, as in the case of the FTSE and the HSI markets, where the highest predicted future expected Hypv portfolios were associated with lesser predictive confidence. We regard the association between higher predicted future Hypv and higher estimated predictability, where it exist, as evidence that the ASMS/mHDM variant may indeed generate *anticipatory trajectories of safer and flexible* investment decisions, as intended in OAL, and as exemplified in Fig. 5.5, although there seems to exist an inherent conflict between maximal future Hypv sets and predictive confidence for some PO instances, include the assessed real-world markets.

Finally, it is worth emphasizing that there has been *at least* one value of $K > 0$ for which ASMS/mHDM has outperformed the baseline SMS-EMOA/RDM in terms of $\hat{\mathcal{S}}_{t+1}$, for the whole test suite, as supported by the boxplots results of Fig. 7.15. Considering all 30 settings from the combination of $K \in \{1, 2, 3\}$ with the 10 PO instances, we point out that ASMS/mHDM significantly outperformed SMS-EMOA/RDM in 26 out of the 30 (87%) cases. Thus, irrespective of the dependency of the ASMS performance on the tracking parameters (i.e., K), the significantly greater $\hat{\mathcal{S}}_{t+1}$ values achieved by ASMS/mHDM are indicative that the AS-MOO methodology has attained the research main goal.

7.3.7 Effects of Using the Wrong Predictive Models

The last noteworthy observation from our experimental study regards the effects of using the *wrong* predictive models for the most challenging artificial instance, namely, PO(2,1.0). That instance corresponds to the fastest and highest disruptive investment environment, wherein the joint return distribution parameters change in a completely random way (i.e., with maximum severity) after each two consecutive periods of time.

Under those conditions, we study the effects of using the "worst" and the "best" possible predictive models, considering the underlying linear dynamics assumption of the Kalman Filter and the Dirichlet MAP estimation, on *flexibility* – i.e., on the observed out-of-sample average expected future Hypv.

The "worst" predictive models for PO(2,1.0) are those resulting from setting the window size parameter to $K = 0$ in SMS, which leads to the assumption of a static environment, yielding the following constant predictive rule in both the decision and the objective space:

> *The next state is exactly equal to the current state estimation.*

This is an obviously wrong assumption for an environment which is changing fast towards completely unpredictable new states.

Conversely, the "best" predictive models for PO(2,1.0) – constrained to our Gauss-Markov (see **chapter 2**, section 2.3.4) and linear dynamics (see **chapter 6**, section 6.1.1 and Eq. (6.28)) assumptions – are those resulting from setting the window size parameter to $K = 1$ (ASMS), which leads to the assumption of a dynamic environment, (roughly) yielding the following predictive rule in both the decision and the objective space:

> *The next state lies in the direction estimated from the difference between the previous and the current state estimation.*

This is clearly the best possible value for K among those tested ($K \in \{0, 1, 2, 3\}$), since it is expected to correctly capture the direction of change, as measured by POCID in Eq (7.13), in roughly 50% of the investiment periods. The reason is that, with $K = 1$, the implemented predictive models in the decision and objective spaces are able to correct the captured trend information faster than the environment is actually changing (disruptions at every $\tau = 2$ periods); whereas the learned trend information will have a chance of being correct at roughly 50% of the periods which *immediately follows disruptive random changes* in the joint return distribution.

From Fig. 7.14 (c), we observe that the average POCID values for both types of decision makers (random and max Hypv) were roughly 44.0%, for SMS ($K = 0$), and 50.5%, for ASMS (with $K = 1$). Note that the result for $K = 1$ confirmed our prior expectations, whereas the result for $K = 0$ reflects the specific proportion of periods in the artificial PO(2,1.0) instance for which the joint return distribution has gone "up" and "down", both in terms of expected return and risk.

Moreover, from Fig. 7.15 (c), we observe that the average expected future Hypv values ($\hat{\mathcal{S}}_{t+1}$) for both types of decision makers (random and max Hypv) were roughly 15.4, for SMS ($K = 0$), and 16.0, for ASMS (with $K = 1$). Note also from that Figure that assuming the "right" predictive model has lead ASMS to outperform SMS on $\hat{\mathcal{S}}_{t+1}$ with statistical significance, which is confirmed by the significance of the window size factor in the ANOVA results in Table 7.3.3.

Finally, from Fig. 7.18 (d), we observe that the average expected wealth in the long-run after 25 portfolio rebalancing decisions, for each combination of decision makers (random and max Hypv) and of the "wrong" and "right" predictive models were roughly (ordered from the worst to the best results): $\$18,000$, for SMS/mHDM ($K = 0$); $\$20,100$, for SMS/RDM ($K = 0$); $\$21,000$, for ASMS/RDM (with $K = 1$); and $\$23,000$, for ASMS/mHDM (with $K = 1$).

Thus, in terms of the average expected long-run investor wealth, we observe that the worst result was obtained *when maximizing flexibility* under the "wrong" prediction models (SMS/mHDM, $K = 0$). In fact, under the "wrong" assumption of a static environment for P(2,1.0), selecting random portfolio rebalancing decisions from the evolved stochastic Pareto Frontier (SMS/RDM, $K = 0$) has lead to significantly higher wealth, when compared to SMS/mHDM ($K = 0$).

Remarkably, when the "right" prediction models were assumed (by setting $K = 1$), *maximizing flexibility* (ASMS/mHDM, with $K = 1$) has significantly outperformed all other decision-making methods and prediction assumptions.

Therefore, we conclude from this discussion that:

1. Maximizing flexibility does not always lead to improved performance – and can actually lead to decreased performance – in terms of the expected realized long-run accumulated value which results from the trajectory of the implemented decisions. That negative result was observed only for the most challenging artificial instance, PO(2,1.0), by purposefully forcing the mHDM variant to use completely "wrong" prediction models; and

2. Conversely, by using the "right" prediction models, maximizing flexibility can lead to the best performance in terms of the expected realized long-run accumulated value resulting from the trajectory of flexible choices.

7.4 Concluding Remarks

The results of the proposed AS-MOO methodology application for portfolio selection in time-varying and stochastic investment environments, considering both simulated and real-world markets, showed that the proposed anticipatory algorithm was competent in obtaining sets of mutually Pareto-incomparable solutions yielding significantly higher average hypervolume, when evaluated for out-of-sample future data, compared to the baseline algorithms (Fig. 7.15). In addition, the following remarks are supported:

- Varying predictability levels were estimated for portfolios representing different portions of the stochastic Pareto frontier, what enabled effective self-adaptation of the anticipation rates encoding time preferences (Fig. 7.7–7.6);

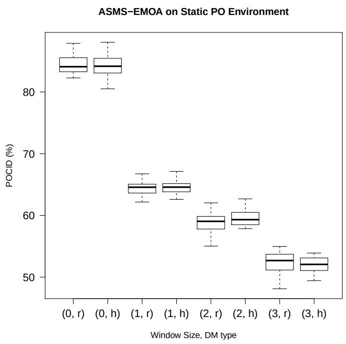
- The window sizes that allowed for the best Bayesian tracking performance were consistent with the different dynamical disruption patterns simulated for each benchmark scenario (Fig. 7.14); and

- The flexible decisions predicted to lead to higher future Hypv were sometimes associated to higher predictable portfolios (Figs. 7.7–7.6 and of Figs. 7.26–7.35), whose implementation has lead to coherent Pareto-flexible sets (part (b) of Figs. 7.16–7.25) that were significantly correlated with improvements in future Hypv (Table 7.3.3), although the inverse association has also been observed for certain instances.

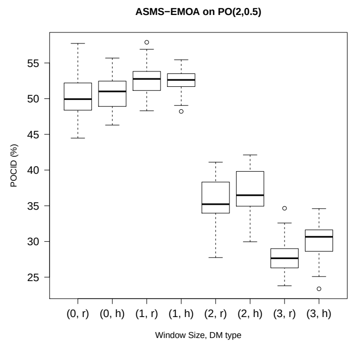
Therefore, the significantly greater sample average hypervolume obtained for out-of-sample future test data by the anticipatory decision-making approach strongly suggest the proposed AS-MOO methodology has attained the research main goal, in the context of the assessed portfolio selection application, which is that of improving the future quality and the diversity of choice of an investor who is unsure about his/her attitude to risky portfolios as well as about his/her willingness for near-term returns.

The successful widespread application of the proposed Anticipatory Stochastic MOO (AS-MOO) models and algorithms will require competent system identification and proper modeling assumptions that may capture the dynamics of the problem under consideration. On the other hand, this provides several research opportunities for improving the anticipatory EMOO models and algorithms proposed in this thesis.

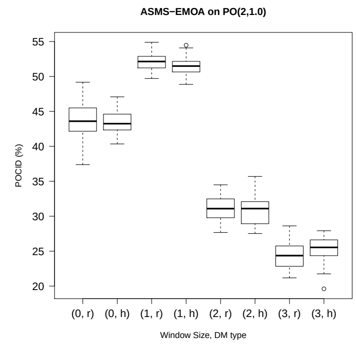
The next chapter presents concluding remarks, limitations of the research scope, as well as suggestions for future works on flexible anticipatory multi-objective machine learning models and algorithms for sequential MCDM under uncertainty.



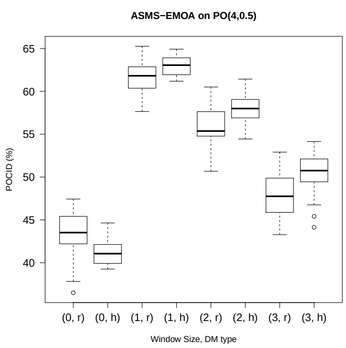
(a) Static mean PO.



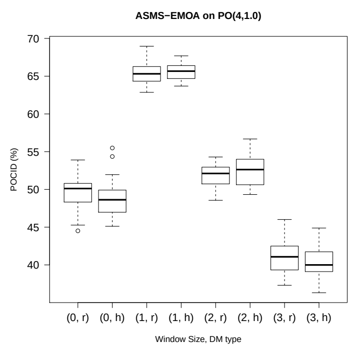
(b) PO(2,0.5)



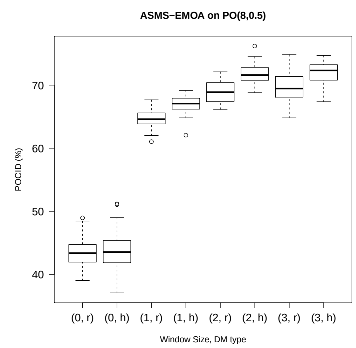
(c) PO(2,1.0)



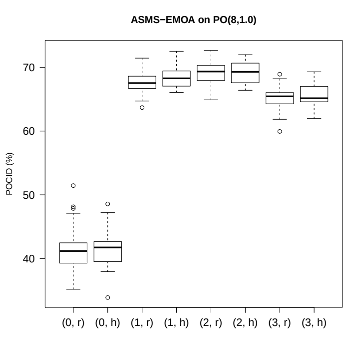
(d) PO(4,0.5)



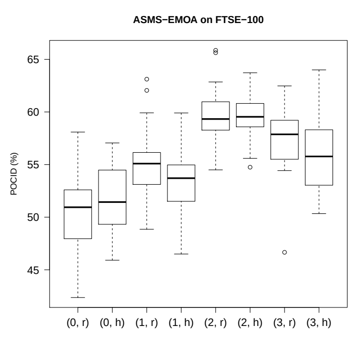
(e) PO(4,1.0)



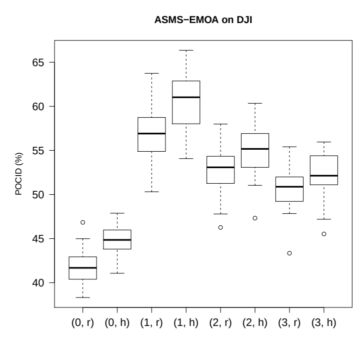
(f) PO(8,0.5)



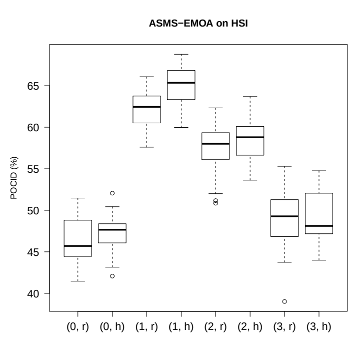
(g) PO(8,1.0)



(h) FTSE

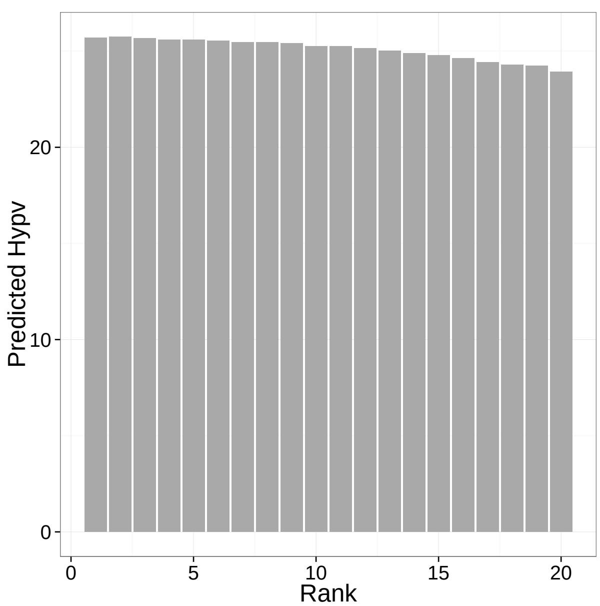


(i) DJI

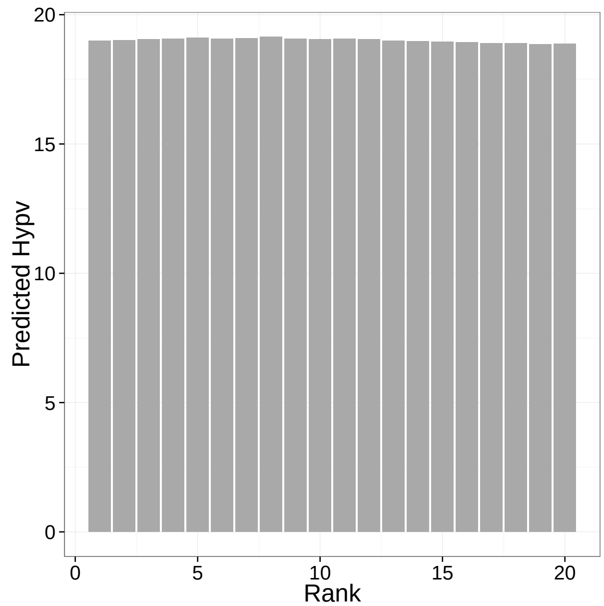


(j) HSI

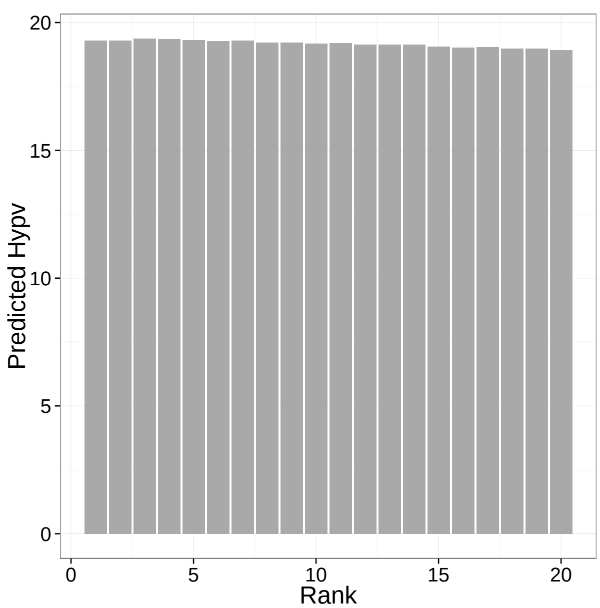
Figure 7.14: Boxplots of Prediction Of Change in Direction (POCID). In the legends, h stands for mHDM and r for RDM within the notation (K, DM).



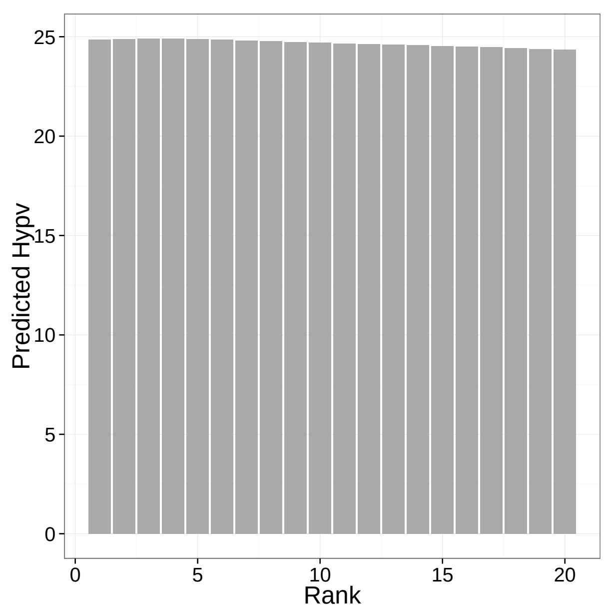
(a) Predicted Hypv (K=0) in static mean PO



(b) Predicted Hypv (K=1) in static mean PO

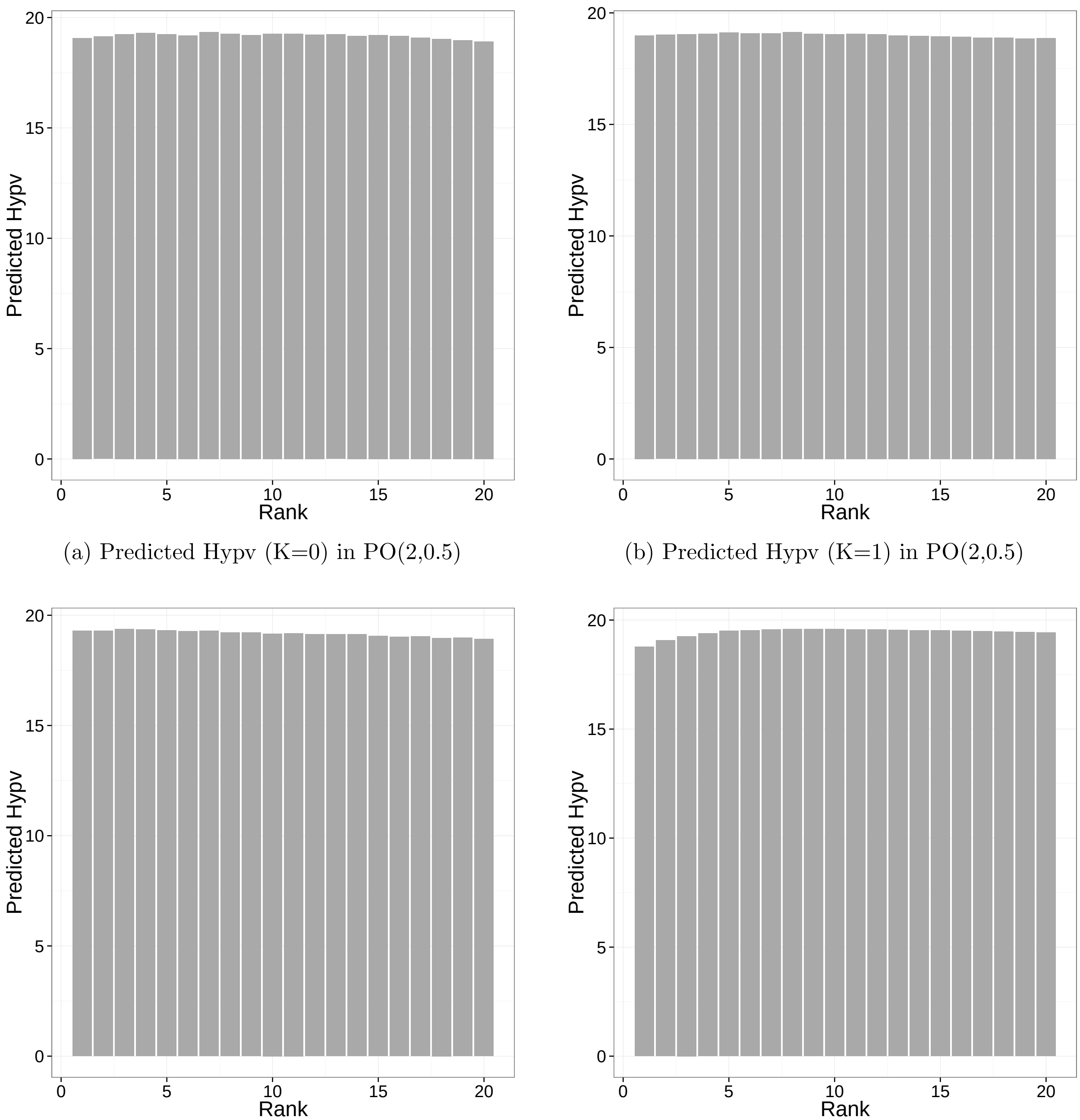


(c) Predicted Hypv (K=2) in static mean PO



(d) Predicted Hypv (K=3) in static mean PO

Figure 7.26: Bar plots showing the predicted future expected Hypv in ASMS/mHDM for the obtained ordered Pareto-flexible sets, averaged over all periods and experiments (for static mean PO).



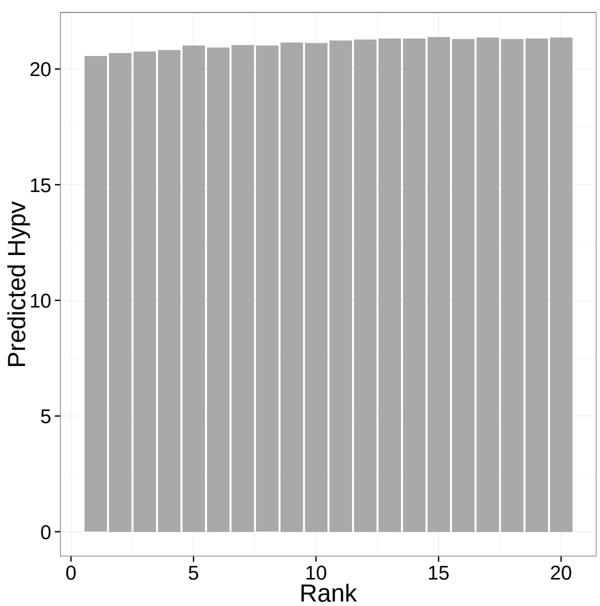
(a) Predicted Hypv (K=0) in PO(2,0.5)

(b) Predicted Hypv (K=1) in PO(2,0.5)

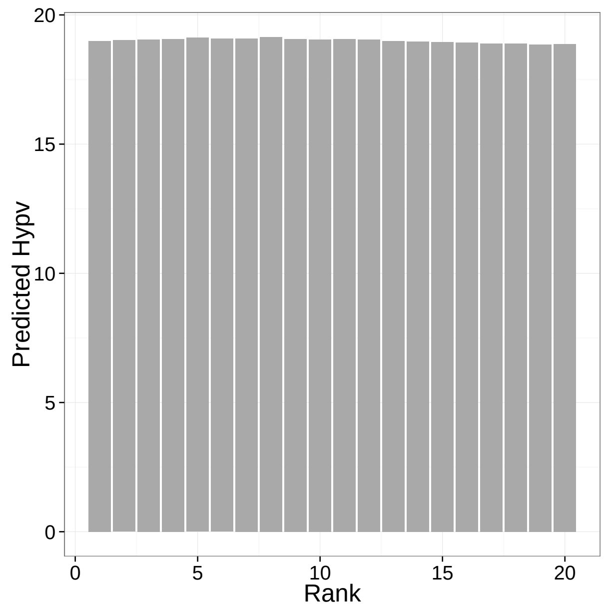
(c) Predicted Hypv (K=2) in PO(2,0.5)

(d) Predicted Hypv (K=3) in PO(2,0.5)

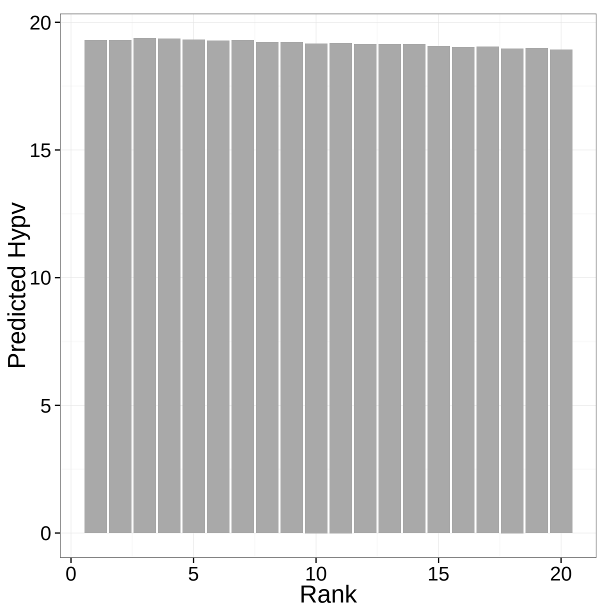
Figure 7.27: Bar plots showing the predicted future expected Hypv in ASMS/mHDM for the obtained ordered Pareto-flexible sets, averaged over all periods and experiments (for PO(2,0.5)).



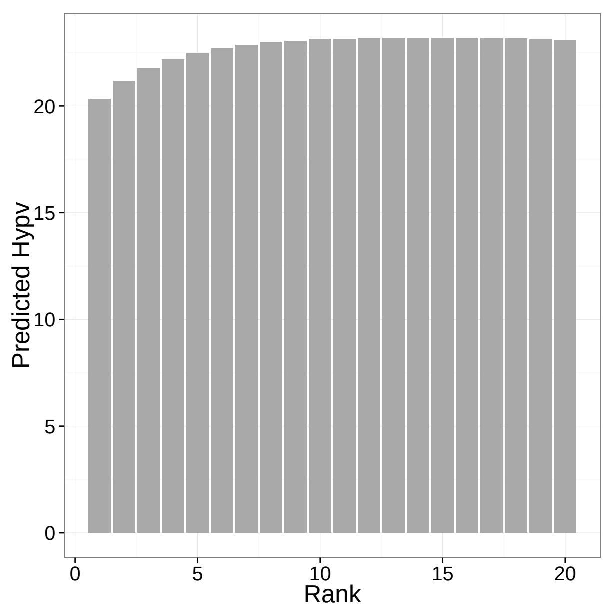
(a) Predicted Hypv (K=0) in PO(2,1.0)



(b) Predicted Hypv (K=1) in PO(2,1.0)

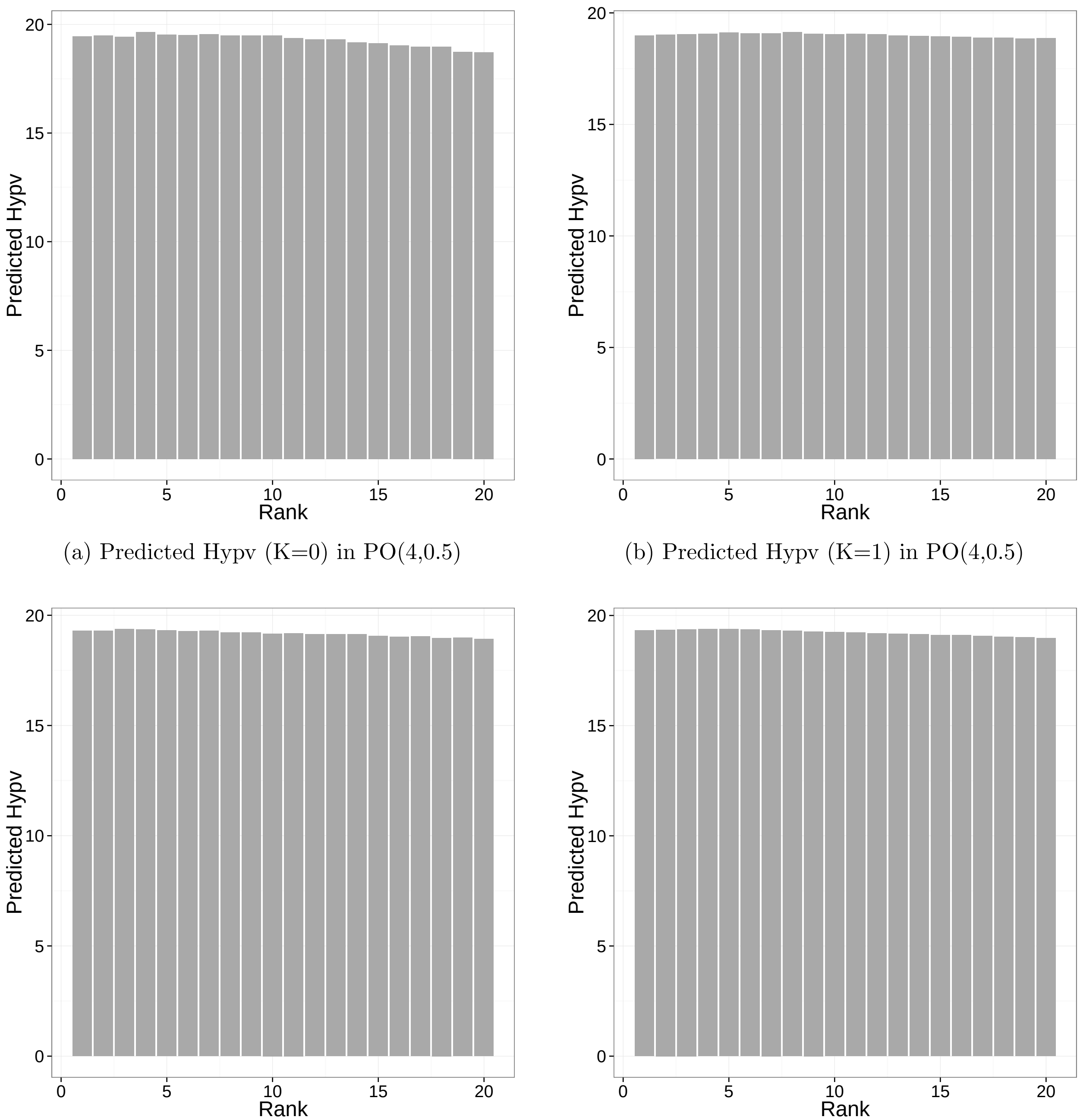


(c) Predicted Hypv (K=2) in PO(2,1.0)



(d) Predicted Hypv (K=3) in PO(2,1.0)

Figure 7.28: Bar plots showing the predicted future expected Hypv in ASMS/mHDM for the obtained ordered Pareto-flexible sets, averaged over all periods and experiments (for PO(2,1.0)).



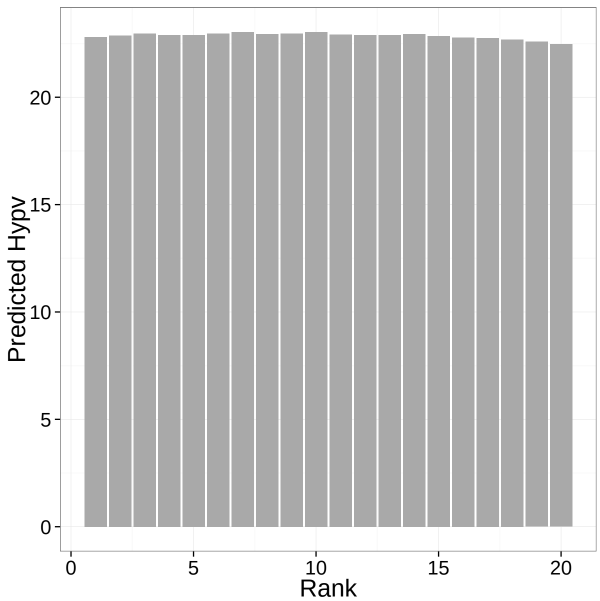
(a) Predicted Hypv (K=0) in PO(4,0.5)

(b) Predicted Hypv (K=1) in PO(4,0.5)

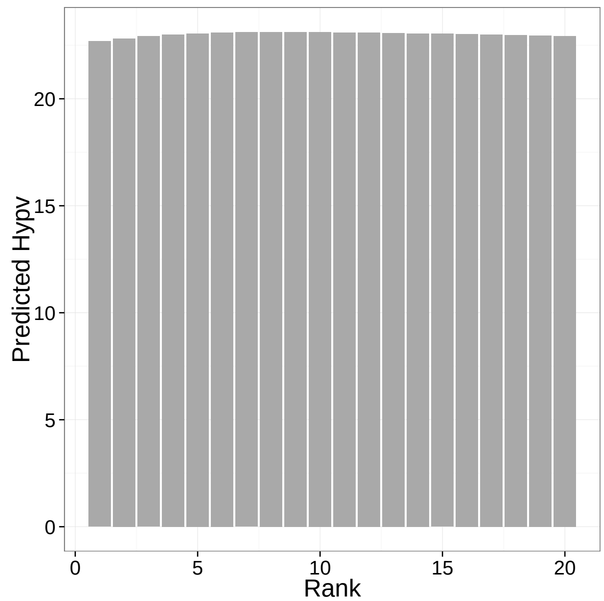
(c) Predicted Hypv (K=2) in PO(4,0.5)

(d) Predicted Hypv (K=3) in PO(4,0.5)

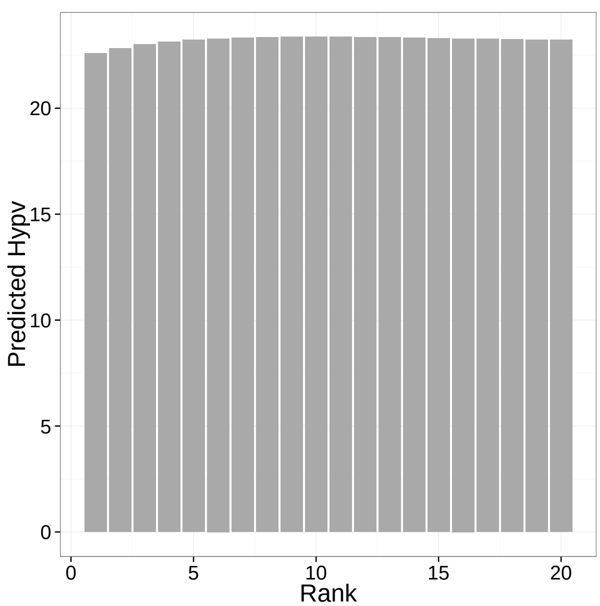
Figure 7.29: Bar plots showing the predicted future expected Hypv in ASMS/mHDM for the obtained ordered Pareto-flexible sets, averaged over all periods and experiments (for PO(4,0.5)).



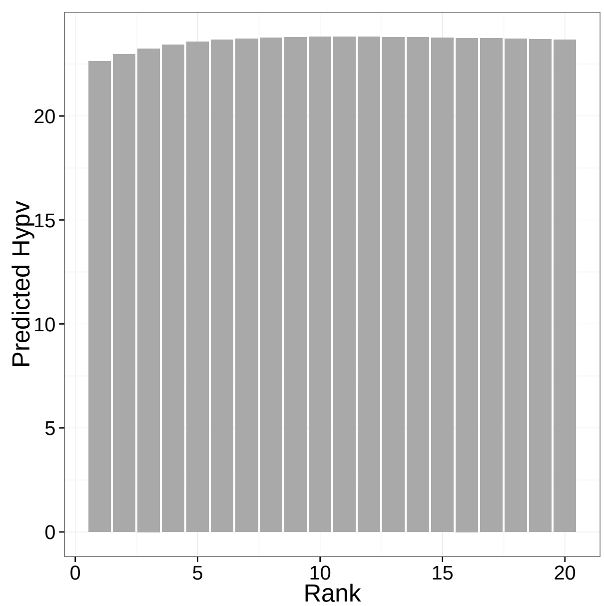
(a) Predicted Hypv (K=0) in PO(4,1.0)



(b) Predicted Hypv (K=1) in PO(4,1.0)

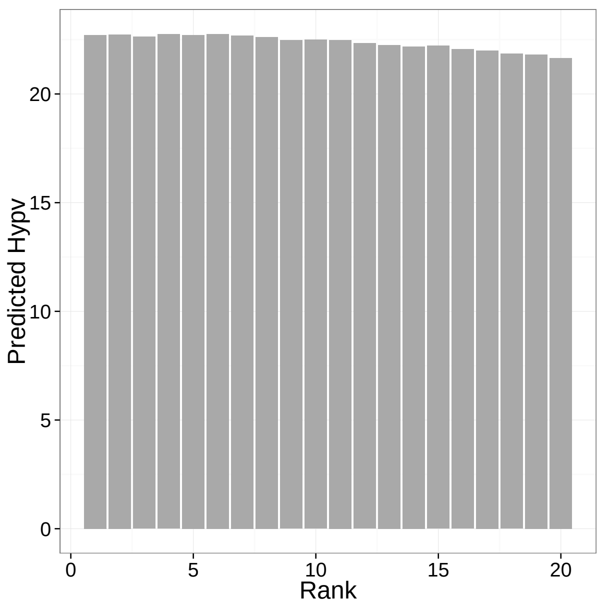


(c) Predicted Hypv (K=2) in PO(4,1.0)

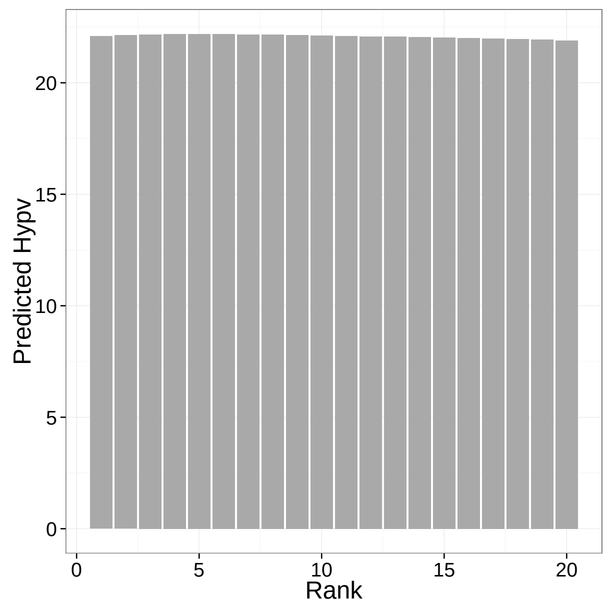


(d) Predicted Hypv (K=3) in PO(4,1.0)

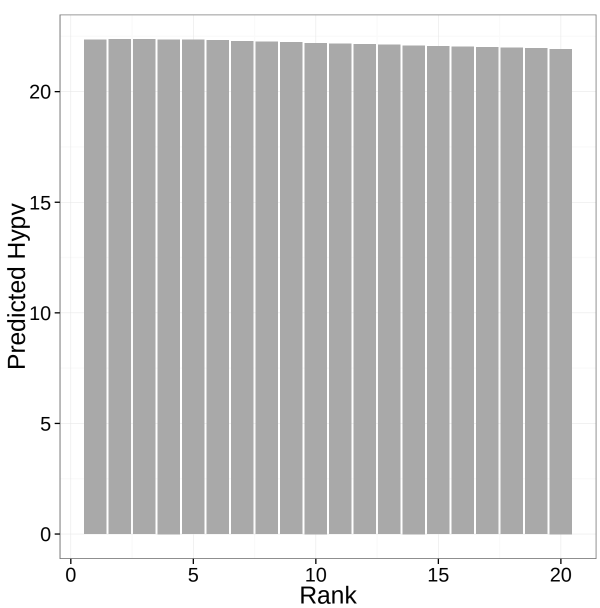
Figure 7.30: Bar plots showing the predicted future expected Hypv in ASMS/mHDM for the obtained ordered Pareto-flexible sets, averaged over all periods and experiments (for PO(4,1.0)).



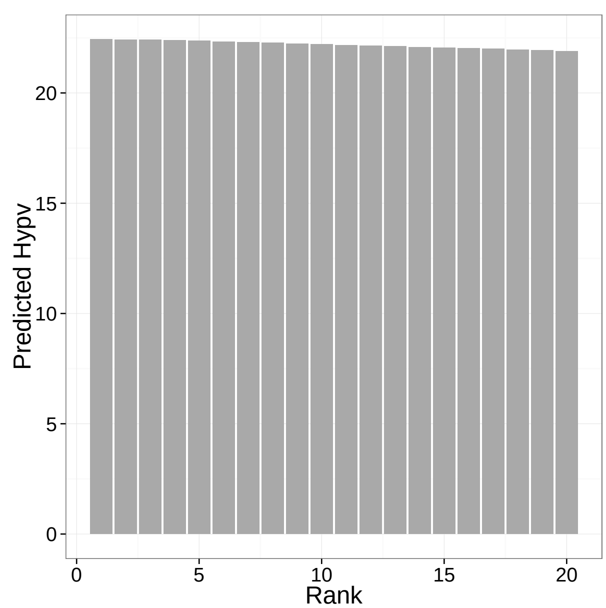
(a) Predicted Hypv (K=0) in PO(8,0.5)



(b) Predicted Hypv (K=1) in PO(8,0.5)

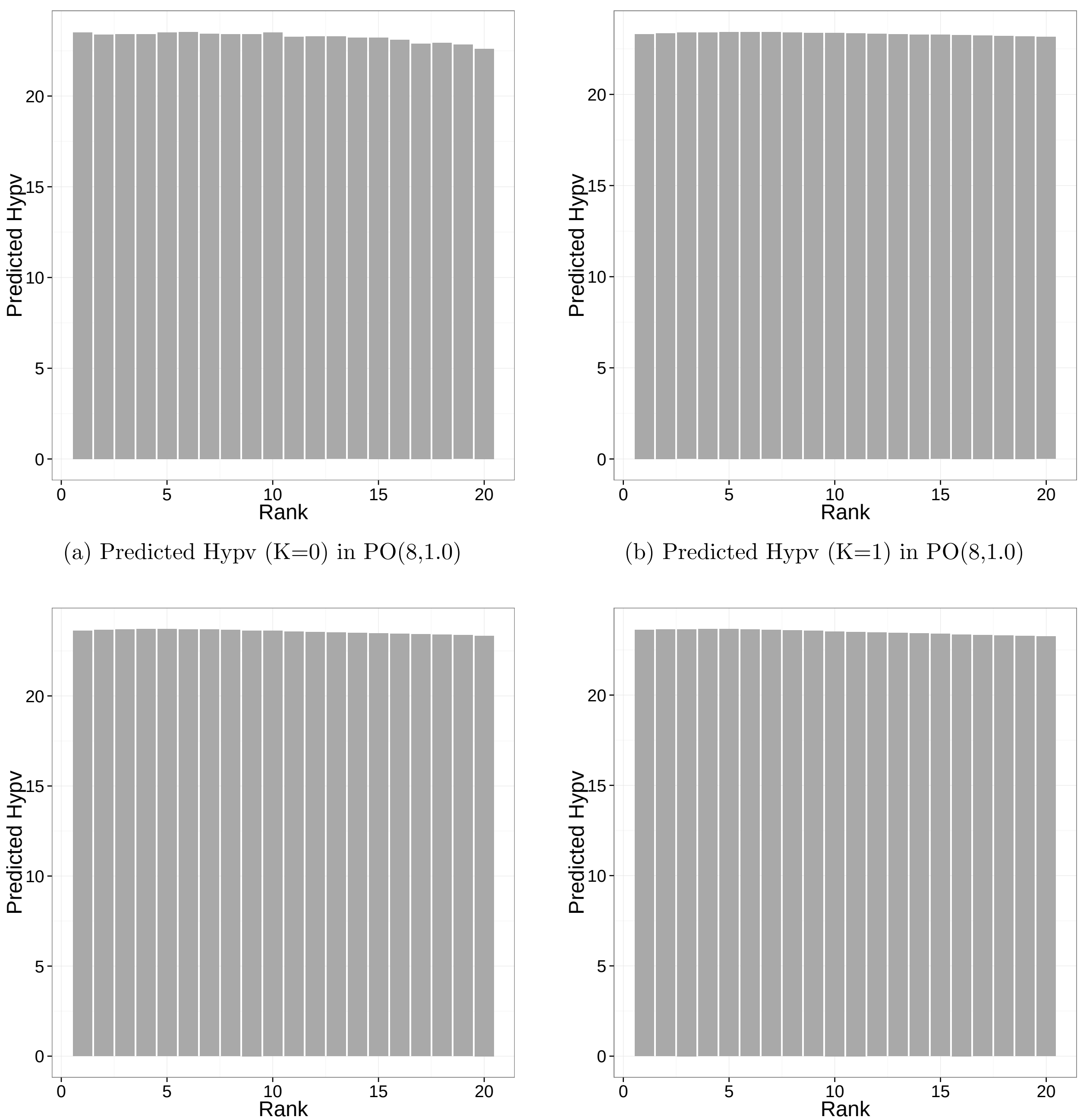


(c) Predicted Hypv (K=2) in PO(8,0.5)



(d) Predicted Hypv (K=3) in PO(8,0.5)

Figure 7.31: Bar plots showing the predicted future expected Hypv in ASMS/mHDM for the obtained ordered Pareto-flexible sets, averaged over all periods and experiments (for PO(8,0.5)).



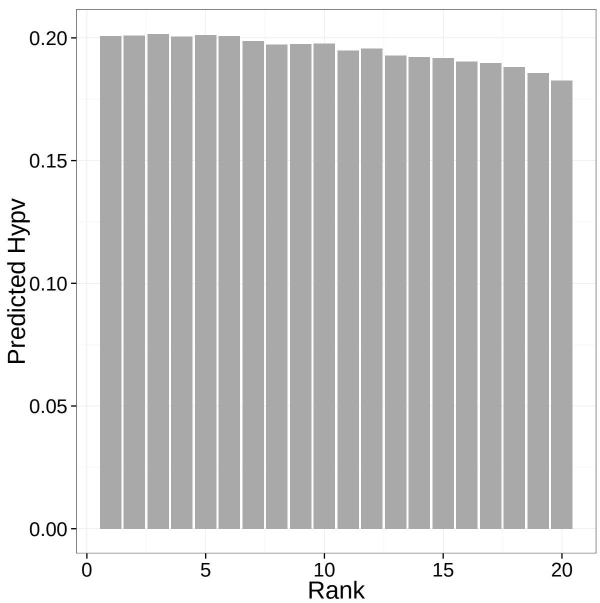
(a) Predicted Hypv (K=0) in PO(8,1.0)

(b) Predicted Hypv (K=1) in PO(8,1.0)

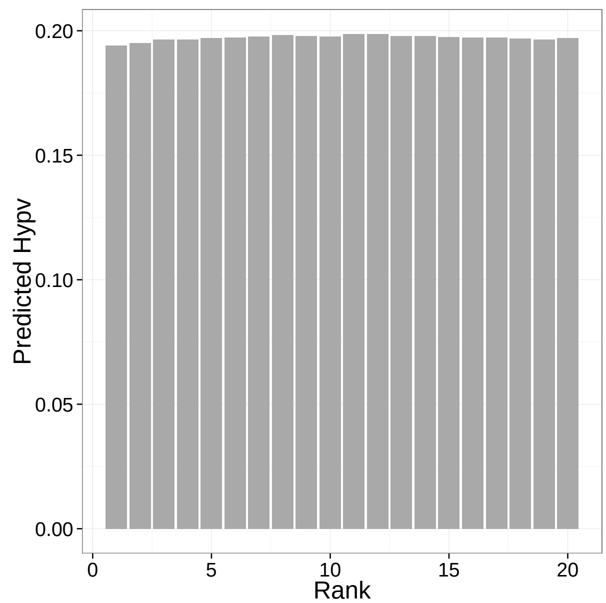
(c) Predicted Hypv (K=2) in PO(8,1.0)

(d) Predicted Hypv (K=3) in PO(8,1.0)

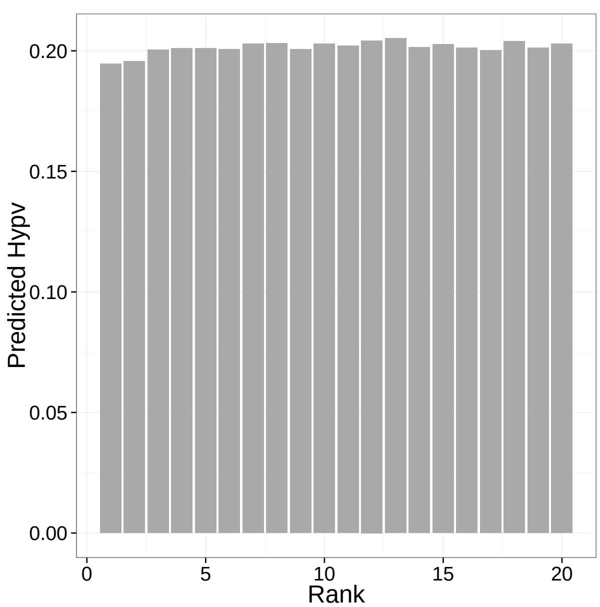
Figure 7.32: Bar plots showing the predicted future expected Hypv in ASMS/mHDM for the obtained ordered Pareto-flexible sets, averaged over all periods and experiments (for PO(8,1.0)).



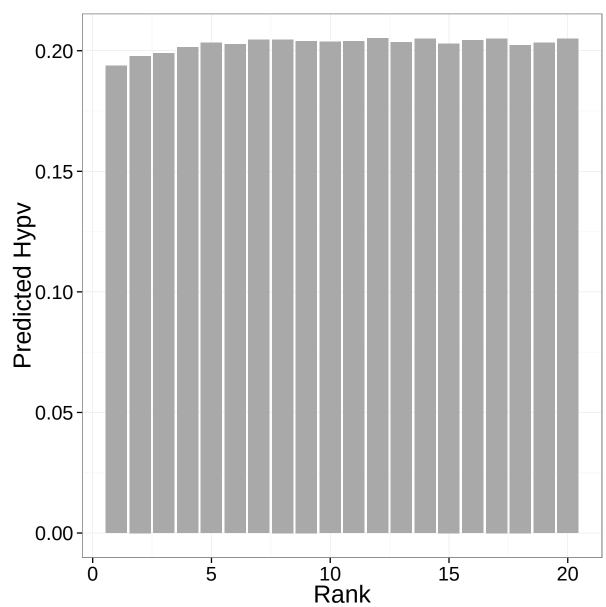
(a) Predicted Hypv (K=0) in FTSE



(b) Predicted Hypv (K=1) in FTSE

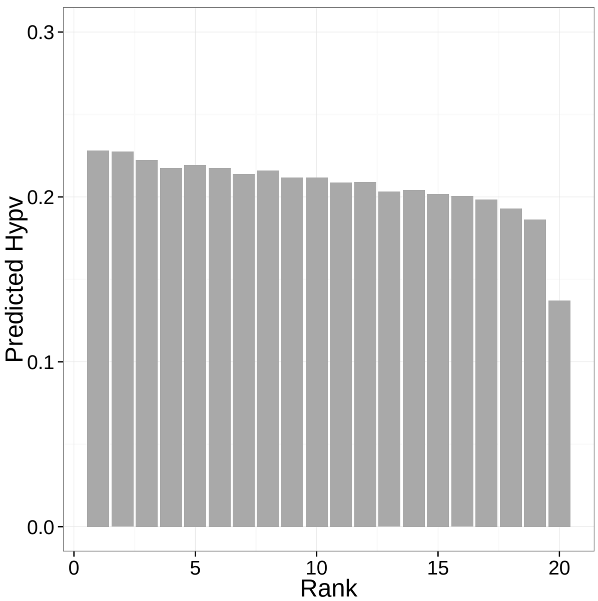


(c) Predicted Hypv (K=2) in FTSE

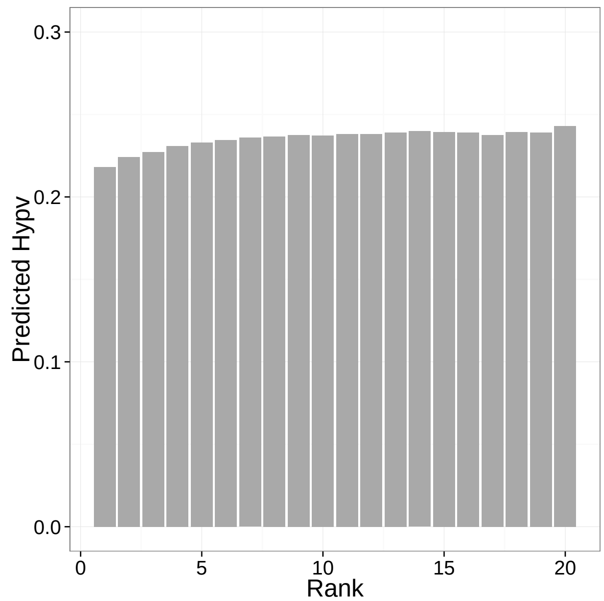


(d) Predicted Hypv (K=3) in FTSE

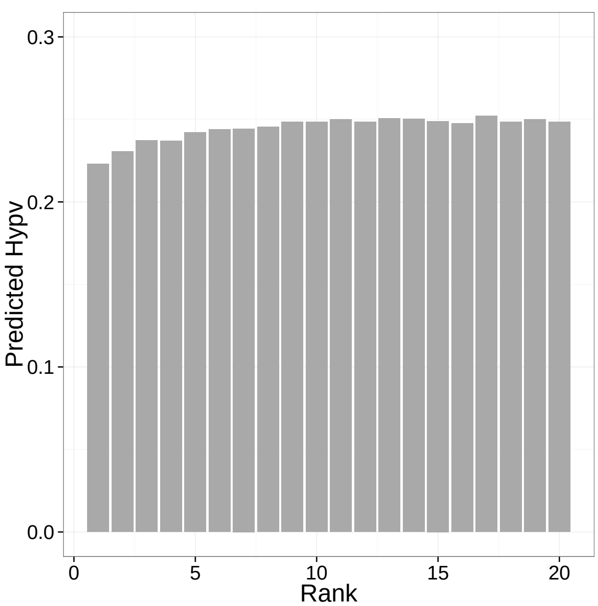
Figure 7.33: Bar plots showing the predicted future expected Hypv in ASMS/mHDM for the obtained ordered Pareto-flexible sets, averaged over all periods and experiments (for FTSE-100).



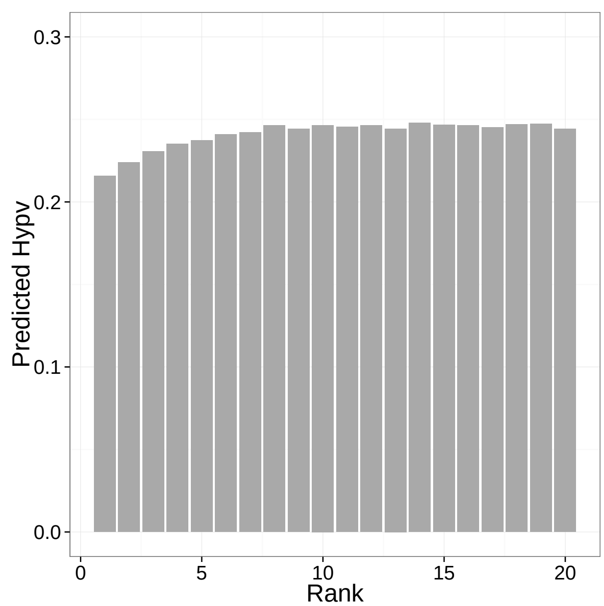
(a) Predicted Hypv (K=0) in DJI



(b) Predicted Hypv (K=1) in DJI

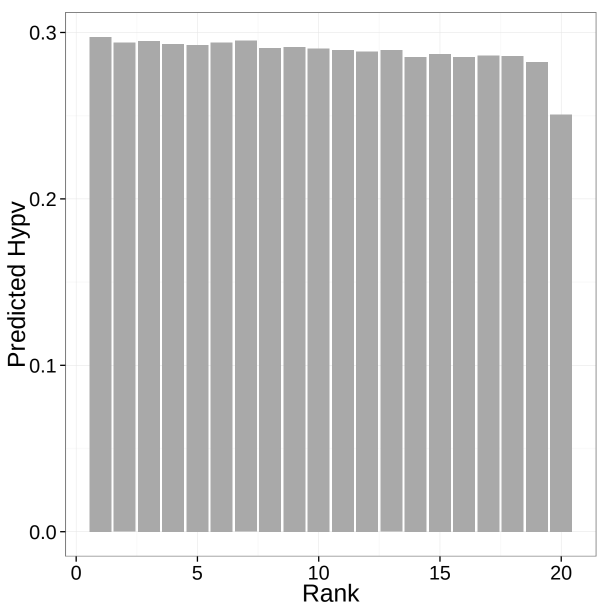


(c) Predicted Hypv (K=2) in DJI

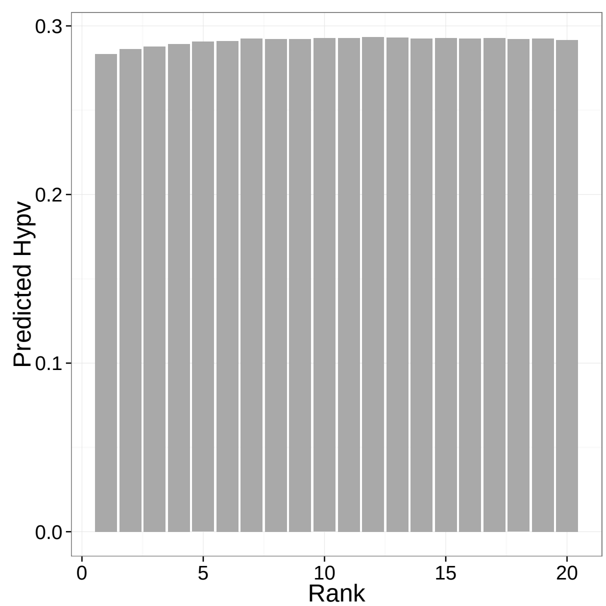


(d) Predicted Hypv (K=3) in DJI

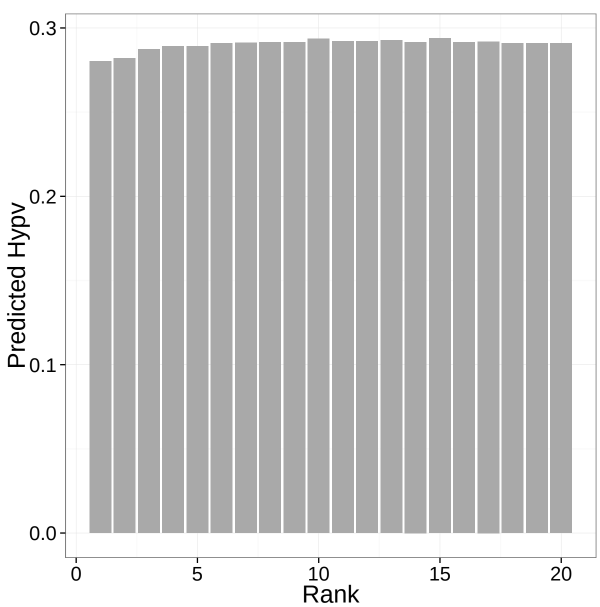
Figure 7.34: Bar plots showing the predicted future expected Hypv in ASMS/mHDM for the obtained ordered Pareto-flexible sets, averaged over all periods and experiments (for DJI).



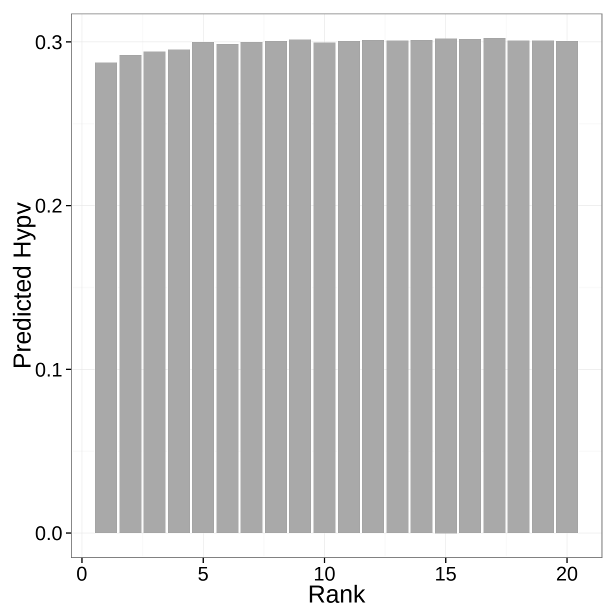
(a) Predicted Hypv (K=0) in HSI



(b) Predicted Hypv (K=1) in HSI



(c) Predicted Hypv (K=2) in HSI



(d) Predicted Hypv (K=3) in HSI

Figure 7.35: Bar plots showing the predicted future expected Hypv in ASMS/mHDM for the obtained ordered Pareto-flexible sets, averaged over all periods and experiments (for HSI).

Chapter 8

Conclusion

It is good to have an end to journey towards; but it is the journey that matters, in the end.

– Ursula Kroeber Le Guin

The opposite of love is not hate, it's indifference. The opposite of art is not ugliness, it's indifference. The opposite of faith is not heresy, it's indifference. And the opposite of life is not death, it's indifference.

– Elie Wiesel

This thesis has made the case for an anticipatory methodology to alleviate the need for disclosing preferences when selecting a solution under multiple conflicting views from a set of alternatives designed for operating in dynamic and noisy environments. Anticipating *flexible provisional options* predicted to maximize future diversity of choice was then deemed as a reasonable strategy to robustly postpone the Decision Maker (DM) preferences specification under uncertainty, therefore contributing to the *automation* of sequential Multiple Criteria Decision-Making (MCDM) processes.

Although a mix of ad-hoc techniques could be pursued for coping with both stochasticity and non-stationarity, this thesis has argued for principled models that could alleviate the design effort needed to put different techniques to operate together. A fundamental premise that motivated this research approach is that predictive knowledge could be combined, under a Bayesian framework, with the observations about past and current trade-off system states, so that multi-objective metaheuristics could regularly operate under the influence of the future expectation about the value of the candidate trade-off solutions.

That is to say, no additional ad-hoc operation would be needed for triggering anticipatory behavior in metaheuristics. The thesis then argued that one principled way to incorporate anticipation in metaheuristics is by updating the current value of a solution, i.e., its objective vector evaluation, in a way it can reflect the expected trade-off consequences of its actuation when deployed in future environments. In this way, we believe that many problems of real-world relevance can be more realistically tackled if modeled as anticipatory stochastic multi-objective ones.

To the best of our knowledge, this thesis was the first to investigate models for simultaneously handling the optimization of multiple conflicting, time-varying, cost-adjusted, and noisy objective functions. The resulting anticipatory problem solving tools were thus designed for (a) approximating flexible stochastic Pareto frontiers over time; (b) anticipating changes in multiple noisy objective functions; and (c) adjusting for costs in time-linkage settings.

From a technical standpoint, the thesis proposed the association of the online maximization of a theoretically sound set-based quality indicator in Multi-Objective Optimization (MOO) – the Hypervolume (Hypv) – with the decision-theoretic notions of preference for (a) flexibility; (b) for long-term predictability; and (c) for trajectories of minimal historical predictive error. It was then postulated that implementing solutions predicted to maximize the expected Hypv of future Pareto sets could be used as a strategy for postponing preference specification, thus automating the resolution of conflicts.

The premise was that, if preferences cannot be reliably determined in such uncertain environments, the DM would prefer implementing flexible solutions (i) predicted to keep his/her set of future options open; (ii) whose future consequences attain higher degrees of predictability; and (iii) whose historical observed consequences were similar to prior expectations.

This scenario motivated the design of online anticipatory learning methods to self-adjust time preferences in terms of the perceived temporal uncertainty. Because there are no proposals in Evolutionary MOO (EMOO) meeting the posed problem requirements, an existing Hypv-based EMOO algorithm (SMS-EMOA [32]) was extended and adapted to our methodology. Moreover, a portfolio selection random instance generator was designed for providing controllable scenarios to experimentally verify our anticipatory methodology in a problem of real-world relevance.

The main experimental study reported in this thesis in **chapter 7**, with the proposed Anticipatory $\mathcal{S}$-Metric Selection Evolutionary Multi-Objective Algorithm (ASMS-EMOA), has shown that, by performing Bayesian tracking in both the objective and the search spaces under linear dynamics assumptions, flexible portfolio rebalancing decisions leading to improved future quality and diversity of trade-off investment choices could be effectively anticipated in all tested artificial and real-world scenarios.

Given our preliminary yet successful experience with the application of our anticipatory methodology for the portfolio selection application, we hope that our proposals in ASMS-EMOA for handling noisy and time-varying environments will encourage other instantiations of algorithms and learning methods capable of dealing with AS-MOO problems. It is clear for us that the AS-MOO formulation is general enough to incorporate different estimation and prediction models, besides other principled rules to integrate additional predictive knowledge into EMOO metaheuristics. Provided that one can adequately fit manageable dynamical models (e.g. Bayesian ones), even if considering non-parametric estimation procedures such as particle filters (e.g. [75, 198]), we believe AS-MOO can become a useful model for real-world multistage stochastic multi-objective optimization problems.

In the next section, the main results and contributions of this thesis are summarized, whereas a discussion on the limitations of the proposed models and problem solving tools, as well on future research paths, are described in the final sections.

8.1 Main Results and Contributions

The main original contributions of this thesis are thus enumerated as follows:

Contribution 1 An open-ended research question of how to make decisions when the DM preferences cannot be reliably elicited has been proposed. Before this thesis, most works in the literature have taken a different path, trying instead to address the question of how the DM preferences could be approximated or partially represented despite inaccuracies in the elicitation process e.g. [21]. Therefore, to the best of our knowledge, we can say that this thesis provides the first attempt to mitigate the need for eliciting preferences without compromising the range and quality of future decision paths. In other words, our proposals provide the means of acting robustly under missing preferences in *uncertain environments*;

Contribution 2 Decision-theoretic accounts of preference for flexibility [140] have been integrated into Multiple Criteria Decision-Making (MCDM). Although flexibility and freedom of choice [169] are classical concepts in areas such as economics and social choice, it was not clear before this thesis how they could be accommodated in MCDM approaches;

Contribution 3 A model for representing problems wherein multiple conflicting noisy optimization criteria must be simultaneously improved over time has been proposed. This was named the Anticipatory Stochastic MOO (AS-MOO) model;

Contribution 4 The AS-MOO model has been formulated in terms of a set-based quality indicator known as Hypervolume. This indicator was not only shown to be consistent with the notions of preference for flexibility and diversity of choice, but has also explicitly allowed for a MOO model whose search space is the set of all finite subsets of the original feasible search space. The AS-MOO model thus falls in the category of set-based MOO [227]. Moreover, it is a recurrence EMOO/MCDM model whose solution for the current decision period depends on the solutions for future stages, what requires predictive approaches to approximate AS-MOO solutions;

Contribution 5 Bayesian tracking has been integrated into both the objective and the search spaces of the so-called ASMS-EMOA proposal. While an external predictive approach could have been used to directly track stock prices in the portfolio selection application (e.g. [89]), this thesis favored Rosen's view of anticipatory systems [174] as those which are provided with *internal* predictive models. More precisely, we provided ASMS-EMOA with the ability of (a) independently tracking the cost-adjusted performance of individual portfolios in the objective space with a Kalman Filter, and of (b) tracking the changes in portfolio compositions for different portions of the stochastic Pareto frontier with a dynamical Dirichlet distribution maximum a posteriori estimation procedure. By doing so, ASMS-EMOA was able to intensify the search for more *predictable* portfolio compositions, what would not be possible if tracking all assets prices over time at once;

Contribution 6 The concept of Online Anticipatory Learning (OAL) has been devised to self-adjust temporal preferences in response to the estimated uncertainty about the Pareto dominance relations between current and predictive objective function distributions.

Besides the six aforementioned contributions, we can highlight other additional original advancements that emerged from this thesis:

Advancement 1 Preliminary experiments in **chapter 4** with a preemptive, non-anticipatory approach were designed to assess whether the hypervolume indicator could accommodate regularization terms in order to evolve portfolios exhibiting desirable characteristics in deterministic environments such as: improved diversification; lower cardinality; and higher stability between investment rounds. The results were positive in this regard, and the proposed regularized hypervolume selection algorithm (RSMS-EMOA) was able to approximate deterministic Pareto frontiers in dynamic portfolio selection environments;

Advancement 2 A time-linkage free AS-MOO model was also conceived in **chapter 5**. In Azevedo and Von Zuben [15], a version of ASMS-EMOA using only the Kalman Filter for tracking objective vectors was designed. Preliminary experiments showed that anticipatory learning can also be effective to handle the case where costs do not incur from past choices over the changing objective functions [15];

Advancement 3 In **chapter 7**, a random instance generator was devised for multistage portfolio selection. In Engle [80], artificial dynamical models representing changes in stock prices were also investigated, although using more complicated non-linear forms. In this thesis, we have focused on linear trends and abrupt changes in the joint asset return distributions, what required the generation of random covariance matrices. To the best of our knowledge, this is the first artificial test suite proposed in the literature that can support the assessment of dynamic stochastic MOO algorithms;

Advancement 4 A dynamical Dirichlet distribution maximum a posteriori estimation procedure was presented in **chapter 6**, for tracking mean vectors in a simplex search space, adapted from Bertucelli and How [30]. We envision that this procedure can be useful for other applications in machine learning, such as tracking changing social networks topic models in text mining applications (e.g. [17]);

Advancement 5 One of the main innovations in the proposed ASMS-EMOA was an analytical expression for the computation of the *expected anticipatory hypervolume contributions* of each candidate solution, under the assumptions that the objective vectors can be modeled as multivariate Gaussians. We consider this an important step for handling stochastic multi-objective optimization in general with hypervolume-based methods by taking advantage of existing deterministic state-of-the art EMOO algorithms (SMS-EMOA [32]). We recall that the pairwise covariance terms appearing in Theorem 6.3.1 merge trade-off information between multiple neighboring anticipatory distributions in the objective space. The expected anticipatory $\mathcal{S}$-metric contribution thus allows for the *mutual interference* and *exchange of trade-off information* between parallel neighboring anticipatory paths.

For a list of all detailed minor contributions, see the summaries at the end of each chapter.

Despite the seemingly intractable quest of simultaneously coping with various sources of uncertainty, our premise was that research on dynamic stochastic multi-objective models can dramatically leverage the effective automation of real-world decision-making systems. Furthermore, we believe that valuable insights can be gained when coordinating different uncertainty handling techniques to operate together, as it was the case with the experiments described in **chapter 7**. Our hope is that this thesis may have contributed to those goals.

8.1.1 Final Philosophical Remarks

One thing that is noteworthy from the experimental results presented in **chapter 7** is that, while there exists a powerful theoretical framework – the Bayes theorem – to rationally update prior probabilistic beliefs in face of new evidence in order to reason about the real world, there is no such framework when it comes to decision-making. In other words, an autonomous agent cannot simply regret and backtrack from decisions implemented in the past upon observing negative rewards, unless he/she is willing to pay the resulting costs.

What is interesting when assessing the usefulness of the Bayesian approach to reason under uncertainty, is that, when revising prior beliefs by processing contradictory evidence, there is usually no cost involved, except perhaps that of acquiring more information. While data acquisition may be very costly in certain contexts (e.g. in oil prospecting over international waters), the use of Bayesian computational methods to *approximately* update the existing predictive models upon the receival of new data is virtually cost-free[1]. On the other hand, when revising previously taken decisions and plans in execution, there are usually high costs involved, as it is the case with the portfolio selection domain.

Therefore, since decision makers cannot always completely change past decisions, and based on the experimental findings of **chapter 7**, it may be reasonable to state that, under severe uncertainty, it might not be enough to rely on Bayes rule to search for the next "optimal" action. This thesis thus made the case that a wiser decision maker would be better off in highly uncertain environments by taking advantage of his/her prior beliefs and of available predictive knowledge for effectively anticipating flexible choices, hence keeping open his/her future set of options. In this way, each evolved multiple plan and learned anticipatory trajectory in the proposed ASMS-EMOA can be interpreted both (1) as a hedge against future failure; and (2) as an attempt of augmenting the autonomous capabilities of metaheuristics.

8.2 Limitations of the Anticipatory Methodology

Despite the promising results obtained with the proposed anticipatory methodology in terms of uncertainty handling in EMOO/MCDM by improving the DM future diversity of choice, the successful widespread application of the AS-MOO models and algorithms will require competent system identification and proper modeling assumptions that may capture the dynamics of the problem under consideration. Thus, the problem solving tools devised in this thesis are most

[1] *Exact* Bayesian inference has nevertheless exponential computational cost.

appropriate for problems where e.g. the linear Gauss-Markov assumption may be a good fit for tracking dynamic random objective vectors in the performance space of an EMOO algorithm.

Besides, the current version of ASMS-EMOA is only applicable for problems whose search space is defined over the $(d-1)$–Simplex, where d is the number of search variables, although many real-world applications require optimization over the simplex (see **chapter 1**).

As far as the computational cost of the proposed tools goes (bilinear complexity in terms of the number of Pareto-efficient solutions and of the anticipation horizon), we consider it to be manageable when compared to the exponential complexity that would be required for obtaining near-optimal solutions to the AS-MOO model. In fact, a typical run of ASMS-EMOA (which was compiled to native Microsoft Windows code from the C++ language) for the most demanding dataset (FTSE with 87 available assets, executed with a population size of 20 Pareto-efficient portfolios, a window size of $K = 3$, an anticipation horizon of $H = 2$, and using the hypervolume-based decision-maker) – using only one processor of an Intel i7 machine running at 3.5 GHz – required about 3 hours of total non-dedicated processing time in a Microsoft Windows 8 operational system, for the whole number of 24 investment environments. On average, that means about 7.5 minutes per period were required to evolve an anticipatory approximation to the stochastic Pareto frontier and to identify the maximal flexible investment choice. Naturally, we believe that more efficient implementations and more cleaver usage of algorithms and data structures could substantially reduce this running time.

On the other hand, such narrow scope and coding limitations provides several research opportunities for improving the anticipatory EMOO models and algorithms proposed in this thesis and to extend them for other application scenarios cases.

8.3 Suggestions of Further Research on AS-MOO Solvers

The proposed ASMS-EMOA can be extended in a straightforward way for handling more than two optimization criteria. The drawback of doing so is the increased computational cost to compute the covariances between pairs of objective function evaluations required for approximating the expected hypervolume contribution of each candidate solution (see **chapter 6**). Besides, additional techniques would be required to handling the loss of selection pressure, as the probability of two arbitrary candidate solutions being mutually non-dominated approaches one with the increase in the number of objectives to be optimized. Moreover, novel anticipatory learning rules to incorporate the estimated predictive uncertainty and past performance errors into the values of the candidate solutions are definitely worth investigating. In the following, we discuss the main envisioned research paths that could advance even further AS-MOO applications.

8.3.1 Anticipating Future Preferences

We recall from **chapter 5** that there is no way envisioned in AS-MOO to compute in advance the future Preferred Feasible Regions (PFRs), and the assumption in AS-MOO is that the PFRs remain constant at future stages. We consider the pursuit of a dynamical behavioral model for the future partial preferences specification of a DM as an interesting open research question. Would the DM be better off by narrowing down his/her preferences when the outcomes resulting from the decisions taken are *close* to what he/she had been expecting? Conversely, should the DM instead flexibilize his/her preferences upon surprising, unforeseen outcomes? It is our opinion that an strategy for dynamically linking preferences specification with feedback error from the utilized Bayesian tracking tools could be worth investigating for automating the MCDM process and for using the anticipatory multi-objective capabilities of an AS-MOO solver in order to cope with such uncertain outcomes.

The investigation of behavioral models that could explain how a DM would adjust the PFR over time as a function of the observed past outcomes is therefore an interesting open problem. Despite the absence of such a model, in the approximate AS-MOO solvers proposed in this thesis, this challenge could be handled in the same way as in how the objective vectors and trade-off solutions are predicted: by using Bayesian predictive models for tracking the coefficients appearing in the linear equations describing the hyperplanes encoded in the proposed PFR representations. If any trend regarding the DMs preferences evolution is captured (e.g. progressively narrowing preferences down towards specific directions), this information could be in principle captured for enabling the AS-MOO solver to obtain approximate ϵ-feasible maximal flexible choices w.r.t. the predicted future PFRs.

8.3.2 Regret Bounds for Online Hypervolume Maximization

The AS-MOO model can be interpreted under a more general MOO framework of online hypervolume maximization. Let then $\mathbf{f}_1, \cdots, \mathbf{f}_T$ be a sequence of vector-valued objective functions, where $\mathbf{f}_t = (f_1^{(t)} \cdots f_m^{(t)})^\intercal$. For any candidate set of μ mutually non-dominated solutions at time t, $\mathbf{U}_t^\mu = \left\{\mathbf{u}_t^{(1)}, \cdots, \mathbf{u}_t^{(\mu)}\right\}$, the associated hypervolume value is $\mathcal{S}(\mathbf{U}_t^\mu)$. A general online MOO hypervolume maximization framework can be described as:

1. For each $t = 1, 2, \cdots$, do
   (a) Choose a set of μ mutually non-dominated solutions, $\hat{\mathbf{U}}_t^\mu \in \Omega^\mu$
   (b) Receive m objective-functions $\mathbf{f}_t : \Omega \mapsto \mathcal{R}^m$
   (c) Observe the hypervolume $\mathcal{S}_t(\hat{\mathbf{U}}_t^\mu)$

Given a fixed learning rate η, by using recently derived analytical expressions for the gradient of the hypervolume in Emmerich et al. [79], one could update the candidate trade-off set using a stochastic gradient rule:

$$\hat{\mathbf{U}}_{t+1}^\mu = \mathbf{U}_{t-1}^\mu - \eta \nabla \mathcal{S}_t(\mathbf{U}_t^\mu). \tag{8.1}$$

The *regret* term is then defined as the difference between the optimal hypervolume value that would be obtained for the best fixed set of mutually non-dominated solutions for all decision periods and the sequence output by such online stochastic gradient update rule:

$$\mathrm{Regret}_T\left(\left\{\hat{\mathbf{U}}_1^\mu, \cdots, \hat{\mathbf{U}}_T^\mu\right\}\right) = \max_{\mathbf{U}^\mu \subset \Omega} \sum_{t=1}^{T} \mathcal{S}_t(\mathbf{U}^\mu) - \sum_{t=1}^{T} \mathcal{S}_t(\hat{\mathbf{U}}_t^\mu). \tag{8.2}$$

Although the stochastic gradient online rule is a myopic approach, the theoretical study of the attainable regret bounds could dramatically advance the incipient field of online multi-objective optimization. While an anticipatory learning approach would be in principle even harder to analyze from a theoretical perspective, the experimental observed regret would still be worth investigating. The main difficulty for a regret analysis is the fact that the hypervolume maximization leads to a non-convex optimization problem (see **chapter 3**). Nevertheless, provided that all objective functions are convex, there are convex reduction procedures that could be in principle applicable for approximating the optimal hypervolume locally, with convergence guarantees. We are currently studying such issues in order to advance research on *anticipatory online multi-objective optimization*.

8.3.3 Anticipatory Anomaly Prevention

With the proposed flexible anticipatory MCDM framework, it is possible to estimate the probability of any current available trade-off option becoming unavailable (i.e., becoming non-essential or dominated) at future decision stages. This information could be used to trigger recommended anticipatory actions for a human decision-maker to remove, at his/her own will, such options in an interactive MCDM framework. Therefore, applications of *multiple criteria anomaly detection* in data streams could benefit from the proposed AS-MOO methodology when there are multiple conflicting views for describing what is an anomalous pattern in applications such as communication networks quality of service monitoring and fraud detection.

8.3.4 Anticipatory Resource Allocation

Moreover, many engineering problems arising in diverse areas can be modeled as a resource allocation problem and described in terms of an anticipatory stochastic dynamic multi-objective optimization problem in *simplex spaces*. Examples include load balancing in complex network and cloud computing environments and vendor-managed inventory routing problems.

In addition, machine learning model building applications that include training multiple classifiers/predictors in an ensemble could also benefit from the AS-MOO methodology applied in simplex spaces. One simple way to combine conflicting classifiers and predictors outputs is to combine such outputs as a convex combination.

8.3.5 Anticipatory Hyper-Heuristics

The ASMS-EMOA algorithm operating over simplex spaces can also be straightforwardly turned into a *hyper-heuristic* for allocating (or investing) computational resources to different heuristics/algorithms operating jointly to solve any given search/decision-making problem by e.g. simultaneously minimizing the expected variance (risk) and maximizing the expected average trade-off heuristic quality. Thus, we foresee that the AS-MOO methodology could allow for a general artificial intelligence *cognitive framework*, supported by flexible anticipatory multi-objective hyper-heuristics for autonomously *generating and investing* computational resources in evolved flexible heuristics for changing and noisy environments.

8.3.6 Automation of Flexible and Sustainable Planning Systems

The term "Sustainable development" has been defined by the World Commission on Environment and Development as

> (...) "*development that meets the needs of the present without compromising the ability of future generations to meet their own needs.*"

The main topic of inquiry envisioned in this thesis for the continuty of research on the automation of sequencial MCDM processes regards the study of the observed trade-offs which may emerge from the *simulation* of multiple conflicting planning *meta-preferences*, namely, *preference for flexibility* and *preference for sustainability.* This investigation is under way.

For instance, research on the automation of *flexible* production/distribution supply chains has been gaining relevance as consumer market demands shift *rapidly and unpredictably* in response to increasing pressures toward (a) fine-grained mass customization; and (b) *accelerated* distribution of products and services. Nevertheless, the social, environmental, and economical long-term costs of flexibilizing production/distribution plans – by means of industrial automation – have also been leading to increasing *sustainability* concerns. The rising real-world *tension* between flexibility and sustainability can be therefore pointed out as a *fundamental motivator* for the continuity of the research conveyed in this thesis.

Starting from a supply chain of N_p producer *and* N_c consumer nodes, a feasible future research goal promoted in this thesis is to foster *open-ended* studies and investigations related to the following sequence of four interdisciplinary and partially unexplored research questions:

1. How can *preference for flexibility* and *preference for sustainability* be axiomatically defined?

2. How to measure the flexibility and sustainability levels of a given sequence of production/distribution decision outcomes as a function of the realized demand?

3. How to design novel theoretically sound AS-MOO *models and algorithms* leading to improved flexibility and sustainability levels in complex production/distribution supply chains under uncertainty?

4. To what extent sustainability maximization *regulates* flexibility maximization (and vice-versa)?

Assuming from question (4) that *causality* between sustainability maximization and flexibility maximization can be established with statistical significance, investigations related to the following additional *long-term* research question can be addressed:

5. Is it possible *in principle* to find a *control law* that outputs flexibility and sustainability levels leading to the (not necessarily monotonic) minimization of the distance to a reference *well-balanced joint wealth allocation* over N_p producers *and* N_c consumers?

Complex supply chains must cope with the underlying uncertainty about future decision outcomes in environments subject to statistical fluctuations and time-varying conditions. The promotion of flexibility in such systems can be thus intuitively justified as *an adaptive response*, on the production side, to *unpredicability*. Sustainability can be, by its turn, justified as *a planned response*, on the consumer end, *to the costs of the accelerated flexibilization of production/distribution*.

On the other hand, whenever the predicability levels of future consumer demands are *reasonable*, i.e., when sound logical judgments can be performed in spite of eventual innacuracies, *reasoning* over multiple conflicting choice future consequences can provide valuable guidelines for not only (a) *resolving the underlying conflicts* between flexible and sustainable planning, but also for (b) reducing the mismatch between production/distribution and consumer goals.

8.3.7 Sustainable Planning: Models, Measurements, and Concepts

Specific interpretations of sustainability accounts it as a planning approach to "(...) *integrate economic, environmental, social and increasingly institutional issues as well as to consider their interdependencies*" [94]. Moreover, the following main aspects of planning approaches that promote sustainability were identified in Gasparatos et al. [94]:

1. The consideration of *future consequences* of present actions;

2. The acknowledgment of *uncertainties* concerning choice future consequences, acting accordingly to a precautionary strategy;

3. The engagement of the public; and

4. The integration of equity considerations, of both "*intragenerational and intergenerational*" nature.

In Hutchins and Sutherland [122], *social sustainability* decision criteria were suggested. This thesis envisions that such indicators can be integrated within mathematical multi-objective optimization models supporting automated planning decisions from input historical data. We also believe that controlling for faulty processes and machinery failures can be key to optimize such criteria. Three such indicators are [122]: (a) reduction of the number of personal work accidents; (b) reduction of the number of disease incidence on workers; and (c) and number of days of workers absence do to injuries.

The trade-offs between decision criteria regarding the *social sustainability* and economical performance in Supply Chain Network Design (SCND) have been experimentally by e.g. the recent work of Devika el a. [73]. The impact of planning for maximizing the number of future potential jobs expected to be created by candidate SCN designs on profit was then investigated using Monte Carlo simulation of the demand and bio-inspired meta-heuristics for optimizing such designs.

A many-objective stochastic optimization model promoting the sustainable planning of long-term investments on a water distribution system portfolio in the region of Grande Valley, Texas, has been experimentally assessed in Kasprzyk et al. [130]. In that work, Monte Carlo simulation was used along with 33-years historical data on water demand for predicting the future long-term drought risks of candidate investment allocations for a 10-years ahead planning. The trade-offs between *drought risks reduction* and other conflicting objectives such as costs reduction could be visualized, what greatly aided the investment choice processes of human decision-makers.

Environmental sustainability has been also recently addressed within a multi-objective linear programing study in Soysal et al. [196] in the context of food logistics (distribution) networks using historical data and the estimated CO2 emission per distance unit traveled by several truck fleet options. In addition, several estimated toxic residuals resulting from different types of food processes production were set to be minimized. Profit and costs criteria were shown to be in conflict with both CO2 emission and toxic residuals reducement.

Finally, we highlight interesting observations in Buchholz et al. [42], wherein 35 candidate subjective sustainability criteria for planning both (a) bioenergy distribution networks; and (b) bioenergy investments were evaluated by 137 experts:

- "*Only two criteria, energy balance and greenhouse gas balance, were perceived as critical by more than half of the respondents*";
- "*Seven of the 12 criteria scored as most important focused on environmental issues, four were social and only one was economic*";
- "*The spatial scale the experts worked at and their profession explained most of the differences in importance ranking between experts, while regional focus had minimal effect*"; and
- "*Criteria that were ranked low for importance, were characterized by a lack of consensus, suggesting the need for further debate regarding their inclusion in sustainability assessments*".

From those observations, the lack of consensus is striking because, even in a relatively well-defined niche area in which sustainable considerations are pervasive, several psychological factors arising from the respondents positions, levels of expertise, and professions may deeply influence their perspectives and set of preferences toward defining a very much pursued sustainable agenda for bioenergy production and distribution systems.

The main conclusion of Buchholz et al. [42] is this: "*The selection of these criteria can vary depending on individual's expertise, geographical region where they work, and spatial scale they are focused on.*"

It is then clear that, despite incipient efforts to define domain-specific sustainable definitions and measurements, a higher-level axiomatic account of *preference for sustainability* can greatly contribute to more sound optimization models and algorithms promoting the sustainable automation of flexible multi-criteria planning systems.

References

[1] F. B. Abdelaziz. Solution approaches for the multiobjective stochastic programming. *European Journal of Operational Research*, 216(1):1 – 16, 2012.

[2] F. B. Abdelaziz, B. Aouni, and R. E. Fayedh. Multi-objective stochastic programming for portfolio selection. *European Journal of Operational Research*, 177(3):1811 – 1823, 2007.

[3] C. C. Aggarwal. *Data Streams: Models and Algorithms*. Springer, 1st edition, 2007.

[4] F. A. Alqallaf, K. P. Konis, R. D. Martin, and R. H. Zamar. Scalable robust covariance and correlation estimates for data mining. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery*, pages 14–23. ACM, 2002.

[5] K. P. Anagnostopoulos and G. Mamanis. A portfolio optimization model with three objectives and discrete variables. *Computers & Operations Research*, 37:1285–1297, 2010.

[6] C. Aranha, C. R. B. Azevedo, and H. Iba. Money in trees: How memes, trees, and isolation can optimize financial portfolios. *Inf. Sci.*, 182(1):184–198, 2012.

[7] D. Ariely. *Predictably irrational*. HarperCollins New York, 2008.

[8] R. Arlegi and J. Nieto. Incomplete preferences and the preference for flexibility. *Mathematical Social Sciences*, 41(2):151 – 165, 2001.

[9] A. Auger, J. Bader, D. Brockhoff, and E. Zitzler. Theory of the hypervolume indicator: optimal μ-distributions and the choice of the reference point. In *Proceedings of the tenth ACM SIGEVO workshop on Foundations of genetic algorithms*, pages 87–102. ACM, 2009.

[10] A. Auger, J. Bader, D. Brockhoff, and E. Zitzler. Hypervolume-based multiobjective optimization: Theoretical foundations and practical implications. *Theoretical Computer Science*, 425:75–103, 2012.

[11] A. Azadeh, M. H. Farahani, H. Eivazy, S. Nazari-Shirkouhi, and G. Asadipour. A hybrid meta-heuristic algorithm for optimization of crew scheduling. *Applied Soft Computing*, 13(1):158 – 164, 2013.

[12] C. R. B. Azevedo and A. F. R. Araujo. Generalized immigration schemes for dynamic evolutionary multiobjective optimization. In *IEEE Congress on Evolutionary Computation*, pages 2032–2039, 2011.

[13] C. R. B. Azevedo and A. F. R. Araújo. Non-dominance landscapes for guiding diversity generation in dynamic environments. In *VIII Best MSc Dissertation/PhD Thesis Contest in Artificial Intelligence (CTDIA). Proceedings of The Brazilian Conference on Intelligent System*, 2012.

[14] C. R. B. Azevedo and F. J. Von Zuben. Anticipatory stochastic multi-objective optimization for uncertainty handling in portfolio selection. In *Evolutionary Computation (CEC), 2013 IEEE Congress on*, pages 157–164, June 2013.

[15] C. R. B. Azevedo and F. J. Von Zuben. Regularized hypervolume selection for robust portfolio optimization in dynamic environments. In *Evolutionary Computation (CEC), 2013 IEEE Congress on*, pages 2146–2153, June 2013.

[16] J. Bader and E. Zitzler. Robustness in hypervolume-based multiobjective search. Technical Report 317, Computer Engineering and Networks Laboratory (TIK), ETH Zurich, Zurich, 2010.

[17] A. Banerjee and S. Basu. Topic Models over Text Streams: A Study of Batch and Online Unsupervised Learning. In *Proceedings of the SIAM International Conference on Data Mining (SDM)*, pages 437–442, 2007.

[18] A. Banerjee, I. S. Dhillon, J. Ghosh, and S. Sra. Clustering on the unit hypersphere using Von Mises-Fisher distributions. *J. Mach. Learn. Res.*, 6:1345–1382, Dec. 2005.

[19] A. Banerjee and H. Shan. Latent Dirichlet conditional naive-bayes models. In *Data Mining, 2007. ICDM 2007. Seventh IEEE International Conference on*, pages 421–426, 2007.

[20] A. K. Barbey, F. Krueger, and J. Grafman. Structured event complexes in the medial prefrontal cortex support counterfactual representations for future planning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521):1291–1300, 2009.

[21] R. Battiti and A. Passerini. Brain-computer evolutionary multiobjective optimization: a genetic algorithm adapting to the decision maker. *IEEE Trans. Evol. Comp*, 14(5):671–687, Oct. 2010.

[22] H. H. Bauschke and J. M. Borwein. Joint and separate convexity of the bregman distance. In Y. C. Dan Butnariu and S. Reich, editors, *Inherently Parallel Algorithms in Feasibility and Optimization and their Applications*, volume 8 of *Studies in Computational Mathematics*, pages 23 – 36. Elsevier, 2001.

[23] R. Bellman. A markovian decision process. *Indiana University Mathematics Journal*, 6(4):679–684, 1957.

[24] I. Bengtsson, K. Zyczkowski, et al. *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge University Press, 2006.

[25] S. Benjaafar, T. L. Morin, and J. J. Talavage. The strategic value of flexibility in sequential decision making. *European Journal of Operational Research*, 82(3):438 – 457, 1995.

[26] C. H. Bennett. Notes on Landauer's principle, reversible computation, and Maxwell's demon. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics*, 34(3):501–510, 2003.

[27] N. Bertschinger, E. Olbrich, N. Ay, and J. Jost. Autonomy: An information theoretic perspective. *Biosystems*, 91(2):331 – 345, 2008. Modelling Autonomy Modelling Autonomy.

[28] D. P. Bertsekas and J. N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, 1st edition, 1996.

[29] D. Bertsimas and R. Shioda. Algorithm for cardinality-constrained quadratic optimization. *Computational Optimization and Applications*, 43(1):1–22, 2009.

[30] L. F. Bertuccelli and J. P. How. Estimation of non-stationary Markov Chain transition models. In *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pages 55–60, 2008.

[31] N. Beume, C. M. Fonseca, M. López-Ibáñez, L. Paquete, and J. Vahrenhold. On the complexity of computing the hypervolume indicator. *Evolutionary Computation, IEEE Transactions on*, 13(5):1075–1082, 2009.

[32] N. Beume, B. Naujoks, and M. Emmerich. SMS-EMOA: Multiobjective selection based on dominated hypervolume. *European Journal of Operational Research*, 181:1653–1669, 2007.

[33] Z. Bingul. Adaptive genetic algorithms applied to dynamic multiobjective problems. *Applied Soft Computing*, 7(3):791 – 799, 2007.

[34] L. Boltzmann. *Vorlesungen über Gastheorie*, volume 1. Verlag von Johann Ambrosius Barth, 1896.

[35] P. A. N. Bosman and H. La Poutré. Inventory management and the impact of anticipation in evolutionary stochastic online dynamic optimization. In *IEEE Congress on Evolutionary Computation*, pages 268–275. IEEE, 2007.

[36] P. A. N. Bosman and H. La Poutré. Learning and anticipation in online dynamic optimization with evolutionary algorithms: the stochastic case. In *Proceedings of the 9th annual conference on Genetic and evolutionary computation*, GECCO '07, pages 1165–1172, New York, NY, USA, 2007. ACM.

[37] S. P. Boyd and L. Vandenberghe. *Convex Optimization.* Berichte über verteilte messysteme. Cambridge University Press, 2004.

[38] J. Branke and D. Mattfeld. Anticipation and flexibility in dynamic scheduling. *International Journal of Production Research*, 43(15):3103–3129, 2005.

[39] J. Branke, M. Middendorf, G. Noeth, and M. M. Dessouky. Waiting strategies for dynamic vehicle routing. *Transportation Science*, 39(3):298–312, 2005.

[40] J. P. Brans and P. Vincke. A preference ranking organisation method: (the PROMETHEE method for multiple criteria decision-making). *Management Science*, 31(6):pp. 647–656, 1985.

[41] K. Bringmann and T. Friedrich. Approximation quality of the hypervolume indicator. *Artif. Intell.*, 195:265–290, Feb. 2013.

[42] T. Buchholz, V. A. Luzadis, and T. A. Volk. Sustainability criteria for bioenergy systems: results from an expert survey. *Journal of Cleaner Production*, 17, Supplement 1(0):S86 – S98, 2009. International Trade in Biofuels.

[43] F. Bunea, A. B. Tsybakov, and M. H. Wegkamp. Aggregation for Gaussian regression. *The Annals of Statistics*, 35(4):1674–1697, 2007.

[44] K. M. Burjorjee. Explaining optimization in genetic algorithms with uniform crossover. In *Proceedings of the twelfth workshop on Foundations of genetic algorithms XII*, FOGA XII '13, pages 37–50, New York, NY, USA, 2013. ACM.

[45] E. K. Burke, M. Gendreau, M. Hyde, G. Kendall, G. Ochoa, E. Özcan, and R. Qu. Hyper-heuristics: A survey of the state of the art. *Journal of the Operational Research Society*, 64(12):1695–1724, 2013.

[46] J. Butans. *Addressing real-time control problems in complex environments using dynamic multi-objective evolutionary approaches.* PhD thesis, Cranfield University, 2011.

[47] M. V. Butz, O. Sigaud, and P. Gerard. Anticipatory behavior: Exploiting knowledge about the future to improve current behavior. In M. Butz, O. Sigaud, and P. Gerard, editors, *Anticipatory Behavior in Adaptive Learning Systems*, volume 2684 of *Lecture Notes in Computer Science*, pages 1–10. Springer Berlin Heidelberg, 2003.

[48] W. B. Cannon. *Bodily changes in pain, hunger, fear, and rage.* New York: Appleton-Century-Crofts, 1929.

[49] Y. Cao, T. Jiang, and T. Girke. A maximum common substructure-based algorithm for searching and predicting drug-like compounds. *Bioinformatics*, 24(13):i366–i374, 2008.

[50] R. T. Cardoso, R. H. Takahashi, and F. R. Cruz. Pareto optimal solutions for stochastic dynamic programming problems via Monte Carlo simulation. *Journal of Applied Mathematics*, 2013, 2013.

[51] G. Caron, P. Saveant, and M. Schoenauer. Multiobjective tactical planning under uncertainty for air traffic flow and capacity management. In *Evolutionary Computation (CEC), 2013 IEEE Congress on*, pages 1548–1555, 2013.

[52] T. Chai, Y. Jin, and B. Sendhoff. Evolutionary complex engineering optimization: Opportunities and challenges [guest editorial]. *Computational Intelligence Magazine, IEEE*, 8(3):12–15, 2013.

[53] T.-J. Chang, N. Meade, J. E. Beasley, and Y. M. Sharaiha. Heuristics for cardinality constrained portfolio optimisation. *Computers & Operations Research*, 27(13):1271 – 1302, 2000.

[54] P. Cheeseman. An inquiry into computer understanding. *Computational Intelligence*, 4(2):58–66, 1988.

[55] J. Chen, B. Xin, Z. Peng, L. Dou, and J. Zhang. Optimal contraction theorem for exploration–exploitation tradeoff in search and optimization. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 39(3):680–691, May 2009.

[56] R. F. Coelho. Co-evolutionary optimization for multi-objective design under uncertainty. *Journal of Mechanical Design*, 135(2), February 2013. Article Number: 021006.

[57] R. F. Coelho and P. Bouillard. Multi-objective reliability-based optimization with stochastic metamodels. *Evol. Comput.*, 19(4):525–560, Dec. 2011.

[58] M. Corazza and D. Favaretto. On the existence of solutions to the quadratic mixed-integer mean–variance portfolio selection problem. *European Journal of Operational Research*, 176(3):1947–1960, 2007.

[59] T. M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.

[60] Y. Crama and M. Schyns. Simulated annealing for complex portfolio selection problems. *European Journal of Operational Research*, 150:546–571, 2003.

[61] S. Dasgupta, C. Papadimitriou, and U. Vazirani. *Algorithms*. McGraw-Hill Education, 2006.

[62] T. Dasu, S. Krishnan, S. Venkatasubramanian, and K. Yi. An information-theoretic approach to detecting changes in multi-dimensional data streams. Technical report, Duke University, 2005.

[63] J. Dattorro. *Convex Optimization & Euclidean Distance Geometry*. Meboo Publishing, 2005.

[64] R. Dawkins. The evolved imagination: Animals as models of their world. *Natural History magazine*, 104:8–11, 1995.

[65] L. N. De Castro and F. J. Von Zuben. *Recent developments in biologically inspired computing*. Igi Global, 2005.

[66] K. Deb. Introduction to evolutionary multiobjective optimization. In J. Branke, K. Deb, K. Miettinen, and R. Slowinski, editors, *Multiobjective Optimization*, volume 5252 of *Lecture Notes in Computer Science*, pages 59–96. Springer Berlin / Heidelberg, 2008.

[67] K. Deb and R. B. Agrawal. Simulated binary crossover for continuous search space. *Complex Systems*, 9:115–148, 1995.

[68] K. Deb, N. U. Bhaskara Rao, and S. Karthik. Dynamic multi-objective optimization and decision-making using modified NSGA-II: A case study on hydro-thermal power scheduling. In *Proceedings of 4th international conference on Evolutionary multi-criterion optimization*, EMO'07, pages 803–817, 2007.

[69] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evolutionary Computation, IEEE Transactions on*, 6(2):182–197, Apr 2002.

[70] K. Deb, R. E. Steuer, R. Tewari, and R. Tewari. Bi-objective portfolio optimization using a customized hybrid NSGA-II procedure. *Lecture Notes in Computer Science*, 6576:358–373, 2011.

[71] V. DeMiguel and F. J. Nogales. Portfolio selection with robust estimation. *Operations Research*, 57(3):560–577, 2009.

[72] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.

[73] K. Devika, A. Jafarian, and V. Nourbakhsh. Designing a sustainable closed-loop supply chain network based on triple bottom line approach: A comparison of metaheuristics hybridization techniques. *European Journal of Operational Research*, 235(3):594 – 615, 2014.

[74] P. Di Barba. Dynamic multiobjective optimization: A way to the shape design with transient magnetic fields. *IEEE Transactions on Magnetics*, 44(6):962–965, june 2008.

[75] A. Doucet, N. de Freitas, and N. Gordon. An introduction to sequential Monte Carlo methods. In *Sequential Monte Carlo Methods in Practice*, Statistics for Engineering and information science, chapter 1, pages 1–13. Springer, Berlin, 2001.

[76] D. Dubois and H. Prade. Possibilistic abduction. In B. Bouchon-Meunier, L. Valverde, and R. R. Yager, editors, *IPMU '92 – Advanced Methods in Artificial Intelligence*, volume 682 of *Lecture Notes in Computer Science*, pages 1–12. Springer Berlin Heidelberg, 1993.

[77] J. Eblen, C. Phillips, G. Rogers, and M. Langston. The maximum clique enumeration problem: algorithms, applications, and implementations. *BMC Bioinformatics*, 13(Suppl 10):S5, 2012.

[78] C. Elkan. The paradoxical controversy over fuzzy logic. *IEEE Expert: Intelligent Systems and Their Applications*, 9(4):47–49, Aug. 1994.

[79] M. Emmerich and A. Deutz. Time complexity and zeros of the hypervolume indicator gradient field. In O. Schuetze, C. A. Coello Coello, A.-A. Tantar, E. Tantar, P. Bouvry, P. D. Moral, and P. Legrand, editors, *EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation III*, volume 500 of *Studies in Computational Intelligence*, pages 169–193. Springer International Publishing, 2014.

[80] R. Engle. *Anticipating Correlations: A New Paradigm for Risk Management*. The Econometric and Tinbergen Institutes Lectures. Princeton University Press, 2009.

[81] L. Euler. Solutio facilis problematum quorundam geometricorum difficillimorum. *Novi Comm. Acad. Scie. Petropolitanae*, 11:139–157, 1765.

[82] F. J. Fabozzi, D. Huang, and G. Zhou. Robust portfolios: contributions from operations research and finance. *Annals of Operations Research*, 176(1):191–220, 2010.

[83] F. J. Fabozzi, P. N. Kolm, D. Pachamanova, and S. M. Focardi. *Robust Portfolio Optimization and Management*. Wiley, New Jersey, 2007.

[84] M. Farina, K. Deb, and P. Amato. Dynamic multiobjective optimization problems: test cases, approximations, and applications. *Evolutionary Computation, IEEE Transactions on*, 8(5):425–442, 2004.

[85] P. A. V. Ferreira. Modelagem, análise e projeto de sistemas com múltiplos objetivos. Tese de livre-docência, UNICAMP, 1998.

[86] M. Fleischer. The measure of Pareto optima: applications to multi-objective metaheuristics. In *Evolutionary Multi-Criterion Optimization. Second International Conference, EMO 2003*, pages 519–533, 2003.

[87] J. Fox. *The Myth of the Rational Market*. HarperCollins, 2009.

[88] G. Frahm and U. Jaekel. Tyler's m-estimator, random matrix theory, and generalized elliptical distributions with applications to finance. Discussion Papers in Statistics and Econometrics 2/07, University of Cologne, Department for Economic and Social Statistics, 2007.

[89] F. D. Freitas, A. F. D. Souza, and A. R. de Almeida. Prediction-based portfolio optimization model using neural networks. *Neurocomputing*, 72(10–12):2155 – 2170, 2009.

[90] H. Fu, B. Sendhoff, K. Tang, and X. Yao. Finding robust solutions to dynamic optimization problems. In *Applications of Evolutionary Computation*, volume 7835 of *Lecture Notes in Computer Science*, pages 616–625. Springer Berlin Heidelberg, 2013.

[91] N. Furukawa. Characterization of optimal policies in vector-valued Markovian decision processes. *Mathematics of Operations Research*, 5(2):271–279, 1980.

[92] N. Furukawa. Recurrence set-relations in stochastic multiobjective dynamic decision processes. *Mathematische Operationsforschung und Statistik. Series Optimization*, 13(1):113–121, 1982.

[93] S. García, D. Quintana, I. M. Galván, and P. Isasi. Time-stamped resampling for robust evolutionary portfolio optimization. *Expert Systems with Applications*, 39:10722–10730, 2012.

[94] A. Gasparatos, M. El-Haram, and M. Horner. A critical review of reductionist approaches for assessing the progress towards sustainability. *Environmental Impact Assessment Review*, 28(4–5):286–311, 2008.

[95] R. E. Giachetti, P. Damodaran, S. Mestry, and C. Prada. Optimization-based decision support system for crew scheduling in the cruise industry. *Computers & Industrial Engineering*, 64(1):500 – 510, 2013.

[96] C.-K. Goh and K. C. Tan. A competitive-cooperative coevolutionary paradigm for dynamic multiobjective optimization. *IEEE Transactions on Evolutionary Computation*, 13(1):103–127, 2009.

[97] D. E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison-Wesley, 1989.

[98] D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics Of Operations Research*, 28(1):1–38, 2003.

[99] A. Greenspan. Testimony to house of representatives committee on government oversight and reform. https://house.resource.org/110/gov.house.ogr.20081023_hrs15REF2154.pdf. Accessed: 2014-07-08.

[100] M. S. Grewal and A. P. Andrews. *Kalman Filtering: theory and practice using MATLAB.* Wiley, New Jersey, 2008.

[101] V. Griffith and C. Koch. Quantifying synergistic mutual information. In M. Prokopenko, editor, *Guided Self-Organization: Inception*, volume 9 of *Emergence, Complexity and Computation*, pages 159–190. Springer Berlin Heidelberg, 2014.

[102] W. Gutjahr and A. Pichler. Stochastic multi-objective optimization: a survey on non-scalarizing methods. *Annals of Operations Research*, pages 1–25, 2013.

[103] L. Gyorfi, F. Udina, and H. Walk. Experiments on Universal Portfolio Selection Using Data From Real Markets. Upf working paper, Universitat Pompeu Fabra, Department of Economics and Business, 2007.

[104] M. R. Haley, H. J. Paarsch, and C. H. Whiteman. Smoothed safety first and the holding of assets. *Quantitative Finance*, 13(2):167–176, 2013.

[105] W. D. Hamilton and S. P. Brown. Autumn tree colours as a handicap signal. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1475):1489–1493, 2001.

[106] E.-H. Han and G. Karypis. Centroid-based document classification: Analysis and experimental results. In *Proceedings of the 4th European Conference on Principles of Data Mining and Knowledge Discovery*, PKDD '00, pages 424–431, London, UK, UK, 2000. Springer-Verlag.

[107] R. Hanel, S. Thurner, and M. Gell-Mann. How multiplicity determines entropy and the derivation of the maximum entropy principle for complex systems. *Proceedings of the National Academy of Sciences*, 111(19):6905–6910, 2014.

[108] A. Hart. *Anticipations, Uncertainty, and Dynamic Planning.* Reprints of economic classics. A.M. Kelley, 1940.

[109] G. Hassan and C. D. Clack. Multiobjective robustness for portfolio optimization in volatile environments. In *Proceedings of the 10th annual conference on Genetic and evolutionary computation*, pages 1507–1514. ACM, 2008.

[110] I. Hatzakis and D. Wallace. Dynamic multi-objective optimization with evolutionary algorithms: a forward-looking approach. In *Proceedings of the 8th annual conference on Genetic and evolutionary computation*, pages 1201–1208. ACM, 2006.

[111] W. Heisenberg. über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. *Zeitschrift für Physik*, 43(3-4):172–198, 1927.

[112] M. Helbig and A. P. Engelbrecht. Archive management for dynamic multi-objective optimisation problems using vector evaluated particle swarm optimisation. In A. E. Smith, editor, *Proceedings of the 2011 IEEE Congress on Evolutionary Computation*, pages 2040–2047, New Orleans, USA, 5-8 June 2011. IEEE Computational Intelligence Society, IEEE Press.

[113] M. Helbig and A. P. Engelbrecht. Performance measures for dynamic multi-objective optimisation algorithms. *Information Sciences*, 250(0):61 – 81, 2013.

[114] M. Helbig and A. P. Engelbrecht. Benchmarks for dynamic multi-objective optimisation algorithms. *ACM Comput. Surv.*, 46(3):37:1–37:39, Jan. 2014.

[115] M. Helbig and A. P. Engelbrecht. Population-based metaheuristics for continuous boundary-constrained dynamic multi-objective optimisation problems. *Swarm and Evolutionary Computation*, 14(0):31 – 47, 2014.

[116] M. Henig. Vector-valued dynamic programming. *SIAM Journal on Control and Optimization*, 21(3):490–499, 1983.

[117] R. B. Holmes. On random correlation matrices. *SIAM J. Matrix Anal. Appl.*, 12(2):239–272, Mar. 1991.

[118] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

[119] D. W. Hubbard. *The Failure of Risk Management: Why It's Broken and How to Fix It*. Wiley, 2009.

[120] P. J. Huber and E. M. Ronchetti. *Robust Statistics*. Wiley, New Jersey, 2009.

[121] E. J. Hughes. Evolutionary multi-objective ranking with uncertainty and noise. In E. Zitzler, L. Thiele, K. Deb, C. A. Coello Coello, and D. Corne, editors, *Evolutionary Multi-Criterion Optimization*, volume 1993 of *Lecture Notes in Computer Science*, pages 329–343. Springer Berlin Heidelberg, 2001.

[122] M. J. Hutchins and J. W. Sutherland. An exploration of measures of social sustainability and their application to supply chain decisions. *Journal of Cleaner Production*, 16(15):1688 – 1698, 2008. Sustainability and Supply Chain Management.

[123] D. E. Hutchison and G. G. Pennacchi. Measuring rents and interest rate risk in imperfect financial markets: The case of retail bank deposits. *Journal of Financial and Quantitative Analysis*, 31(03):399–417, 1996.

[124] I. R. James and J. E. Mosimann. A new characterization of the Dirichlet distribution through neutrality. *The Annals of Statistics*, pages 183–189, 1980.

[125] Y. Jin, K. Tang, X. Yu, B. Sendhoff, and X. Yao. A framework for finding robust optimal solutions over time. *Memetic Computing*, 5(1):3–18, 2013.

[126] D. Jones and M. Tamiz. Detection and restoration of Pareto inefficiency. In *Practical Goal Programming*, volume 141 of *International Series in Operations Research & Management Science*, pages 95–112. Springer US, 2010.

[127] L. P. Kaelbling, M. L. Littman, and A. P. Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285, 1996.

[128] L. P. Kaelbling and T. Lozano-Pérez. Integrated task and motion planning in belief space. *The International Journal of Robotics Research*, 32(9-10):1194–1227, 2013.

[129] D. Kahneman. *Thinking, Fast and Slow*. Doubleday Canada, Limited, 2013.

[130] J. R. Kasprzyk, P. M. Reed, B. R. Kirsch, and G. W. Characklis. Managing population and drought risks using many-objective water portfolio planning under uncertainty. *Water Resources Research*, 45(12):n/a–n/a, 2009.

[131] D. Kikuti, F. G. Cozman, and R. S. Filho. Sequential decision making with partially ordered preferences. *Artificial Intelligence*, 175:1346–1365, May 2011.

[132] A. J. Kleywegt, V. S. Nori, and M. W. P. Savelsbergh. The stochastic inventory routing problem with direct deliveries. *Transportation Science*, 36(1):94–118, Feb. 2002.

[133] R. Klötzler. Multiobjective dynamic programming. *Mathematische Operationsforschung und Statistik. Series Optimization: A Journal of Mathematical Programming and Operations Research*, 9(3):423–426, 1978.

[134] P. N. Kolm, R. Tutuncu, and F. J. Fabozzi. 60 years of portfolio optimization: Practical challenges and current trends. *European Journal of Operational Research*, 234(2):356 – 371, 2014.

[135] W. Koo, C. Goh, and K. Tan. A predictive gradient strategy for multiobjective evolutionary algorithms in a fast changing environment. *Memetic Computing*, 2:87–110, 2010.

[136] T. C. Koopmans. On flexibility of future preference. Cowles Foundation Discussion Papers 150, Cowles Foundation for Research in Economics, Yale University, 1962.

[137] B. Koszegi. Utility from anticipation and personal equilibrium. *Economic Theory*, 44(3):415–444, 2010.

[138] C. Kotropoulos and V. Moschou. Self organizing maps for reducing the number of clusters by one on simplex subspaces. In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 5, pages V–V, 2006.

[139] M. Kremer, S. Minner, and L. N. Van Wassenhove. On the preference to avoid ex post inventory errors. *Production and Operations Management*, 23(5):773–787, 2014.

[140] D. M. Kreps. A representation theorem for "preference for flexibility". *Econometrica*, 47(3):pp. 565–577, 1979.

[141] H. W. Kuhn and A. W. Tucker. Nonlinear programming. In *Second Berkeley Symposium on Mathematical Statistics and Probability*, 1951.

[142] V. Kumar. Entropic measures of manufacturing flexibility. *International Journal of Production Research*, 25(7):957–966, 1987.

[143] L. W. Labaree and B. Franklin. *Mr. Franklin: a selection from his personal letters*. Yale University Press, 1956.

[144] W.-P. Lee. Towards agent-based decision making in the electronic marketplace: interactive recommendation and automated negotiation. *Expert Systems with Applications*, 27(4):665 – 679, 2004.

[145] J. Lehman and K. O. Stanley. Evolvability is inevitable: Increasing evolvability without the pressure to adapt. *PLoS ONE*, 8(4):e62186, 04 2013.

[146] J. Liesiö and A. Salo. Scenario-based portfolio selection of investment projects with incomplete probability and utility information. *European Journal of Operational Research*, 217(1):162 – 172, 2012.

[147] K. Lwin, R. Qu, and G. Kendall. A learning-guided multi-objective evolutionary algorithm for constrained portfolio optimization. *Applied Soft Computing*, 2014.

[148] D. Ly and H. Lipson. Optimal experiment design for coevolutionary active learning. *Evolutionary Computation, IEEE Transactions on*, PP(99):1–1, 2013.

[149] M. Mandelbaum and J. Buzacott. Flexibility and decision making. *European Journal of Operational Research*, 44(1):17 – 27, 1990.

[150] R. Mansini and M. G. Speranza. Heuristic algorithms for the portfolio selection problem with minimum transaction lots. *European Journal of Operational Research*, 114:219–233, 1999.

[151] H. Markowitz. *Portfolio Selection: Efficient Diversification of Investments*. University of Chicago., 1955.

[152] H. M. Markowitz. *Portfolio Selection: Efficient Diversification of Investments*. Wiley, New Jersey, 1991.

[153] R. A. Maronna and R. H. Zamar. Robust estimates of location and dispersion of high-dimensional datasets. *Technometrics*, 44(4):307–317, 2002.

[154] R. K. Merton. *Social Theory and Social Structure*. New York: Free Press, 1968.

[155] T. Mitsa. *Temporal Data Mining*. Chapman & Hall/CRC Data Mining and Knowledge Discovery Series. Taylor & Francis, 2010.

[156] M. Mørup and L. K. Hansen. An exact relaxation of clustering. Technical report, Department of Iinformatics and Mathematical Modeling, Technical University of Denmark, 2009.

[157] T. S. Motzkin and E. G. Straus. Maxima for graphs and a new proof of a theorem of turán. *Canad. J. Math*, 17(4):533–540, 1965.

[158] A. Muller. Stochastic ordering of multivariate normal distributions. *Annals of the Institute of Statistical Mathematics*, 53:567–575, 2001.

[159] A. Muruganantham, Y. Zhao, S. B. Gee, X. Qiu, and K. C. Tan. Dynamic multiobjective optimization using evolutionary algorithm with Kalman filter. *Procedia Computer Science*, 24(0):66 – 75, 2013.

[160] K. W. Ng, G. L. Tian, and M. L. Tang. *Dirichlet and Related Distributions: Theory, Methods and Applications*. Wiley Series in Probability and Statistics. Wiley, 2011.

[161] V. Nollau. A stochastic decision model with vector-valued reward. *Optimization*, 16(5):733–742, 1985.

[162] B. Nyhan and J. Reifler. Blank slates or closed minds? *American Journal of Political Science*, 2014 (to appear).

[163] S. Y. Park and A. K. Bera. Maximum entropy autoregressive conditional heteroskedasticity model. *Journal of Econometrics*, 150(2):219–230, 2009.

[164] G. Parmigiani. Decision theory: Bayesian. In N. J. Smelser and P. B. Baltes, editors, *International Encyclopedia of the Social & Behavioral Sciences*, pages 3327 – 3334. Pergamon, Oxford, 2001.

[165] S. Parsons. *Qualitative Methods for Reasoning Under Uncertainty*. Number v. 13 in Artificial Intelligence Series. Morgan Kaufmann Publishers, 2001.

[166] J. Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann series in representation and reasoning. Morgan Kaufmann Publishers, 1988.

[167] V. Picheny. Multiobjective optimization using Gaussian process emulators via stepwise uncertainty reduction. *Statistics and Computing*, pages 1–16, 2013.

[168] W. B. Powell. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Wiley Series in Probability and Statistics. Wiley, 2007.

[169] C. Puppe. An Axiomatic Approach to "Preference for Freedom of Choice". *Journal of Economic Theory*, 68(1):174–199, January 1996.

[170] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 1st edition, 2006.

[171] G. Rizzolatti, L. Fadiga, V. Gallese, and L. Fogassi. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3(2):131 – 141, 1996.

[172] C. Robert. *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. Springer Texts in Statistics. U.S. Government Printing Office, 2001.

[173] D. Roijers, P. Vamplew, S. Whiteson, and R. Dazeley. A survey of multi-objective sequential decision-making. *Journal of Artificial Intelligence Research*, 48:67–113, 2013.

[174] R. Rosen. *Anticipatory systems: philosophical, mathematical, and methodological foundations*. Pergamon Pr, 1st edition, 1985.

[175] C. Rossi, M. Abderrahim, and J. C. Díaz. Tracking moving optima using Kalman-based predictions. *Evol. Comput.*, 16(1):1–30, Mar. 2008.

[176] S. Rota Bulò and M. Pelillo. A game-theoretic approach to hypergraph clustering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(6):1312–1327, 2013.

[177] A. D. Roy. Safety first and the holding of assets. *Econometrica*, 20(3):431–449, 1952.

[178] R. Ruiz-Torrubiano and A. Suarez. Hybrid approaches and dimensionality reduction for portfolio selection with cardinality constraints. *Computational Intelligence Magazine, IEEE*, 5(2):92–107, 2010.

[179] S. Salomon, G. Avigad, P. J. Fleming, and R. C. Purshouse. Optimization of adaptation - a multi-objective approach for optimizing changes to design parameters. In R. Purshouse, P. Fleming, C. Fonseca, S. Greco, and J. Shaw, editors, *Evolutionary Multi-Criterion Optimization*, volume 7811 of *Lecture Notes in Computer Science*, pages 21–35. Springer Berlin Heidelberg, 2013.

[180] P. A. Samuelson. Proof that Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review*, 6:41–49, 1965.

[181] P. A. Samuelson. Proof that properly discounted present values of assets vibrate randomly. *Bell Journal of Economics*, 4(2):369–374, 1973.

[182] J. Schmidhuber. Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *Autonomous Mental Development, IEEE Transactions on*, 2(3):230–247, Sept 2010.

[183] J. Schönberger and H. Kopfer. Online decision making and automatic decision model adaptation. *Computers & Operations Research*, 36(6):1740 – 1750, 2009.

[184] M. T. Schultz, K. N. Mitchell, and B. K. Harper. Decision Making Under Uncertainty. Final technical report, Engineer Research and Development Center, US Army Corps of Engineers, 2010.

[185] H.-P. Schwefel. *Evolutionsstrategie und numerische Optimierung*. PhD thesis, Technische Universität Berlin, 1975.

[186] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[187] S. Shalev-Shwartz. Online learning and online convex optimization. *Found. Trends Mach. Learn.*, 4(2):107–194, Feb. 2012.

[188] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[189] W. F. Sharpe. The Sharpe ratio. *The Journal of Portfolio Management*, 21(1):49 – 58, 1994.

[190] A. Simões and E. Costa. Prediction in evolutionary algorithms for dynamic environments. *Soft Computing*, pages 1–27, 2013.

[191] S. Sivalinghem, M. M. Kasumovic, A. C. Mason, M. C. B. Andrade, and D. O. Elias. Vibratory communication in the jumping spider phidippus clarus: polyandry, male courtship signals, and mating success. *Behavioral Ecology*, 21(6):1308–1314, 2010.

[192] A. M. J. Skulimowski. Solving vector optimization problems via multilevel analysis of foreseen consequences. *Foundations of Control Engineering*, 10:25–38, 1985.

[193] A. M. J. Skulimowski. Freedom of choice and creativity in multicriteria decision making. In *Knowledge, Information, and Creativity Support Systems*, pages 190–203. Springer, 2011.

[194] A. M. J. Skulimowski. Hybrid anticipatory networks. In L. Rutkowski, M. Korytkowski, R. Scherer, R. Tadeusiewicz, L. Zadeh, and J. Zurada, editors, *Artificial Intelligence and Soft Computing*, volume 7268 of *Lecture Notes in Computer Science*, pages 706–715. Springer Berlin Heidelberg, 2012.

[195] M. Small. *Applied Nonlinear Time Series Analysis: Applications in Physics, Physiology and Finance*. World Scientific series in nonlinear science, Series A. World Scientific, 2005.

[196] M. Soysal, J. Bloemhof-Ruwaard, and J. van der Vorst. Modelling food logistics networks with emission considerations: The case of an international beef supply chain. *International Journal of Production Economics*, 152(0):57 – 70, 2014. Sustainable Food Supply Chain Management.

[197] M. Spence. Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374, 1973.

[198] P. Stinis. Stochastic global optimization as a filtering problem. *J. Comput. Phys.*, 231(4):2002–2014, Feb. 2012.

[199] P. D. Stroud. Kalman-extended genetic algorithm for search in nonstationary environments with noisy fitness evaluations. *Evolutionary Computation, IEEE Transactions on*, 5(1):66–77, 2001.

[200] G. Sugihara, R. May, H. Ye, C.-h. Hsieh, E. Deyle, M. Fogarty, and S. Munch. Detecting causality in complex ecosystems. *Science*, 338(6106):496–500, 2012.

[201] E. J. Sullivan. A.D. Roy: the forgotten father of portfolio theory. *Research in the History of Economic Thought and Methodology*, 29:73–82, 2011.

[202] F. Takens. Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980*, pages 366–381. Springer, 1981.

[203] Q. Tang, Y. Peng, X. Zhang, and C. Zhao. On graph-lagrangians of hypergraphs containing dense subgraphs. *Journal of Optimization Theory and Applications*, pages 1–26, 2013.

[204] J. R. Taylor. *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. A series of books in physics. University Science Books, 1997.

[205] J. Teich. Pareto-front exploration with uncertain objectives. In E. Zitzler, L. Thiele, K. Deb, C. A. Coello Coello, and D. Corne, editors, *Evolutionary Multi-Criterion Optimization*, volume 1993 of *Lecture Notes in Computer Science*, pages 314–328. Springer Berlin Heidelberg, 2001.

[206] H. Thomas and J. L. Stoddart. Leaf senescence. *Annual Review of Plant Physiology*, 31(1):83–111, 1980.

[207] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. Intelligent robotics and autonomous agents series. Mit Press, 2005.

[208] S. Thrun, M. Montemerlo, and A. Aron. Probabilistic terrain analysis for high-speed desert driving. In *Robotics: Science and Systems*, pages 16–19, 2006.

[209] R. M. Unger. *The self awakened: Pragmatism unbound*. Harvard University Press, 2007.

[210] J. I. van Hemert, C. Van Hoyweghen, E. Lukschandl, and K. Verbeeck. A "futurist" approach to dynamic environments. In *Proceedings of the Workshop on Evolutionary Algorithms for Dynamic Optimization Problems at the Genetic and Evolutionary Computation Conference*, pages 35–38. ACM, 2001.

[211] P. Van Hentenryck and R. Bent. *Online Stochastic Combinatorial Optimization*. MIT Press, 2006.

[212] N. Van Hoeck, N. Ma, L. Ampe, K. Baetens, M. Vandekerckhove, and F. Van Overwalle. Counterfactual thinking: an fmri study on changing the past for a better future. *Social Cognitive and Affective Neuroscience*, 2012.

[213] K. Van Moffaert, M. M. Drugan, and A. Nowé. Hypervolume-based multi-objective reinforcement learning. In R. Purshouse, P. Fleming, C. Fonseca, S. Greco, and J. Shaw, editors, *Evolutionary Multi-Criterion Optimization*, volume 7811 of *Lecture Notes in Computer Science*, pages 352–366. Springer Berlin Heidelberg, 2013.

[214] W. N. Venables and B. D. Ripley. *Modern Applied Statistics with S*. Springer, New York, 2002.

[215] B. Viswanathan, V. V. Aggarwal, and K. P. K. Nair. Multiple criteria Markov decision processes. *TIMS Studies in the management sciences*, 6:263–272, 1977.

[216] J. Von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton university press, 1944.

[217] M. Voorneveld. Characterization of Pareto dominance. *Operations Research Letters*, 31(1):7 – 11, 2003.

[218] W. Wang and M. Sebag. Hypervolume indicator and dominance reward based multi-objective Monte-Carlo tree search. *Machine Learning*, 92(2-3):403–429, 2013.

[219] D. J. White. Multi-objective infinite-horizon discounted Markov decision processes. *Journal of Mathematical Analysis and Applications*, 89(2), 1982.

[220] W. Wiesemann, D. Kuhn, and B. Rustem. Multi-resource allocation in stochastic project scheduling. *Annals of Operations Research*, 193(1):193–220, 2012.

[221] A. D. Wissner-Gross and C. E. Freer. Causal entropic forces. *Phys. Rev. Lett.*, 110:168702, Apr 2013.

[222] D. Wolpert, Z. Ghahramani, and M. Jordan. An internal model for sensorimotor integration. *Science*, 269(5232):1880–1882, 1995.

[223] S. Yang and R. Tinos. Hyper-selection in dynamic environments. In *IEEE World Congress on Computational Intelligence*, pages 3185–3192, 2008.

[224] L. A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1(1):3 – 28, 1978.

[225] Z. Zhang. Multiobjective optimization immune algorithm in dynamic environments and its application to greenhouse control. *Applied Soft Computing*, 8(2):959 – 971, 2008.

[226] A. Zhou, Y. Jin, and Q. Zhang. A population prediction strategy for evolutionary dynamic multiobjective optimization. *Cybernetics, IEEE Transactions on*, 44(1):40–53, Jan 2014.

[227] E. Zitzler, J. Knowles, and L. Thiele. Quality assessment of Pareto set approximations. In J. Branke, K. Deb, K. Miettinen, and R. Slowinski, editors, *Multiobjective Optimization*, volume 5252 of *Lecture Notes in Computer Science*, pages 373–404. Springer Berlin Heidelberg, 2008.

[228] E. Zitzler, M. Laumanns, L. Thiele, C. M. Fonseca, and V. G. d. Fonseca. Why quality assessment of multiobjective optimizers is difficult. In *Proceedings of the Genetic and Evolutionary Computation Conference*, GECCO '02, pages 666–674, San Francisco, CA, USA, 2002. Morgan Kaufmann Publishers Inc.

[229] E. Zitzler, L. Thiele, and J. Bader. On set-based multiobjective optimization. *Evolutionary Computation, IEEE Transactions on*, 14(1):58–79, 2010.

Appendices

Benjamin Franklin's Multiple Criteria Decision-Making Method
To Joseph Priestley
London, September 19, 1772

Dear Sir,

In the Affair of so much Importance to you, wherein you ask my Advice, I cannot for want of sufficient Premises, advise you what to determine, but if you please I will tell you how.

When these difficult Cases occur, they are difficult chiefly because while we have them under Consideration all the Reasons pro and con are not present to the Mind at the same time; but sometimes one Set present themselves, and at other times another, the first being out of Sight. Hence the various Purposes or Inclinations that alternately prevail, and the Uncertainty that perplexes us.

To get over this, my Way is, to divide half a Sheet of Paper by a Line into two Columns, writing over the one Pro, and over the other Con. Then during three or four Days Consideration I put down under the different Heads short Hints of the different Motives that at different Times occur to me for or against the Measure. When I have thus got them all together in one View, I endeavour to estimate their respective Weights; and where I find two, one on each side, that seem equal, I strike them both out: If I find a Reason pro equal to some two Reasons con, I strike out the three. If I judge some two Reasons con equal to some three Reasons pro, I strike out the five; and thus proceeding I find at length where the Ballance lies; and if after a Day or two of farther Consideration nothing new that is of Importance occurs on either side, I come to a Determination accordingly.

And tho' the Weight of Reasons cannot be taken with the Precision of Algebraic Quantities, yet when each is thus considered separately and comparatively, and the whole lies before me, I think I can judge better, and am less likely to take a rash Step; and in fact I have found great Advantage from this kind of Equation, in what may be called Moral or Prudential Algebra.

Wishing sincerely that you may determine for the best, I am ever, my dear Friend,

Yours most affectionately,

B. Franklin [143]