EPFL

Thèse n° 10 298

Scalable Stochastic Optimization: Scenario Reduction with Guarantees

Présentée le 18 septembre 2020

au Collège du management de la technologie
Chaire analyse de risque et optimisation
Programme doctoral en management de la technologie

pour l'obtention du grade de Docteur ès Sciences

par

Kilian SCHINDLER

Acceptée sur proposition du jury

Prof. D. Foray, président du jury
Prof. D. Kuhn, directeur de thèse
Prof. W. Wiesemann, rapporteur
Prof. M. Sim, rapporteur
Prof. K. Younge, rapporteur

École polytechnique fédérale de Lausanne

2020

Plans are nothing, planning is everything.
— Dwight D. Eisenhower

To my family, including Rachel. . .

Acknowledgements

I would like to take the opportunity to express my gratitude to Prof. Daniel Kuhn, Prof. Wolfram Wiesemann, and Prof. Napat Rujeerapaiboon for all the extensive joint work at the base of the various papers written together.

I have benefited tremendously from their extensive knowledge in the field of distributionally robust methods for dynamic data-driven optimization under uncertainty, their highly effective and professional working ethics, and their unrelenting desire to stretch knowledge to new frontiers. It is through our work together that I was rapidly "lifted" into the world of producing and publishing academic papers with a world-class quality. This has been a true honor for me, and hereby would like to express to all three of them my recognition and everlasting appreciation. I do hope that in some ways they have also enjoyed working with a less experienced and less gifted person as myself.

Over the last few years, I have also benefited from the discussions, remarks, questions, and analytical insights from many colleagues at the Risk Analytics and Optimization team. I would like to wish each and every one of them all possible success with their current and future work.

My gratitude also goes to my family for their permanent support, and for happily embracing the notion that life at EPFL and Pully became "my home".

Finally, all my thankfulness and endless love to Rachel Lacroix - my much beloved life companion. Her endless caring and support makes both my life and my work at EPFL all the more meaningful. It is pure joy to share our extensive conversations and discussions about our respective PhD-related work.

Lausanne, May 16, 2020 Kilian Schindler

Abstract

Stochastic optimization is a popular modeling paradigm for decision-making under uncertainty and has a wide spectrum of applications in management science, economics and engineering. However, the stochastic optimization models one faces in practice are intractable, and numerical solutions necessitate approximations. The mainstream approach for making a stochastic optimization model amenable to numerical solution is to discretize the probability distribution of the uncertain problem parameters. However, both the accuracy of the approximation as well as the computational burden of solving the approximate problem scale with the number of scenarios of the approximate distribution. An effective means to ease the computational burden is to use *scenario reduction*, which replaces an accurate initial distribution accommodating many scenarios with a simpler distribution supported on only few scenarios that is close to the initial distribution with respect to a probability metric.

Using the Wasserstein distance as measure of proximity between distributions, we provide new insights into the fundamental limitations of scenario reduction, and we propose the first polynomial-time constant-factor approximations for a popular scenario reduction problem from the literature. As scenario reduction is equivalent to clustering, it suffers from two well-known shortcomings. Namely, it suffers from outlier sensitivity and may produce highly unbalanced clusters. To mitigate both shortcomings, we formulate a joint outlier detection and clustering problem, where the clusters must satisfy certain cardinality constraints. We cast this problem as a mixed-integer linear program (MILP) that admits tractable semidefinite and linear programming relaxations. We propose deterministic rounding schemes that transform the relaxed solutions to feasible solutions for the MILP. We also prove that these solutions are optimal in the MILP if a cluster separation condition holds. Finally, we develop a highly efficient scenario reduction method for a large-scale hydro scheduling problem. Specifically, we study the optimal operation of a fleet of interconnected hydropower plants that sell energy on both the spot and the reserve markets, and we propose a two-layer stochastic programming framework for its solution. The outer layer problem (the *planner's problem*) optimizes the end-of-day reservoir filling levels over one year, whereas the inner layer problem (the *trader's problem*) selects optimal hourly market bids within each day. We prove that the trader's problem admits a scenario reduction that dramatically reduces its complexity without loss of optimality, which in turn facilitates an efficient approximation of the planner's problem.

Key words: Stochastic Optimization; Robust Optimization; Scenario Reduction; Clustering; Hydro Power; Reserve Market

Zusammenfassung

Die stochastische Optimierung umfasst beliebte Methoden zur Modellierung von Entscheidungsproblemen unter Unsicherheit und hat eine Vielzahl möglicher Anwendungen sowohl in der Betriebs- und der Volkswirtschaftslehre als auch in den Ingenieurwissenschaften. In der Praxis haben stochastische Optimierungsmodelle aber meist keine analytische Lösung, und numerische Lösungsverfahren beruhen deshalb fast immer auf Approximationen. Das am weitesten verbreitete Approximationsverfahren beruht auf einer Diskretisierung der Wahrscheinlichkeitsverteilung der unsicheren Problemparameter. Allerdings nehmen sowohl die Genauigkeit als auch die Komplexität des resutierenden diskretisierten Optimierungsmodells mit der Anzahl der Szenarien in der diskreten Verteilung zu. Eine effektive Methode um den Rechenaufwand zu verringern ist die *Szenarioreduktion*, welche eine anfängliche genaue Verteilung mit vielen Szenarien durch eine nahe Verteilung mit nur wenigen Szenarien ersetzt, wobei man die Nähe zwischen Verteilungen mittels einer Wahrscheinlichkeitsmetrik misst.

In dieser Dissertation zeigen wir zuerst die fundamentalen Grenzen der Szenarioreduktion auf, wenn man als Wahrscheinlichkeitsmetrik die Wasserstein-Distanz verwendet. Wir beschreiben zudem das erste polynomielle Approximationsverfahren mit einer garantierten Güte für ein weitverbreitetes Szenarioreduktionsproblem aus der Literatur. Aufgrund seiner Äquivalenz zum Clusteringproblem weist die Szenarioreduktion zwei wohlbekannte Defizite auf: Sie ist sensibel ggü. Ausreissern, und sie neigt zur Bildung höchst unausgeglichener Cluster. Wir adressieren beide Nachteile durch die Formulierung eines simultanen Ausreissererfassungs- und Clusteringproblems, in welchem die Cluster bestimmte Kardinalitätsvorgaben erfüllen müssen. Wir formulieren dieses Problem als ein gemischt-ganzzahliges Optimierungsproblem, welches effizient lösbare semidefinite und lineare Relaxierungen aufweist. Wir entwickeln deterministische Rundungsverfahren, welche die Lösungen der Relaxierungen in zulässige Lösungen des ursprünglichen Optimierungsproblems überführt. Weiterhin beweisen wir, dass diese Lösungen im ursprünglichen gemischt-ganzzahligen Optimierungsproblem optimal sind, wenn eine bestimmte Cluster-Separationsbedingung erfüllt ist. Schliesslich entwickeln wir eine höchst effiziente Szenarioreduktionsmethode für die Bewirtschaftung grosser Wasserkraftwerke. In diesem Kontext analysieren wir die optimale Bewirtschaftung einer Gruppe von Wasserkraftwerken die Energie sowohl auf dem Spot- als auch auf dem Reservemarkt anbieten, und wir entwickeln ein zweischichtiges stochastiches Optimierungsverfahren zur Lösung dieses Problems. Die äussere Schicht unseres Verfahrens (das *Planungsproblem*) optimiert

die täglichen Reservoir-Füllmengen über den Zeitraum eines Jahres, wohingegen die innere Schicht unseres Verfahrens (das *Handelsproblem*) die optimalen stündlichen Marktgebote innerhalb eines Tages auswählt. Wir zeigen, dass das Handelsproblem einer Szenarioreduktion zugänglich ist, welche eine verlustfreie Komplexitätsreduktion ermöglicht, und dass das resultierende reduzierte Handelsproblem wiederum eine effiziente Approximation des Planungsproblems erlaubt.

Stichwörter: Stochastische Optimierung; Robuste Optimierung; Szenarioreduktion; Clustering; Wasserkraft; Reservemärkte

Contents

Acknowledgements i

Abstract (English/Deutsch) iii

List of figures ix

List of tables xi

1 Introduction 1
- 1.1 Scenario Reduction . . . 1
- 1.2 Clustering . . . 3
- 1.3 Contributions and Structure of the Thesis . . . 5
- 1.4 Statement of Originality . . . 6

2 Scenario Reduction Revisited: Fundamental Limits and Guarantees 7
- 2.1 Introduction . . . 7
- 2.2 Fundamental Limits of Scenario Reduction . . . 11
  - 2.2.1 Fundamental Limits for the Type-2 Wasserstein Distance . . . 13
  - 2.2.2 Fundamental Limits for the Type-1 Wasserstein Distance . . . 15
  - 2.2.3 Discussion . . . 17
- 2.3 Guarantees for Discrete Scenario Reduction . . . 20
  - 2.3.1 Guarantees for the Type-2 Wasserstein Distance . . . 21
  - 2.3.2 Guarantees for the Type-1 Wasserstein Distance . . . 24
- 2.4 Solution Methods for Scenario Reduction Problems . . . 26
  - 2.4.1 Heuristics for the Discrete Scenario Reduction Problem . . . 27
  - 2.4.2 Constant-Factor Approximation for the Scenario Reduction Problem . . . 32
  - 2.4.3 Mixed-Integer Reformulations of the Discrete and Continuous Scenario Reduction Problems . . . 33
- 2.5 Numerical Experiments . . . 35
  - 2.5.1 Color Quantization . . . 35
  - 2.5.2 Capital Budgeting . . . 37

3 Cardinality-Constrained Clustering and Outlier Detection via Conic Optimization **41**
3.1 Introduction . . . 41
3.2 Problem Formulation and Analysis . . . 45
3.3 Cardinality-Constrained Clustering without Outliers . . . 49
3.3.1 Convex Relaxations and Rounding Algorithm . . . 49
3.3.2 Tighter Relaxations for Balanced Clustering . . . 52
3.3.3 Comparison to existing SDP Relaxations . . . 54
3.3.4 Perfect Recovery Guarantees . . . 58
3.4 Cardinality-Constrained Clustering with Outliers . . . 61
3.5 Numerical Experiments . . . 65
3.5.1 Cardinality-Constrained K-Means Clustering (Real-World Data) . . . 66
3.5.2 Cardinality-Constrained K-Means Clustering (Synthetic Data) . . . 68
3.5.3 Outlier Detection . . . 68
3.6 Conclusion . . . 68

4 A Planner-Trader Decomposition for Multi-Market Hydro Scheduling **71**
4.1 Introduction . . . 71
4.2 Individual and Collective Bidding Models . . . 75
4.3 Planner-Trader Decomposition . . . 80
4.4 Reduction of the Individual Trading Model . . . 87
4.5 Reduction of the Collective Trading Model . . . 96
4.6 Numerical Solution of the Reduced Collective Bidding Model . . . 102
4.7 Case Study: Gasteiner Tal Cascade . . . 106
4.7.1 Problem Parameters . . . 106
4.7.2 Results . . . 108

5 Conclusion **115**
5.1 Future Research Avenues . . . 116

Bibliography **117**

Curriculum Vitae

List of Figures

2.1 Comparison between $C_2(\hat{\mathbb{P}}_n, m)$ and $\overline{C}_2(n, m)$. 20
2.2 Distributions that maximize the ratio between $D_l(\hat{\mathbb{P}}_n, 1)$ and $C_l(\hat{\mathbb{P}}_n, 1)$. 22
2.3 Outputs of different color quantization algorithms. 37
2.4 Optimality gaps of `DPCV`, `LOC-1` and `LOC-2` relative to `MILP`. 38
2.5 Approximation errors and runtimes of the reduced capital budgeting problem. 39

3.1 Sensitivity of the (un)constrained K-means clustering problem to outliers. . . . 42
3.2 Elbow plot for the dataset depicted in Figure 3.1. 66
3.3 Comparison of different algorithms for (cardinality-constrained) K-means clustering. 69
3.4 Outlier detection for breast cancer diagnosis. 70

4.1 Relations between different versions of the trader's problem. 94
4.2 Relations between different variants of the trader's problem. 99
4.3 Daily out-of-sample revenues for the spot-reserve and the spot-only model. . . 108
4.4 Cumulative out-of-sample revenues of the spot-reserve and the spot-only model. 109
4.5 Predicted vs. realized daily revenues in the spot-reserve model. 110
4.6 Daily filling levels of the Bockhartsee, Nassfeld and Remsach reservoirs. 111

List of Tables

2.1 Runtimes of different methods for discrete scenario reduction. 36

3.1 Comparison of Recovery Guarantees for K-means Clustering Relaxations. . . . 44
3.2 Overview of the main dataset characteristics. 67
3.3 Performance of $\mathscr{R}_{\mathrm{LP}}$, $\mathscr{R}_{\mathrm{SDP}}$, Bennett et al., and Peng and Wei. 67

4.1 Average daily revenues and percentage losses of various decision rule architectures.110

1 Introduction

Decision problems in management, economics or engineering almost always involve exogenous uncertain parameters. These parameters may be uncertain due to measurement errors or simply because they are not yet known at the time when the problem is formulated and solved. Indeed, when a manufacturer decides on a production plan for next year, the demands of the products and the prices of the raw materials are not yet available and can at best be estimated statistically. In this case, one aims to find decisions that perform well under all possible realizations of the uncertain parameters. If the uncertain problem parameters are governed by a known probability distribution, then the determination of an optimal policy amounts to solving a *stochastic optimization model* or *stochastic program*. While stochastic programming offers great modeling power and can therefore be used in a vast number of applications, the resulting mathematical optimization problems are hard. If the probability distribution governing the uncertain problem parameters is continuous, then we are confronted with an optimization problem over an infinite-dimensional function space. As exact analytical solutions are usually out of the question, modelers need to resort to approximations. The mainstream approach for making a stochastic optimization model amenable to numerical solution is to discretize the underlying continuous probability distribution.

The computational burden of solving the approximate stochastic program scales with the number of discretization points or *scenarios* in the approximate distribution. The fewer scenarios, the faster we can solve the approximate problem. However, large numbers of scenarios lead to better solution accuracy. The main aim of this thesis is to study the fundamental trade-off between the number of scenarios and the accuracy of the approximate distribution.

1.1 Scenario Reduction

A discrete approximate distribution that is close to the true continuous distribution can conveniently be generated by drawing a large number n of samples from the true distribution. Alternatively, it could be constructed directly from n real historical observations of the uncertain parameters. The larger n, the better the approximation of the true distribution, but

the longer it takes to solve the underlying stochastic program. An effective means to ease the computational burden is to use *scenario reduction*, an approach pioneered by Dupačová et al. (2003), which aims to approximate the initial n-point distribution with a simpler m-point distribution with $m \ll n$ that is as close as possible to the initial distribution with respect to a probability metric; see also Heitsch and Römisch (2003). The modern stability theory of stochastic programming shows that the *Wasserstein distance* is the natural choice for this probability metric. Indeed, Römisch and Schultz (1991) demonstrate that the optimal values of two-stage stochastic programs with random right-hand sides are Lipschitz continuous in the distribution of the uncertainty if the space of probability distributions is equipped with a Wasserstein metric. They further prove that the optimal solution mapping is Hölder continuous in the Wasserstein metric. Similar continuity results are established for chance constrained programs. A stronger quantitative stability theory of stochastic programming relies on the Fortet-Mourier metric, which generalizes the Wasserstein metric but offers improved convergence rates; see Rachev and Römisch (2002). For a survey of classical stability results in stochastic programming see Dupačová (1990) or Römisch and Wets (2007).

In the following we distinguish continuous scenario reduction, where the atoms of the reduced distribution may be chosen freely, and discrete scenario reduction, where the atoms of the reduced distribution must be chosen from among those of the input distribution. While the continuous scenario reduction problem offers more flexibility and is therefore guaranteed to find (weakly) better approximations to the initial empirical distribution, the existing stochastic programming literature has exclusively focused on the discrete scenario reduction problem. This is partly due to the fact that continuous scenario reduction generates new scenarios that can be improbable in view of the true probability distribution, especially if its support is disconnected or non-convex. As in discrete scenario reduction the support points are fixed and the scenario reduction problem can be reformulated as a linear program over the probabilities, which admits an explicit solution (Dupačová et al. 2003). The continuous scenario reduction problem, on the other hand, is intractable as it encapsulates the $\mathcal{NP}$-hard metric K-median problem (Kariv and Hakimi 1979) and the $\mathcal{NP}$-hard K-means clustering problem with $K = m$ (Mahajan et al. 2009, Aloise et al. 2009) as special cases.

Heitsch and Römisch (2003) have shown that the discrete scenario reduction problem admits a reformulation as a mixed-integer linear program (MILP), which can be solved to global optimality for $n \lesssim 10^3$ using off-the-shelf solvers. For larger instances, however, one must resort to approximation algorithms. Most large-scale discrete scenario reduction problems are nowadays solved with a greedy heuristic that was originally devised by Dupačová et al. (2003) and further refined by Heitsch and Römisch (2003). For example, this heuristic is routinely used for scenario reduction in the context of power systems operations; see, *e.g.*, Parvania and Fotuhi-Firuzabad (2010), Ruiz et al. (2009) or Conejo et al. (2010a) and the references therein. Despite its practical success, we will show in Chapter 2 of this thesis that this heuristic fails to provide a constant-factor approximation for the discrete scenario reduction problem.

1.2 Clustering

Clustering aims to partition a set of datapoints into a set of clusters so that datapoints in the same cluster are more similar to one another than to those in other clusters. As such, clustering—as a method of unsupervised machine learning—can be viewed as a synonym for scenario reduction, as used in stochastic programming. In fact, any clustering algorithm can be used for scenario reduction and vice versa. Among the myriad of clustering approaches from the literature, *K-means clustering* stands out for its long history dating back to 1957 as well as its impressive performance in various application domains, ranging from market segmentation and recommender systems to image segmentation and feature learning (Jain 2010b).

An important variant of the standard clustering problem is the *cardinality-constrained clustering problem*, which is defined as the task of partitioning n datapoints into K clusters of prescribed sizes, so as to minimize the sum of the squared intra-cluster distances. The cardinality-constrained clustering problem is of interest for the following reasons. Firstly, it has been shown by Bennett et al. (2000) and Chen et al. (2006) that the algorithms commonly employed for the *un*constrained K-means clustering problem frequently produce suboptimal solutions where some of the clusters contain very few or even no datapoints. In this context, cardinality constraints can act as a regularizer that avoids local minima of poor quality. Secondly, many application domains require the clusters to be of comparable size. This is the case, among others, in distributed clustering (where different computer clusters should contain similar numbers of network nodes), market segmentation (where each customer segment will subsequently be addressed by a marketing campaign) and document clustering (where topic hierarchies should display a balanced view of the available documents); see Banerjee and Ghosh (2006) and Balcan et al. (2013). Finally, and perhaps most importantly, K-means clustering is highly sensitive to outliers. A comprehensive and principled treatment of outlier detection methods can be found in the book of Aggarwal (2013).

To date only two solution approaches have been proposed for the cardinality-constrained clustering problem. Bennett et al. (2000) combine a classical local search heuristic for the unconstrained K-means clustering problem due to Lloyd (1982) with the repeated solution of linear assignment problems to solve a variant of problem (3.1) that imposes lower bounds on the cluster sizes. Banerjee and Ghosh (2006) solve the balanced version of problem (3.1), where all clusters must have the same size, by sampling a subset of the datapoints, performing a clustering on this subset, and subsequently populating the resulting clusters with the remaining datapoints while adhering to the cardinality constraints. Balanced clustering is also considered by Malinen and Fränti (2014) and Costa et al. (2017). Malinen and Fränti (2014) proceed similarly to Bennett et al. (2000) but take explicit advantage of the Hungarian algorithm to speed up the cluster assignment step within the local search heuristic. Costa et al. (2017) propose a variable neighborhood search heuristic that starts from a random partition of the datapoints into balanced clusters and subsequently searches for better solutions in the neighborhood obtained by an increasing number of datapoint swaps between two clusters.

Although all of these heuristics tend to quickly produce solutions of high quality, they are not known to be polynomial-time algorithms, they do not provide bounds on the suboptimality of the identified solutions, and their performance may be sensitive to the choice of the initial solution. Moreover, neither of these local search schemes accounts for outliers.

In recent years, several conic optimization schemes have been proposed to alleviate the shortcomings of these local search methods for the unconstrained K-means clustering problem (Peng and Wei 2007, Awasthi et al. 2015). Peng and Wei (2007) develop two semidefinite programming relaxations of the unconstrained K-means clustering problem. Their weaker relaxation admits optimal solutions that can be characterized by means of an eigenvalue decomposition. They further use this eigenvalue decomposition to set up a modified K-means clustering problem where the dimensionality of the datapoints is reduced to $K-1$ (provided that their original dimensionality was larger than that). To obtain an upper bound, they solve this K-means clustering problem of reduced dimensionality, which can be done either exactly by enumerating Voronoi partitions, as described in Inaba et al. (1994), or by approximation methods such as those in Hasegawa et al. (1993). Using either approach, the runtime grows polynomially in the number of datapoints n but not in the number of desired clusters K. Hence, this method is primarily suitable for small K. Similar conic approximation schemes have been developed by Elhamifar et al. (2012) and Nellore and Ward (2015) in the context of unconstrained exemplar-based clustering.

Awasthi et al. (2015) and Iguchi et al. (2017) develop probabilistic recovery guarantees for the stronger semidefinite relaxation of Peng and Wei (2007) when the data is generated by a stochastic ball model (*i.e.,* datapoints are drawn randomly from rotation symmetric distributions supported on unit balls). More specifically, they use primal-dual arguments to establish conditions on the cluster separation under which the semidefinite relaxation of Peng and Wei (2007) recovers the underlying clusters with high probability as the number of datapoints n increases. The condition of Awasthi et al. (2015) requires less separation in low dimensions, while the condition of Iguchi et al. (2017) is less restrictive in high dimensions. In addition, Awasthi et al. (2015) consider a linear programming relaxation of the unconstrained K-means clustering problem, and they derive similar recovery guarantees for this relaxation as well.

Two more papers study the recovery guarantees of conic relaxations under a stochastic block model (*i.e.,* the dataset is characterized by a similarity matrix where the expected pairwise similarities of points in the same cluster are higher than those of points in different clusters). Ames (2014) considers the densest K-disjoint-clique problem whose aim is to split a given complete graph into K subgraphs such as to maximize the sum of the average similarities of the resulting subgraphs. K-means clustering can be considered as a specific instance of this broader class of problems. By means of primal-dual arguments, the author derives conditions on the means in the stochastic block model such that his semidefinite relaxation recovers the underlying clusters with high probability as the cardinality of the smallest cluster increases. Vinayak and Hassibi (2016) develop a semidefinite relaxation and regularize it with the trace of the cluster assignment matrix. Using primal-dual arguments they show that, for specific

ranges of the regularization parameter, their regularized semidefinite relaxation recovers the true clusters with high probability as the cardinality of the smallest cluster increases. The probabilistic recovery guarantees of Ames (2014) and Vinayak and Hassibi (2016) can also be extended to datasets containing outliers.

1.3 Contributions and Structure of the Thesis

This thesis is structured as follows.

Chapter 2 takes a fresh look at the classical results in the scenario reduction literature. We study both continuous scenario reduction, where the atoms of the reduced distribution may be chosen freely, and discrete scenario reduction, where the atoms of the reduced distribution must be chosen from among those of the input distribution. Using the Wasserstein distance as measure of proximity between distributions, we characterize those distributions that are least susceptible to scenario reduction, *i.e.*, that have maximum Wasserstein distance to their closest m-point distributions for some prescribed $m < n$. We also provide sharp bounds on the added benefit of continuous over discrete scenario reduction. Finally, we propose the first polynomial-time constant-factor approximations for both discrete and continuous scenario reduction as well as the first exact exponential-time algorithms for continuous scenario reduction. The material of Chapter 2 has been published in the following journal paper.

- Napat Rujeerapaiboon, Kilian Schindler, Daniel Kuhn, Wolfram Wiesemann. Scenario reduction revisited: Fundamental limits and guarantees. *Mathematical Programming* (in press, published online on 6 April 2018).

Chapter 3 views the scenario reduction problem from the perspective of clustering and addresses two major shortcomings of the mainstream K-means clustering algorithm: the approach suffers from outlier sensitivity and may produce highly unbalanced clusters. This is undesirable when K-means clustering is used to simplify the distribution of a stochastic program because outliers could have a detrimental effect on its optimal solutions and because unbalanced clusters give rise to non-uniform approximating distributions that eventually make inefficient use of the available computational resources. To mitigate both shortcomings, we formulate a joint outlier detection and clustering problem, which assigns a prescribed number of datapoints of the empirical input distribution to an auxiliary outlier cluster and performs cardinality-constrained K-means clustering on the residual dataset, treating the cluster cardinalities as a given input. We cast this problem as a mixed-integer linear program (MILP) that admits tractable semidefinite and linear programming relaxations. We propose deterministic rounding schemes that transform the relaxed solutions to feasible solutions for the MILP. We also prove that these solutions are optimal in the MILP if a cluster separation condition holds. The material underlying Chapter 2 has been published in the following journal paper.

- Napat Rujeerapaiboon, Kilian Schindler, Daniel Kuhn, Wolfram Wiesemann. Size Matters: Cardinality-Constrained Clustering and Outlier Detection via Conic Optimization. *SIAM Journal on Optimization* 29(2), 1211–1239, 2019.

Chapter 4 develops a highly efficient scenario reduction method for a large-scale hydro scheduling problem. Specifically, we study the optimal operation of a fleet of interconnected hydropower plants that sell energy on both the spot and the reserve markets, and we propose a two-layer stochastic programming framework for its solution. The outer layer problem (the *planner's problem*) optimizes the end-of-day reservoir filling levels over one year, whereas the inner layer problem (the *trader's problem*) selects optimal hourly market bids within each day. Using an information restriction whereby the planner prescribes the end-of-day reservoir targets one day in advance, we prove that the trader's problem simplifies from an infinite-dimensional stochastic program with 25 stages to a finite two-stage stochastic program with only two scenarios. We thus reduce the scenarios of the problem dramatically without any loss of optimality. Substituting the reduced reformulation back into the outer layer and approximating the reservoir targets by affine decision rules then allows us to simplify the planner's problem from an infinite-dimensional stochastic program with 365 stages to a two-stage stochastic program that can conveniently be solved via the sample average approximation. The material underlying Chapter 4 originates from the following working paper.

- Napat Rujeerapaiboon, Kilian Schindler, Daniel Kuhn, Wolfram Wiesemann. A Planner-Trader Decomposition for Multi-Market Hydro Scheduling. Submitted to *Operations Research*, 2020.

Chapter 5 distills the key insights of the thesis and outlines future research directions.

1.4 Statement of Originality

I hereby certify that the content of this thesis is the product of my own work with some assistance from my supervisor Prof. Daniel Kuhn as well as my co-authors Prof. Wolfram Wiesemann and Prof. Napat Rujeerapaiboon.

2 Scenario Reduction Revisited: Fundamental Limits and Guarantees

The goal of scenario reduction is to approximate a given discrete distribution with another discrete distribution that has fewer atoms. We distinguish continuous scenario reduction, where the new atoms may be chosen freely, and discrete scenario reduction, where the new atoms must be chosen from among the existing ones. Using the Wasserstein distance as measure of proximity between distributions, we identify those n-point distributions on the unit ball that are least susceptible to scenario reduction, *i.e.*, that have maximum Wasserstein distance to their closest m-point distributions for some prescribed $m < n$. We also provide sharp bounds on the added benefit of continuous over discrete scenario reduction. Finally, to our best knowledge, we propose the first polynomial-time constant-factor approximations for both discrete and continuous scenario reduction as well as the first exact exponential-time algorithms for continuous scenario reduction.

2.1 Introduction

The vast majority of numerical solution schemes in stochastic programming rely on a discrete approximation of the true (typically continuous) probability distribution governing the uncertain problem parameters. This discrete approximation is often generated by sampling from the true distribution. Alternatively, it could be constructed directly from real historical observations of the uncertain parameters. To obtain a faithful approximation for the true distribution, however, the discrete distribution must have a large number n of support points or *scenarios*, which may render the underlying stochastic program computationally excruciating.

An effective means to ease the computational burden is to rely on *scenario reduction* pioneered by Dupačová et al. (2003), which aims to approximate the initial n-point distribution with a simpler m-point distribution ($m < n$) that is as close as possible to the initial distribution with respect to a probability metric; see also Heitsch and Römisch (2003). The modern stability theory of stochastic programming indicates that the Wasserstein distance may serve as a natural candidate for this probability metric. Indeed, Römisch and Schultz (1991) show that, under suitable regularity conditions, the optimal values of two-stage stochastic programs with

random right-hand sides are Lipschitz continuous in the distribution of the uncertainty if the space of probability distributions is equipped with a Wasserstein metric. They further prove that the optimal solution mapping is Hölder continuous in the Wasserstein metric. Similar continuity results are established for chance constrained programs. A stronger quantitative stability theory of stochastic programming relies on the Fortet-Mourier metric, which generalizes the Wasserstein metric but offers improved convergence rates; see Rachev and Römisch (2002). For a survey of classical stability results in stochastic programming see Dupačová (1990) or Römisch and Wets (2007).

Our interest in Wasserstein distance-based scenario reduction is also fuelled by recent progress in data-driven distributionally robust optimization, where it has been shown that the worst-case expectation of an uncertain cost over all distributions in a Wasserstein ball can often be computed efficiently via convex optimization (Mohajerin Esfahani and Kuhn 2017, Zhao and Guan 2015, Gao and Kleywegt 2016). A Wasserstein ball is defined as the family of all distributions that are within a certain Wasserstein distance from a discrete reference distribution. As distributionally robust optimization problems over Wasserstein balls are harder to solve than their stochastic counterparts, we expect significant computational savings from replacing the initial n-point reference distribution with a new m-point reference distribution. The benefits of scenario reduction may be particularly striking for two-stage distributionally robust linear programs, which admit tight approximations as semidefinite programs (Hanasusanto and Kuhn 2016).

Suppose now that the initial distribution is given by $\mathbb{P} = \sum_{i\in I} p_i \delta_{\boldsymbol{\xi}_i}$, where $\boldsymbol{\xi}_i \in \mathbb{R}^d$ and $p_i \in [0,1]$ represent the location and probability of the i-th scenario of $\mathbb{P}$ for $i \in I = \{1,\ldots,n\}$. Similarly, assume that the reduced target distribution is representable as $\mathbb{Q} = \sum_{j\in J} q_j \delta_{\boldsymbol{\zeta}_j}$, where $\boldsymbol{\zeta}_j \in \mathbb{R}^d$ and $q_j \in [0,1]$ stand for the location and probability of the j-th scenario of $\mathbb{Q}$ for $j \in J = \{1,\ldots,m\}$. Then, the type-l Wasserstein distance between $\mathbb{P}$ and $\mathbb{Q}$ is defined through

$$d_l(\mathbb{P},\mathbb{Q}) = \left[\min_{\Pi\in\mathbb{R}_+^{n\times m}} \left\{ \sum_{i\in I}\sum_{j\in J} \pi_{ij} \|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j\|^l : \begin{array}{l} \sum_{j\in J}\pi_{ij} = p_i \;\forall i \in I \\ \sum_{i\in I}\pi_{ij} = q_j \;\forall j \in J \end{array} \right\} \right]^{1/l},$$

where $l \geq 1$ and $\|\cdot\|$ denotes some norm on $\mathbb{R}^d$; see, *e.g.*, Heitsch and Römisch (2007) or Pflug and Pichler (2011). The linear program in the definition of the Wasserstein distance can be viewed as a minimum-cost transportation problem, where π_{ij} represents the amount of probability mass shipped from $\boldsymbol{\xi}_i$ to $\boldsymbol{\zeta}_j$ at unit transportation cost $\|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j\|^l$. Thus, $d_l^l(\mathbb{P},\mathbb{Q})$ quantifies the minimum cost of moving the initial distribution $\mathbb{P}$ to the target distribution $\mathbb{Q}$.

For any $\Xi \subseteq \mathbb{R}^d$, we denote by $\mathscr{P}_{\mathrm{E}}(\Xi, n)$ the set of all uniform discrete distributions on Ξ with exactly n distinct scenarios and by $\mathscr{P}(\Xi, m)$ the set of all (not necessarily uniform) discrete distributions on Ξ with at most m scenarios. We henceforth assume that $\mathbb{P} \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$. This assumption is crucial for the simplicity of the results in Sections 2.2 and 2.3, and it is almost surely satisfied whenever $\mathbb{P}$ is obtained via sampling from a continuous probability distribution. Hence, we can think of $\mathbb{P}$ as an *empirical distribution*. To remind us of this

interpretation, we will henceforth denote the initial distribution by $\hat{\mathbb{P}}_n$. Note that the pairwise difference of the scenarios can always be enforced by slightly perturbing their locations, while the uniformity of their probabilities can be enforced by decomposing the scenarios into clusters of close but mutually distinct sub-scenarios with (smaller) uniform probabilities.

We are now ready to introduce the *continuous scenario reduction problem*

$$C_l(\hat{\mathbb{P}}_n, m) \;=\; \min_{\mathbb{Q}} \left\{ d_l(\hat{\mathbb{P}}_n, \mathbb{Q}) : \mathbb{Q} \in \mathscr{P}(\mathbb{R}^d, m) \right\},$$

where the new scenarios $\boldsymbol{\zeta}_j$, $j \in J$, of the target distribution $\mathbb{Q}$ may be chosen freely from within $\mathbb{R}^d$, as well as the *discrete scenario reduction problem*

$$D_l(\hat{\mathbb{P}}_n, m) \;=\; \min_{\mathbb{Q}} \left\{ d_l(\hat{\mathbb{P}}_n, \mathbb{Q}) : \mathbb{Q} \in \mathscr{P}(\mathrm{supp}(\hat{\mathbb{P}}_n), m) \right\},$$

where the new scenarios must be chosen from within the support of the empirical distribution, which is given by the finite set $\mathrm{supp}(\hat{\mathbb{P}}_n) = \{\boldsymbol{\xi}_i : i \in I\}$. The terms discrete and continuous scenario reduction are chosen because of their intimate relationship to discrete and continuous facility location, respectively; see *e.g.*, Drezner and Hamacher (2004). Note that the literature sometimes refers to continuous scenario reduction as *scenario generation* (Löhndorf 2016).

Even though the continuous scenario reduction problem offers more flexibility and is therefore guaranteed to find (weakly) better approximations to the initial empirical distribution, to our best knowledge, the existing stochastic programming literature has exclusively focused on the discrete scenario reduction problem. This is partly due to the fact that continuous scenario reduction generates new scenarios that can be improbable in view of the true probability distribution, especially if its support is disconnected or non-convex.

Note that if the support points $\boldsymbol{\zeta}_j$, $j \in J$, are fixed, then both scenario reduction problems simplify to a linear program over the probabilities q_j, $j \in J$, which admits an explicit solution (Dupačová et al. 2003, Theorem 2). Otherwise, however, both problems are intractable. Indeed, if $l = 1$, then the discrete scenario reduction problem represents a metric k-median problem with $k = m$, which was shown to be $\mathscr{NP}$-hard by Kariv and Hakimi (1979). If $l = 2$ and distances in $\mathbb{R}^d$ are measured by the 2-norm, on the other hand, then the continuous scenario reduction problem constitutes a k-means clustering problem with $k = m$, which is $\mathscr{NP}$-hard even if $d = 2$ or $m = 2$; see Mahajan et al. (2009) and Aloise et al. (2009). For a comprehensive review of data clustering approaches, we refer to Jain et al. (1999).

Heitsch and Römisch (2003) have shown that the discrete scenario reduction problem admits a reformulation as a mixed-integer linear program (MILP), which can be solved to global optimality for $n \lesssim 10^3$ using off-the-shelf solvers. For larger instances, however, one must resort to approximation algorithms. Most large-scale discrete scenario reduction problems are nowadays solved with a greedy heuristic that was originally devised by Dupačová et al. (2003) and further refined by Heitsch and Römisch (2003). For example, this heuristic is routinely used for scenario reduction in the context of power systems operations; see, *e.g.*, Parvania and

Fotuhi-Firuzabad (2010), Ruiz et al. (2009) or Conejo et al. (2010a) and the references therein. Despite its practical success, we will show in Section 2.4 that this heuristic fails to provide a constant-factor approximation for the discrete scenario reduction problem.

This paper extends the theory of scenario reduction along several dimensions.

(i) We establish fundamental performance guarantees for continuous scenario reduction when $l \in \{1,2\}$, *i.e.*, we show that the Wasserstein distance of the initial n-point distribution to its nearest m-point distribution is bounded by $\sqrt{\frac{n-m}{n-1}}$ across all initial distributions on the unit ball in $\mathbb{R}^d$. We show that for $l = 2$ this worst-case performance is attained by some initial distribution, which we construct explicitly. We also provide evidence indicating that this worst-case performance reflects the norm rather than the exception in high dimensions d. Finally, we provide a lower bound on the worst-case performance for $l = 1$.

(ii) We analyze the loss of optimality incurred by solving the discrete scenario reduction problem instead of its continuous counterpart. Specifically, we demonstrate that the ratio $D_l(\hat{\mathbb{P}}_n, m)/C_l(\hat{\mathbb{P}}_n, m)$ is bounded by $\sqrt{2}$ for $l = 2$ and by 2 for $l = 1$. We also show that these bounds are essentially tight.

(iii) We showcase the intimate relation between scenario reduction and k-median clustering. By leveraging existing constant-factor approximation algorithms for k-median clustering problems due to Arya et al. (2004) and the new performance bounds from (ii), we develop the first polynomial-time constant-factor approximation algorithms for both continuous and discrete scenario reduction. We also show that these algorithms can be warmstarted using the greedy heuristic by Dupačová et al. (2003) to improve practical performance.

(iv) We present exact mixed-integer programming reformulations for the continuous scenario reduction problem.

Continuous scenario reduction is intimately related to the optimal quantization of probability distributions, where one seeks an m-point distribution approximating a non-discrete initial distribution. Research efforts in this domain have mainly focused on the asymptotic behavior of the quantization problem as m tends to infinity, see Graf and Luschgy (2000). The ramifications of this stream of literature for stochastic programming are discussed by Pflug and Pichler (2011). Techniques familiar from scenario reduction lend themselves also for scenario tree generation, where one aims to construct a scenario tree with a prescribed branching structure that approximates a given stochastic process with respect to a probability metric, see, *e.g.*, Pflug (2001) and Hochreiter and Pflug (2007).

The rest of this paper unfolds as follows. Section 2.2 seeks to identify n-point distributions on the unit ball that are least susceptible to scenario reduction, *i.e.*, that have maximum Wasserstein distance to their closest m-point distributions, and Section 2.3 discusses sharp bounds

on the added benefit of continuous over discrete scenario reduction. Section 2.4 presents exact exponential-time algorithms as well as polynomial-time constant-factor approximations for scenario reduction. Section 2.5 reports on numerical results for a color quantization and a capital budgeting problem. Unless otherwise specified, below we will always work with the 2-norm on $\mathbb{R}^d$.

Notation: We let $\mathbb{I}$ be the identity matrix, $\mathbf{e}$ the vector of all ones and $\mathbf{e}_i$ the i-th standard basis vector of appropriate dimensions. The ij-th element of a matrix $\mathbf{A}$ is denoted by a_{ij}. For $\mathbf{A}$ and $\mathbf{B}$ in the space $\mathbb{S}^n$ of symmetric $n \times n$ matrices, the relation $\mathbf{A} \succeq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive semidefinite. Generic norms are denoted by $\|\cdot\|$, while $\|\cdot\|_p$ stands for the p-norm, $p \geq 1$. For $\Xi \subseteq \mathbb{R}^d$, we define $\mathscr{P}(\Xi, m)$ as the set of all probability distributions supported on at most m points in Ξ and $\mathscr{P}_{\mathrm{E}}(\Xi, n)$ as the set of all *uniform* distributions supported on exactly n *distinct* points in Ξ. The support of a probability distribution $\mathbb{P}$ is denoted by $\mathrm{supp}(\mathbb{P})$, and the Dirac distribution concentrating unit mass at $\boldsymbol{\xi}$ is denoted by $\delta_{\boldsymbol{\xi}}$.

2.2 Fundamental Limits of Scenario Reduction

In this section we characterize the Wasserstein distance $C_l(\hat{\mathbb{P}}_n, m)$ between an n-point empirical distribution $\hat{\mathbb{P}}_n = \frac{1}{n}\sum_{i=1}^n \delta_{\boldsymbol{\xi}_i}$ and its continuously reduced optimal m-point distribution $\mathbb{Q} \in \mathscr{P}(\mathbb{R}^d, m)$. Since the positive homogeneity of the Wasserstein distance d_l implies that $C_l(\hat{\mathbb{P}}'_n, m) = \lambda \cdot C_l(\hat{\mathbb{P}}_n, m)$ for the scaled distribution $\hat{\mathbb{P}}'_n = \frac{1}{n}\sum_{i=1}^n \delta_{\lambda\boldsymbol{\xi}_i}$, $\lambda \in \mathbb{R}_+$, we restrict ourselves to empirical distributions $\hat{\mathbb{P}}_n$ whose scenarios satisfy $\|\boldsymbol{\xi}_i\|_2 \leq 1$, $i = 1, \ldots, n$. We thus want to quantify

$$\overline{C}_l(n, m) = \max_{\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)} \left\{ C_l(\hat{\mathbb{P}}_n, m) : \|\boldsymbol{\xi}\|_2 \leq 1 \ \ \forall \boldsymbol{\xi} \in \mathrm{supp}(\hat{\mathbb{P}}_n) \right\}, \tag{2.1}$$

which amounts to the worst-case (*i.e.*, largest) Wasserstein distance between any n-point empirical distribution $\hat{\mathbb{P}}_n$ over the unit ball and its optimally selected continuous m-point scenario reduction. By construction, this worst-case distance satisfies $\overline{C}_l(n, m) \geq 0$, and the lower bound is attained whenever $n = m$. One also verifies that $\overline{C}_l(n, m) \leq \overline{C}_l(n, 1) \leq 1$ since the Wasserstein distance to the Dirac distribution $\delta_{\mathbf{0}}$ is bounded above by 1. Our goal is to derive possibly tight upper bounds on $\overline{C}_l(n, m)$ for the Wasserstein distances of type $l \in \{1, 2\}$.

In the following, we denote by $\mathfrak{P}(I, m)$ the family of all m-set partitions of the index set I, *i.e.*,

$$\mathfrak{P}(I, m) = \left\{ \{I_1, \ldots, I_m\} : \emptyset \neq I_1, \ldots, I_m \subseteq I, \ \cup_j I_j = I, \ I_i \cap I_j = \emptyset \ \ \forall i \neq j \right\},$$

and an element of this set (*i.e.*, a specific m-set partition) as $\{I_j\} \in \mathfrak{P}(I, m)$. Our derivations will make extensive use of the following theorem.

Theorem 2.1. For any type-l Wasserstein distance induced by any norm $\|\cdot\|$, the continuous

scenario reduction problem can be reformulated as

$$C_l(\hat{\mathbb{P}}_n, m) = \min_{\{I_j\}\in\mathfrak{P}(I,m)} \left[\frac{1}{n} \sum_{j\in J} \min_{\boldsymbol{\zeta}_j\in\mathbb{R}^d} \sum_{i\in I_j} \|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j\|^l \right]^{1/l}. \tag{2.2}$$

Problem (2.2) can be interpreted as a Voronoi partitioning problem that asks for a Voronoi decomposition of $\mathbb{R}^d$ into m cells whose Voronoi centroids $\boldsymbol{\zeta}_1,\ldots,\boldsymbol{\zeta}_m$ minimize the cumulative l-th powers of the distances to n prespecified points $\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_n$.

of Theorem 2.1. Theorem 2 of Dupačová et al. (2003) implies that the smallest Wasserstein distance between the empirical distribution $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ and any distribution $\mathbb{Q}$ supported on a finite set $\Xi \subset \mathbb{R}^d$ amounts to

$$\min_{\mathbb{Q}\in\mathscr{P}(\Xi,\infty)} d_l(\hat{\mathbb{P}}_n, \mathbb{Q}) = \left[\frac{1}{n} \sum_{i\in I} \min_{\boldsymbol{\zeta}\in\Xi} \|\boldsymbol{\xi}_i - \boldsymbol{\zeta}\|^l \right]^{1/l},$$

where $\mathscr{P}(\Xi,\infty)$ denotes the set of all probability distributions supported on the finite set Ξ. The continuous scenario reduction problem $C_l(\hat{\mathbb{P}}_n, m)$ selects the set $\Xi^\star$ that minimizes this quantity over all sets in $\Xi \subset \mathbb{R}^d$ with $|\Xi| = m$ elements:

$$C_l(\hat{\mathbb{P}}_n, m) = \min_{\{\boldsymbol{\zeta}_j\}\subseteq\mathbb{R}^d} \left[\frac{1}{n} \sum_{i\in I} \min_{j\in J} \|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j\|^l \right]^{1/l}. \tag{2.3}$$

One readily verifies that any optimal solution $\{\boldsymbol{\zeta}_1^\star,\ldots,\boldsymbol{\zeta}_m^\star\}$ to problem (2.3) corresponds to a feasible solution $\{I_1^\star,\ldots,I_m^\star\}$ to problem (2.2) with the same objective value if we identify the set $I_j^\star$ with all observations $\boldsymbol{\xi}_i$ that are closer to $\boldsymbol{\zeta}_j^\star$ than any other $\boldsymbol{\zeta}_{j'}^\star$ (ties may be broken arbitrarily). Likewise, any optimal solution $\{I_1^\star,\ldots,I_m^\star\}$ to problem (2.2) with inner minimizers $\{\boldsymbol{\zeta}_1^\star,\ldots,\boldsymbol{\zeta}_m^\star\}$ translates into a feasible solution $\{\boldsymbol{\zeta}_1^\star,\ldots,\boldsymbol{\zeta}_m^\star\}$ to problem (2.3) with the same objective value. □

Remark 2.2. *(Minimizers of* (2.2)*)* For $l = 2$, the inner minimum corresponding to the set I_j is attained by the *mean* $\boldsymbol{\zeta}_j^\star = \mathrm{mean}(I_j) = \frac{1}{|I_j|}\sum_{i\in I_j} \boldsymbol{\xi}_i$. Likewise, for $l = 1$, the inner minimum corresponding to the set I_j is attained by any *geometric median*

$$\boldsymbol{\zeta}_j^\star = \mathrm{gmed}(I_j) \in \operatorname*{arg\,min}_{\boldsymbol{\zeta}_j\in\mathbb{R}^d} \sum_{i\in I_j} \|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j\|,$$

which can be determined efficiently by solving a second-order cone program whenever a p-norm with rational $p \geq 1$ is considered (Alizadeh and Goldfarb (2003)).

Instead of setting the scenarios of the reduced m-point distribution to the centroids of the Voronoi cells found in (2.2), Löhndorf (2016) suggests to select one *random* scenario from within each Voronoi cell and argues that the resulting reduced distribution captures the dispersion of the original distribution more accurately.

The rest of this section derives tight upper bounds on $\overline{C}_l(n,m)$ for Wasserstein distances of type $l=2$ (Section 2.2.1) as well as upper and lower bounds for Wasserstein distances of type $l=1$ (Section 2.2.2). We summarize and discuss our findings in Section 2.2.3.

2.2.1 Fundamental Limits for the Type-2 Wasserstein Distance

We now derive a revised upper bound on $\overline{C}_l(n,m)$ for the type-2 Wasserstein distance. The result relies on auxiliary lemmas that are relegated to the appendix.

Theorem 2.3. The worst-case type-2 Wasserstein distance satisfies $\overline{C}_2(n,m) \leq \sqrt{\frac{n-m}{n-1}}$.

Note that whenever the reduced distribution satisfies $m>1$, the bound of Theorem 2.3 is strictly tighter than the naïve bound of 1 from the previous section.

Proof of Theorem 2.3. From Theorem 2.1 and Remark 2.2 we observe that

$$\begin{array}{rcll} \overline{C}_2(n,m) & = & \displaystyle\max_{\{\boldsymbol{\xi}_i\}\subseteq\mathbb{R}^d} & \displaystyle\min_{\{I_j\}\in\mathfrak{P}(I,m)} \left[\frac{1}{n}\sum_{j\in J}\sum_{i\in I_j}\left\|\boldsymbol{\xi}_i-\mathrm{mean}(I_j)\right\|_2^2\right]^{1/2} \\ & & \text{s.t.} & \|\boldsymbol{\xi}_i\|_2\leq 1 \quad \forall i\in I. \end{array}$$

Introducing the epigraphical variable τ, this problem can be expressed as

$$\begin{array}{rcll} \overline{C}_2^2(n,m) & = & \displaystyle\max_{\tau\in\mathbb{R},\,\{\boldsymbol{\xi}_i\}\subseteq\mathbb{R}^d} & \displaystyle\frac{1}{n}\tau \\ & & \text{s.t.} & \displaystyle\tau\leq\sum_{j\in J}\sum_{i\in I_j}\left\|\boldsymbol{\xi}_i-\mathrm{mean}(I_j)\right\|_2^2 \quad \forall\{I_j\}\in\mathfrak{P}(I,m) \\ & & & \boldsymbol{\xi}_i{}^\top\boldsymbol{\xi}_i\leq 1 \qquad\qquad\qquad\qquad\quad \forall i\in I. \end{array} \tag{2.4}$$

For each $j\in J$ and $i\in I_j$, the squared norm in the first constraint of (2.4) can be expressed in terms of the inner products between pairs of empirical observations:

$$\begin{aligned} \left\|\boldsymbol{\xi}_i-\mathrm{mean}(I_j)\right\|_2^2 &= \frac{1}{|I_j|^2}\left\||I_j|\boldsymbol{\xi}_i-\sum_{k\in I_j}\boldsymbol{\xi}_k\right\|_2^2 \\ &= \frac{1}{|I_j|^2}\left(|I_j|^2\boldsymbol{\xi}_i{}^\top\boldsymbol{\xi}_i-2|I_j|\sum_{k\in I_j}\boldsymbol{\xi}_i{}^\top\boldsymbol{\xi}_k+\sum_{k\in I_j}\boldsymbol{\xi}_k{}^\top\boldsymbol{\xi}_k+\sum_{\substack{k,k'\in I_j\\ k\neq k'}}\boldsymbol{\xi}_k{}^\top\boldsymbol{\xi}_{k'}\right). \end{aligned}$$

Introducing the Gram matrix

$$\mathbf{S}=[\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_n]^\top[\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_n]\in\mathbb{S}^n,\ \mathbf{S}\succeq\mathbf{0}\ \text{and}\ \mathrm{rank}(\mathbf{S})\leq\min\{n,d\} \tag{2.5}$$

then allows us to simplify the first constraint in (2.4) to

$$\tau \leq \sum_{j\in J} \frac{1}{|I_j|^2} \sum_{i\in I_j} \left(|I_j|^2 s_{ii} - 2|I_j| \sum_{k\in I_j} s_{ik} + \sum_{k\in I_j} s_{kk} + \sum_{\substack{k,k'\in I_j \\ k\neq k'}} s_{kk'} \right).$$

Note that the second constraint in problem (2.4) can now be expressed as $s_{ii} \leq 1$, and hence all constraints in (2.4) are linear in the Gram matrix $\mathbf{S}$.

Our discussion implies that we obtain an upper bound on $\overline{C}_2(n,m)$ by reformulating problem (2.4) as a semidefinite program in terms of the Gram matrix $\mathbf{S}$

$$\begin{array}{ll} \max\limits_{\tau\in\mathbb{R},\, \mathbf{S}\in\mathbb{S}^n} & \frac{1}{n}\tau \\ \text{s.t.} & \tau \leq \sum\limits_{j\in J} \frac{1}{|I_j|^2} \sum\limits_{i\in I_j} \left(|I_j|^2 s_{ii} - 2|I_j| \sum\limits_{k\in I_j} s_{ik} + \sum\limits_{k\in I_j} s_{kk} + \sum\limits_{\substack{k,k'\in I_j \\ k\neq k'}} s_{kk'} \right) \\ & \qquad\qquad \forall \{I_j\} \in \mathfrak{P}(I,m) \\ & \mathbf{S} \succeq \mathbf{0},\ s_{ii} \leq 1 \quad \forall i \in I, \end{array} \tag{2.6}$$

where we have relaxed the rank condition in the definition of the Gram matrix (2.5). Lemma 2.21 in the appendix shows that (2.6) has an optimal solution $(\tau^\star, \mathbf{S}^\star)$ that satisfies $\mathbf{S}^\star = \alpha\mathbb{I} + \beta\mathbf{e}\mathbf{e}^\top$ for some $\alpha, \beta \in \mathbb{R}$. Moreover, Lemma 2.22 in the appendix shows that any matrix of the form $\mathbf{S} = \alpha\mathbb{I} + \beta\mathbf{1}\mathbf{1}^\top$ is positive semidefinite if and only if $\alpha \geq 0$ and $\alpha + n\beta \geq 0$. We thus conclude that (2.6) can be reformulated as

$$\begin{array}{ll} \max\limits_{\tau,\alpha,\beta\in\mathbb{R}} & \frac{1}{n}\tau \\ \text{s.t.} & \tau \leq (n-m)\alpha,\ \alpha+\beta \leq 1 \\ & \alpha \geq 0,\ \alpha + n\beta \geq 0, \end{array} \tag{2.7}$$

where the first constraint follows from the fact that for any set I_j in (2.6), we have

$$\frac{1}{|I_j|^2} \sum_{i\in I_j} \left(|I_j|^2(\alpha+\beta) - 2|I_j|\big(\alpha + |I_j|\beta\big) + \sum_{k\in I_j} (\alpha+\beta) + \sum_{\substack{k,k'\in I_j \\ k\neq k'}} \beta \right) = (|I_j|-1)\alpha,$$

and $\sum_{j\in J}(|I_j|-1)\alpha = (n-m)\alpha$ since $|I| = n$ and $|J| = m$. The statement of the theorem now follows since problem (2.7) is optimized by $\tau^\star = \frac{n(n-m)}{(n-1)}$, $\alpha^\star = \frac{n}{n-1}$ and $\beta^\star = \frac{-1}{n-1}$. □

The proof of Theorem 2.3 shows that the upper bound $\sqrt{\frac{n-m}{n-1}}$ on the worst-case type-2 Wasserstein distance $\overline{C}_2(n,m)$ is tight whenever there is an empirical distribution $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ whose scenarios $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$ correspond to a Gram matrix $\mathbf{S} = [\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n]^\top [\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n] = \frac{n}{n-1}\mathbb{I} - \frac{1}{n-1}\mathbf{e}\mathbf{e}^\top$, which implies $\|\boldsymbol{\xi}_i\|_2 = \sqrt{s_{ii}} = 1$ for all $i \in I$. We now show that such an empirical distribution exists when $d \geq n-1$.

Proposition 2.4. For $d \geq n-1$, there is $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ with $\|\boldsymbol{\xi}\|_2 \leq 1$ for all $\boldsymbol{\xi} \in \mathit{supp}(\hat{\mathbb{P}}_n)$ such that $C_2(\hat{\mathbb{P}}_n, m) = \sqrt{\frac{n-m}{n-1}}$.

Proof. Assume first that $d = n$ and consider the empirical distribution $\hat{\mathbb{P}}_n = \frac{1}{n}\sum_{i=1}^n \delta_{\boldsymbol{\xi}_i}$ defined through

$$\boldsymbol{\xi}_i = y\mathbf{e} + (x-y)\mathbf{e}_i \in \mathbb{R}^n \quad \text{with} \quad x = \sqrt{\frac{n-1}{n}} \text{ and } y = \frac{-1}{\sqrt{n(n-1)}}. \tag{2.8}$$

A direct calculation reveals that $\mathbf{S} = [\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n]^\top [\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n] = \frac{n}{n-1}\mathbb{I} - \frac{1}{n-1}\mathbf{e}\mathbf{e}^\top$.

To prove the statement for $d = n-1$, we note that the n scenarios in (2.8) lie on the $(n-1)$-dimensional subspace $\mathscr{H}$ orthogonal to $\mathbf{e} \in \mathbb{R}^n$. Thus, there exists a rotation that maps $\mathscr{H}$ to $\mathbb{R}^{n-1} \times \{0\}$. As the Gram matrix is invariant under rotations, the rotated scenarios give rise to an empirical distribution $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^{n-1}, n)$ satisfying the statement of the proposition. Likewise, for $d > n$ the linear transformation $\boldsymbol{\xi}_i \mapsto (\mathbb{I}, \mathbf{0})^\top \boldsymbol{\xi}_i$, $\mathbb{I} \in \mathbb{R}^{n\times n}$ and $\mathbf{0} \in \mathbb{R}^{n\times(d-n)}$, generates an empirical distribution $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ that satisfies the statement of the proposition. □

Proposition 2.4 requires that $d \geq n-1$, which appears to be restrictive. We note, however, that this condition is only sufficient (and not necessary) to guarantee the tightness of the bound from Theorem 2.3. Moreover, we will observe in Section 2.2.3 that the bound of Theorem 2.3 provides surprisingly accurate guidance for the Wasserstein distance between practice-relevant empirical distributions $\hat{\mathbb{P}}_n$ and their continuously reduced optimal distributions.

2.2.2 Fundamental Limits for the Type-1 Wasserstein Distance

In analogy to the previous section, we now derive a revised upper bound on $\overline{C}_l(n,m)$ for the type-1 Wasserstein distance.

Theorem 2.5. The worst-case type-1 Wasserstein distance satisfies $\overline{C}_1(n,m) \leq \sqrt{\frac{n-m}{n-1}}$.

Note that this bound is identical to the bound of Theorem 2.3 for $l = 2$.

Proof of Theorem 2.5. Leveraging again Theorem 2.1 and Remark 2.2, we obtain that

$$\begin{array}{rll} \overline{C}_1(n,m) = & \max\limits_{\{\boldsymbol{\xi}_i\}\subseteq\mathbb{R}^d} & \min\limits_{\{I_j\}\in\mathfrak{P}(I,m)} \dfrac{1}{n}\displaystyle\sum_{j\in J}\sum_{i\in I_j} \left\|\boldsymbol{\xi}_i - \mathrm{gmed}(I_j)\right\|_2 \\ & \text{s.t.} & \|\boldsymbol{\xi}_i\|_2 \leq 1 \quad \forall i \in I. \end{array}$$

We show that $\overline{C}_1(n,m) \leq \overline{C}_2(n,m)$ for all n and $m = 1, \ldots, n$, which in turn proves the statement

of the theorem by virtue of Theorem 2.3. To this end, we observe that

$$\begin{aligned}\overline{C}_1(n,m) \leq \max_{\{\boldsymbol{\xi}_i\}\subseteq\mathbb{R}^d} \min_{\{I_j\}\in\mathfrak{P}(I,m)} \frac{1}{n}\sum_{j\in J}\sum_{i\in I_j}\left\|\boldsymbol{\xi}_i-\mathrm{mean}(I_j)\right\|_2 \\ \text{s.t.} \quad \|\boldsymbol{\xi}_i\|_2\leq 1 \quad \forall i\in I \end{aligned} \tag{2.9}$$

$$\begin{aligned}\leq \max_{\{\boldsymbol{\xi}_i\}\subseteq\mathbb{R}^d} \min_{\{I_j\}\in\mathfrak{P}(I,m)} \left[\frac{1}{n}\sum_{j\in J}\sum_{i\in I_j}\left\|\boldsymbol{\xi}_i-\mathrm{mean}(I_j)\right\|_2^2\right]^{1/2} \\ \text{s.t.} \quad \|\boldsymbol{\xi}_i\|_2\leq 1 \quad \forall i\in I, \end{aligned} \tag{2.10}$$

where the first inequality follows from the definition of the geometric median, which ensures that

$$\sum_{i\in I_j}\left\|\boldsymbol{\xi}_i-\mathrm{gmed}(I_j)\right\|_2 \leq \sum_{i\in I_j}\left\|\boldsymbol{\xi}_i-\mathrm{mean}(I_j)\right\|_2 \quad \forall j\in J,$$

and the second inequality is due to the arithmetic-mean quadratic-mean inequality (Steele 2004, Exercise 2.14). The statement of the theorem now follows from the observation that the optimal value of (2.10) is identical to $\overline{C}_2(n,m)$. □

In the next proposition we derive a lower bound on $\overline{C}_1(n,m)$.

Proposition 2.6. For $d\geq n-1$, the worst-case type-1 Wasserstein distance satisfies $\overline{C}_1(n,m)\geq\sqrt{\frac{(n-m)(n-m+1)}{n(n-1)}}$.

Proof. Assume first that $d=n$, and consider the empirical distribution $\hat{\mathbb{P}}_n$ with scenarios defined as in (2.8). Let $\{I_j\}\in\mathfrak{P}(I,m)$ be an arbitrary m-set partition of I. Note that $\mathrm{gmed}(I_j)=\mathrm{mean}(I_j)$ for every $j\in J$ due to the permutation symmetry of the $\boldsymbol{\xi}_i$, which implies that $\mathbf{0}\in\partial f_j(\mathrm{mean}(I_j))$ for each $f_j(\boldsymbol{\zeta})=\sum_{i\in I_j}\left\|y\mathbf{e}+(x-y)\mathbf{e}_i-\boldsymbol{\zeta}\right\|_2$, $j\in J$. Thus, we have

$$\begin{aligned}&\left\|\boldsymbol{\xi}_i-\mathrm{mean}(I_j)\right\|_2 \\ &= \left[\left(x-\frac{x+(|I_j|-1)y}{|I_j|}\right)^2+(|I_j|-1)\left(y-\frac{x+(|I_j|-1)y}{|I_j|}\right)^2\right]^{1/2} \\ &= \left[\frac{|I_j|-1}{|I_j|}\right]^{1/2}(x-y) = \left[\frac{n(|I_j|-1)}{(n-1)|I_j|}\right]^{1/2} \quad \forall i\in I_j,\end{aligned}$$

where the last equality follows from the definitions of x and y in (2.8). By Theorem 2.1 and Remark 2.2 we therefore obtain

$$\begin{aligned}C_1(\hat{\mathbb{P}}_n,m) &= \min_{\{I_j\}\in\mathfrak{P}(I,m)}\frac{1}{n}\sum_{j\in J}\sum_{i\in I_j}\left\|\boldsymbol{\xi}_i-\mathrm{mean}(I_j)\right\|_2 \\ &= \min_{\{I_j\}\in\mathfrak{P}(I,m)}\frac{1}{\sqrt{n(n-1)}}\sum_{j\in J}\sqrt{|I_j|(|I_j|-1)}.\end{aligned}$$

By introducing auxiliary variables $z_j = |I_j| - 1 \in \mathbb{N}_0$, $j \in J$, we find that determining $C_1(\hat{\mathbb{P}}_n, m)$ is tantamount to solving

$$C_1(\hat{\mathbb{P}}_n, m) \;=\; \frac{1}{\sqrt{n(n-1)}} \min_{\{z_j\}\subseteq\mathbb{N}_0} \left\{ \sum_{j\in J} \sqrt{z_j(z_j+1)} : \sum_{j\in J} z_j = n-m \right\}.$$

Observe that the objective function of $z_1 = n-m$ and $z_2 = \ldots = z_m = 0$ evaluates to $\sqrt{(n-m)(n-m+1)}$, which implies that $C_1(\hat{\mathbb{P}}_n, m) \le \sqrt{\frac{(n-m)(n-m+1)}{n(n-1)}}$. Hence, it remains to establish the reverse inequality. To this end, we note that

$$\begin{aligned}
\sum_{j\in J} \sqrt{z_j(z_j+1)} \;&=\; \left[\sum_{j\in J} z_j(z_j+1) + \sum_{\substack{j,j'\in J \\ j\neq j'}} \sqrt{z_j z_{j'}(1+z_j)(1+z_{j'})} \right]^{1/2} \\
&\geq\; \left[\sum_{j\in J} z_j(z_j+1) + \sum_{\substack{j,j'\in J \\ j\neq j'}} z_j z_{j'} \right]^{1/2} \\
&=\; \left[\left(\sum_{j\in J} z_j \right)^2 + \sum_{j\in J} z_j \right]^{1/2} \;=\; \sqrt{(n-m)(n-m+1)},
\end{aligned}$$

and thus the claim follows for $d = n$. The cases $d = n-1$ and $d > n$ can be reduced to the case $d = n$ as in Proposition 2.4. Details are omitted for brevity. □

Proposition 2.6 asserts that $\overline{C}_1(n,m) \gtrsim \frac{n-m}{n-1} = \overline{C}_2^2(n,m)$ whenever $d \geq n-1$. Together with Theorem 2.5, we thus obtain the following relation between the worst-case Wasserstein distances of types $l = 1$ and $l = 2$:

$$\overline{C}_2^2(n,m) \leq \overline{C}_1(n,m) \leq \overline{C}_2(n,m).$$

We conjecture that the lower bound is tighter, but we were not able to prove this.

2.2.3 Discussion

Theorems 2.3 and 2.5 imply that $\overline{C}_l(n,m) \lesssim \sqrt{1-p}$ for large n and for $l \in \{1,2\}$, where $p = \frac{m}{n}$ represents the desired reduction factor. The significance of this result is that it offers *a priori* guidelines for selecting the number m of support points in the reduced distribution. To see this, consider any empirical distribution $\hat{\mathbb{P}}_n = \frac{1}{n}\sum_{i\in I} \delta_{\boldsymbol{\xi}_i}$, and denote by $r \geq 0$ and $\boldsymbol{\mu} \in \mathbb{R}^d$ the radius and the center of any (ideally the smallest) ball enclosing $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$, respectively. In this case, we have

$$C_l(\hat{\mathbb{P}}_n, m) = r \cdot C_l\left(\frac{1}{n} \sum_{i\in I} \delta_{\frac{\boldsymbol{\xi}_i - \boldsymbol{\mu}}{r}}, m \right) \leq r \cdot \overline{C}_l(n,m) \lesssim r \cdot \sqrt{1-p}, \tag{2.11}$$

where the inequality holds because $\|(\boldsymbol{\xi}_i - \boldsymbol{\mu})/r\|_2 \leq 1$ for every $i \in I$. Note that (2.11) enables us to find an upper bound on the smallest m guaranteeing that $C_l(\hat{\mathbb{P}}_n, m)$ falls below a prescribed threshold (*i.e.*, guaranteeing that the reduced m-point distribution remains within some prescribed distance from $\hat{\mathbb{P}}_n$).

Even though the inequality in (2.11) can be tight, which has been established in Proposition 2.4, one might suspect that typically $C_l(\hat{\mathbb{P}}_n, m)$ is significantly smaller than $r \cdot \sqrt{1-p}$ when the points $\boldsymbol{\xi}_i \in \mathbb{R}^d$, $i \in I$, are sampled randomly from a standard distribution, *e.g.*, a multivariate uniform or normal distribution. However, while the upper bound (2.11) can be loose for low-dimensional data, Proposition 2.7 below suggests that it is surprisingly tight in high dimensions—at least for $l = 2$.

Proposition 2.7. For any $\epsilon > 0$ and $\delta > 0$ there exist $c > 0$ and $d \in \mathbb{N}$ such that

$$\mathbb{P}^n \left(\|\boldsymbol{\xi}_i\|_2 \leq 1 \ \forall i \in I \text{ and } C_2\left(\frac{1}{n}\sum_{i\in I} \delta_{\boldsymbol{\xi}_i}, m\right) \geq \sqrt{1-p} - \delta \right) \geq 1 - \epsilon, \tag{2.12}$$

where $p = \frac{m}{n}$, and the support points $\boldsymbol{\xi}_i$, $i \in I$, are sampled independently from the normal distribution $\mathbb{P}$ with mean $\mathbf{0} \in \mathbb{R}^d$ and covariance matrix $(\sqrt{d-1} + c)^{-2}\mathbb{I} \in \mathbb{S}^d$.

Proposition 2.7 can be paraphrased as follows. Sampling the $\boldsymbol{\xi}_i$, $i \in I$, independently from the normal distribution $\mathbb{P}$ yields a (random) empirical distribution $\hat{\mathbb{P}}_n$ that is feasible and δ-suboptimal in (2.1) with probability $1 - \epsilon$. The intuition behind this result is that, in high dimensions, samples drawn from $\mathbb{P}$ are almost orthogonal and close to the surface of the unit ball with high probability. Indeed, these two properties are shared by the worst case distribution (2.8) in high dimensions.

Proof of Proposition 2.7. Theorem 2.1 and Remark 2.2 imply that

$$C_2^2\left(\frac{1}{n}\sum_{i\in I} \delta_{\boldsymbol{\xi}_i}, m\right) = \min_{\{I_j\}\in\mathfrak{P}(I,m)} \frac{1}{n}\sum_{j\in J}\sum_{i\in I_j} \left\|\boldsymbol{\xi}_i - \text{mean}(I_j)\right\|_2^2. \tag{2.13}$$

From the proof of Theorem 2.3 we further know that (2.13) can be expressed as a continuous function $f(\mathbf{S})$ of the Gram matrix $\mathbf{S} = [\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n]^\top [\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n]$, that is,

$$f(\mathbf{S}) = \min_{\{I_j\}\in\mathfrak{P}(I,m)} \frac{1}{n}\sum_{j\in J} \frac{1}{|I_j|^2} \sum_{i\in I_j} \left(|I_j|^2 s_{ii} - 2|I_j| \sum_{k\in I_j} s_{ik} + \sum_{k\in I_j} s_{kk} + \sum_{\substack{k,k'\in I_j \\ k\neq k'}} s_{kk'} \right).$$

An elementary calculation shows that $f(\mathbb{I}) = \frac{n-m}{n} = 1 - p$. Thus, by the continuity of $f(\cdot)$, there exists $\eta \in (0,1)$ with $\sqrt{f(\mathbf{S})} \geq \sqrt{1-p} - \delta$ whenever $\|\mathbf{S} - \mathbb{I}\|_{\max} \leq \eta$.

We are now ready to construct $\mathbb{P}$. First, select $c > 0$ large enough to ensure that

$$\left(1 - \frac{4}{c^2} e^{-\frac{c^2}{4}}\right)^n \geq 1 - \frac{\epsilon}{2}.$$

Then, select $d \in \mathbb{N}$ large enough such that

$$\frac{\sqrt{d-1}-c}{\sqrt{d-1}+c} \geq 1-\eta \quad \text{and} \quad n(n-1)\Phi(-\eta(\sqrt{d-1}+c)) \leq \frac{\epsilon}{2},$$

where $\Phi(\cdot)$ denotes the univariate standard normal distribution function. Observe that the distribution $\mathbb{P}$ is completely determined by c and d. Next, we find that

$$\begin{aligned}
&\mathbb{P}^n\left(\|\boldsymbol{\xi}_i\|_2 \leq 1 \ \forall i \text{ and } C_2\left(\frac{1}{n}\sum_{i\in I}\delta_{\boldsymbol{\xi}_i}, m\right) \geq \sqrt{1-p}-\delta\right) \\
&\geq \mathbb{P}^n\left(\|\boldsymbol{\xi}_i\|_2 \leq 1 \ \forall i \text{ and } \|\mathbf{S}-\mathbb{I}\|_{\max} \leq \eta\right) \\
&= \mathbb{P}^n\left(1-\eta \leq \|\boldsymbol{\xi}_i\|_2 \leq 1 \ \forall i \text{ and } |\boldsymbol{\xi}_i^\top \boldsymbol{\xi}_j| \leq \eta \ \forall i \neq j\right) \\
&\geq \mathbb{P}^n\left(\frac{\sqrt{d-1}-c}{\sqrt{d-1}+c} \leq \|\boldsymbol{\xi}_i\|_2 \leq 1 \ \forall i \text{ and } |\boldsymbol{\xi}_i^\top \boldsymbol{\xi}_j| \leq \eta \cdot \|\boldsymbol{\xi}_j\|_2 \ \forall i \neq j\right) \\
&\geq \mathbb{P}^n\left(\frac{\sqrt{d-1}-c}{\sqrt{d-1}+c} \leq \|\boldsymbol{\xi}_i\|_2 \leq 1 \ \forall i\right) + \mathbb{P}^n\left(|\boldsymbol{\xi}_i^\top \boldsymbol{\xi}_j| \leq \eta \cdot \|\boldsymbol{\xi}_j\|_2 \ \forall i \neq j\right) - 1, \qquad (2.14)
\end{aligned}$$

where the first and second inequalities follow from the construction of η and c, respectively, while the third inequality exploits the union bound. We also have

$$\begin{aligned}
\mathbb{P}^n\left(\frac{\sqrt{d-1}-c}{\sqrt{d-1}+c} \leq \|\boldsymbol{\xi}_i\|_2 \leq 1 \ \forall i\right) &= \mathbb{P}\left(\frac{\sqrt{d-1}-c}{\sqrt{d-1}+c} \leq \|\boldsymbol{\xi}_1\|_2 \leq 1\right)^n \\
&\geq \left(1-\frac{4}{c^2}e^{-\frac{c^2}{4}}\right)^n \geq 1-\frac{\epsilon}{2}, \qquad (2.15)
\end{aligned}$$

where the equality holds due to the independence of the $\boldsymbol{\xi}_i$, while the first and second inequalities follow from Lemma 2.8 in Hopcroft and Kannan (2012) and the construction of c, respectively. By the choice of d, we finally obtain

$$\begin{aligned}
\mathbb{P}^n\left(|\boldsymbol{\xi}_i^\top \boldsymbol{\xi}_j| \leq \eta \cdot \|\boldsymbol{\xi}_j\|_2 \ \forall i \neq j\right) &\geq 1 - \sum_{i\neq j} \mathbb{P}^n\left(|\boldsymbol{\xi}_i^\top \boldsymbol{\xi}_j| \geq \eta \cdot \|\boldsymbol{\xi}_j\|_2\right) \\
&= 1 - n(n-1)\Phi(-\eta(\sqrt{d-1}+c)) \geq 1-\frac{\epsilon}{2}. \qquad (2.16)
\end{aligned}$$

The equality in (2.16) holds due to the rotation symmetry of $\mathbb{P}$, which implies that

$$\mathbb{P}^n\left(|\boldsymbol{\xi}_i^\top \boldsymbol{\xi}_j| \geq \eta \cdot \|\boldsymbol{\xi}_j\|_2\right) = \mathbb{P}\left(|\boldsymbol{\xi}_1^\top \mathbf{e}_1| \geq \eta\right) = 2\Phi(-\eta(\sqrt{d-1}+c)).$$

The claim then follows by substituting (2.15) and (2.16) into (2.14). □

Figure 2.1 compares $\overline{C}_2(m,n)$ with $C_2(\hat{\mathbb{P}}_n, m)$ for $\hat{\mathbb{P}}_n$ uniform or normal, $n = 100$, $m \in \{10,\ldots,100\}$ and $d \in \{25,50,75,100\}$. The n original support points are sampled randomly from the uniform distribution on the unit ball (left panel) and the normal distribution from Proposition 2.7 with $c = 2.97$, which ensures that $\|\boldsymbol{\xi}_i\|_2 \leq 1$ with 95% probability (right panel). Note that

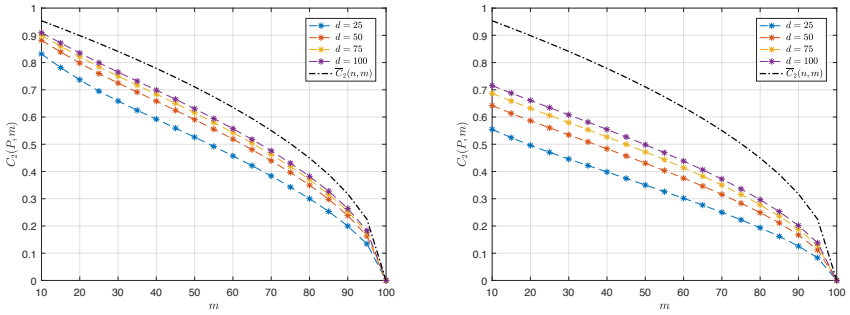


Figure 2.1 – Comparison between $C_2(\hat{\mathbb{P}}_n, m)$ and $\overline{C}_2(n,m)$ under uniform (left panel) and normal (right panel) sampling.

$C_2(\hat{\mathbb{P}}_n, m)$ is random. Thus, all shown values are averaged across 100 independent trials. Figure 2.1 confirms that $C_2(\hat{\mathbb{P}}_n, m)$ approaches the worst-case bound $\overline{C}_2(n,m)$ as the dimension d increases.

2.3 Guarantees for Discrete Scenario Reduction

For n-point empirical distributions $\hat{\mathbb{P}}_n = \frac{1}{n}\sum_{i=1}^n \delta_{\boldsymbol{\xi}_i}$ supported on $\mathbb{R}^d$, we now study the loss of optimality incurred by solving the discrete scenario reduction problem instead of its continuous counterpart. More precisely, we want to determine the point-wise largest lower bound $\underline{\kappa}_l(n,m)$ and the point-wise smallest upper bound $\overline{\kappa}_l(n,m)$ that satisfy

$$\underline{\kappa}_l(n,m)\cdot C_l(\hat{\mathbb{P}}_n, m) \;\le\; D_l(\hat{\mathbb{P}}_n, m) \;\le\; \overline{\kappa}_l(n,m)\cdot C_l(\hat{\mathbb{P}}_n, m) \quad \forall\, \hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n) \tag{2.17}$$

for the Wasserstein distances of type $l \in \{1,2\}$. Note that the existence of finite bounds $\underline{\kappa}_l(n,m)$ and $\overline{\kappa}_l(n,m)$ is not a priori obvious as they do not depend on the dimension d. Moreover, while it is clear that $\underline{\kappa}_l(n,m) \ge 1$ if it exists, it does not seem easy to derive a naïve upper bound on $\overline{\kappa}_l(n,m)$.

Our derivations in this section will use the following result, which is the analogue of Theorem 2.1 for the discrete scenario reduction problem.

Theorem 2.8. For any type-l Wasserstein distance induced by any norm $\|\cdot\|$, the discrete scenario reduction problem can be reformulated as

$$D_l(\hat{\mathbb{P}}_n, m) = \min_{\{I_j\}\in\mathfrak{P}(I,m)} \left[\frac{1}{n}\sum_{j\in J}\min_{\boldsymbol{\zeta}_j\in\{\boldsymbol{\xi}_i : i\in I_j\}}\sum_{i\in I_j}\|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j\|^l\right]^{1/l}.$$

Proof. The proof is similar to the proof of Theorem 2.1 and is therefore omitted. □

The remainder of this section derives lower and upper bounds on $\underline{\kappa}_l(n,m)$ and $\overline{\kappa}_l(n,m)$ for

Wasserstein distances of type $l = 2$ (Section 2.3.1) and $l = 1$ (Section 2.3.2), respectively. To eliminate trivial cases, we assume throughout this section that $n \geq 2$, $m \in \{1, \ldots, n-1\}$ and $d \geq 2$.

2.3.1 Guarantees for the Type-2 Wasserstein Distance

We first bound $\overline{\kappa}_2(n,m)$ in equation (2.17) from above (Theorem 2.9) and below (Proposition 2.10).

Theorem 2.9. The upper bound $\overline{\kappa}_2(n,m)$ in (2.17) satisfies $\overline{\kappa}_2(n,m) \leq \sqrt{2}$ for all n, m.

Proof. The proof proceeds in two steps. We first show that $\overline{\kappa}_2(n,m) \leq \sqrt{2}$ for all n when $m = 1$ (Step 1). Then we extend the result to all n and m (Step 2).

Step 1: Fix any $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$. W.l.o.g., we can assume that $\mathrm{mean}(I) = \mathbf{0}$ and $\frac{1}{n}\sum_{i\in I} \|\boldsymbol{\xi}_i\|_2^2 = 1$ by re-positioning and scaling the atoms $\boldsymbol{\xi}_i$ appropriately. Note that the re-positioning does not affect $C_2(\hat{\mathbb{P}}_n, 1)$ or $D_2(\hat{\mathbb{P}}_n, 1)$, and the positive homogeneity of the Wasserstein distance implies that the scaling affects both $C_2(\hat{\mathbb{P}}_n, 1)$ and $D_2(\hat{\mathbb{P}}_n, 1)$ in the same way and thus preserves their ratio $\overline{\kappa}_2(n, 1)$. Theorem 2.1 and Remark 2.2 then imply that

$$C_2(\hat{\mathbb{P}}_n, 1) \;=\; \left[\frac{1}{n}\sum_{i\in I} \|\boldsymbol{\xi}_i - \mathrm{mean}(I)\|_2^2\right]^{1/2} \;=\; 1.$$

Step 1 is thus complete if we can show that $D_2(\hat{\mathbb{P}}_n, 1) \leq \sqrt{2}$. Indeed, we have

$$\begin{aligned} D_2^2(\hat{\mathbb{P}}_n, 1) \;&=\; \min_{j\in I} \frac{1}{n}\sum_{i\in I} \|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\|_2^2 \;= \min_{j\in I} \frac{1}{n}\sum_{i\in I} (\boldsymbol{\xi}_i - \boldsymbol{\xi}_j)^\top (\boldsymbol{\xi}_i - \boldsymbol{\xi}_j) \\ &=\; \min_{j\in I} \frac{1}{n}\sum_{i\in I} \left(\boldsymbol{\xi}_i^\top \boldsymbol{\xi}_i - 2\boldsymbol{\xi}_i^\top \boldsymbol{\xi}_j + \boldsymbol{\xi}_j^\top \boldsymbol{\xi}_j\right) \;=\; \min_{j\in I} \frac{1}{n}\sum_{i\in I} \left(\boldsymbol{\xi}_i^\top \boldsymbol{\xi}_i + \boldsymbol{\xi}_j^\top \boldsymbol{\xi}_j\right) \\ &=\; \min_{j\in I} \|\boldsymbol{\xi}_j\|_2^2 + \frac{1}{n}\sum_{i\in I} \|\boldsymbol{\xi}_i\|_2^2 \;=\; 1 + \min_{j\in I} \|\boldsymbol{\xi}_j\|_2^2 \;\leq\; 2, \end{aligned} \tag{2.18}$$

where the first equality is due to Theorem 2.8, the fourth follows from $\sum_{i\in I} \boldsymbol{\xi}_i = n \cdot \mathrm{mean}(I) = \mathbf{0}$, and the inequality holds since $\min_{j\in I} \|\boldsymbol{\xi}_j\|_2^2 \leq \frac{1}{n}\sum_{i\in I} \|\boldsymbol{\xi}_i\|_2^2 = 1$.

Step 2: Fix any $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$. Theorem 2.1 and Remark 2.2 imply that

$$C_2(\hat{\mathbb{P}}_n, m) = \min_{\{I_j\}\in\mathfrak{P}(I,m)} \left[\frac{1}{n}\sum_{j\in J}\sum_{i\in I_j} \|\boldsymbol{\xi}_i - \mathrm{mean}(I_j)\|_2^2\right]^{1/2}.$$

For an optimal partition $\{I_j^\star\}$ to this problem, $C_2(\hat{\mathbb{P}}_n, m)$ can be expressed as

$$C_2(\hat{\mathbb{P}}_n, m) = \left[\sum_{j\in J} \frac{|I_j^\star|}{n} C_{2,j}^2\right]^{1/2} \text{ with } C_{2,j} = \left[\frac{1}{|I_j^\star|}\sum_{i\in I_j^\star} \|\boldsymbol{\xi}_i - \text{mean}(I_j^\star)\|_2^2\right]^{1/2}.$$

From our discussion in Step 1 we know that $C_{2,j}$ represents the type-2 Wasserstein distance between the conditional empirical distribution $\hat{\mathbb{P}}_n^j = \frac{1}{|I_j^\star|}\sum_{i\in I_j^\star} \delta_{\boldsymbol{\xi}_i}$ and its closest Dirac distribution, that is, $C_2(\hat{\mathbb{P}}_n^j, 1)$. Analogously, we obtain that

$$\begin{aligned} D_2(\hat{\mathbb{P}}_n, m) &\leq \left[\sum_{j\in J} \frac{|I_j^\star|}{n} D_{2,j}^2\right]^{1/2} \text{ with } D_{2,j} = \left[\min_{j\in I_j^\star} \frac{1}{|I_j^\star|}\sum_{i\in I_j^\star} \|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\|_2^2\right]^{1/2} \\ &\leq \left[\sum_{j\in J} \frac{|I_j^\star|}{n} (2C_{2,j}^2)\right]^{1/2} = \sqrt{2}\, C_2(\hat{\mathbb{P}}_n, m), \end{aligned}$$

where the first inequality holds since the optimal partition $\{I_j^\star\}$ for $C_2(\hat{\mathbb{P}}_n, m)$ is typically suboptimal in $D_2(\hat{\mathbb{P}}_n, m)$, the second inequality follows from the fact that $D_{2,j} = D_2(\hat{\mathbb{P}}_n^j, 1)$ and $D_2(\hat{\mathbb{P}}_n^j, 1) \leq \sqrt{2} C_2(\hat{\mathbb{P}}_n^j, 1)$ due to Step 1, and the identity follows from the definition of $C_{2,j}$. The statement now follows. □

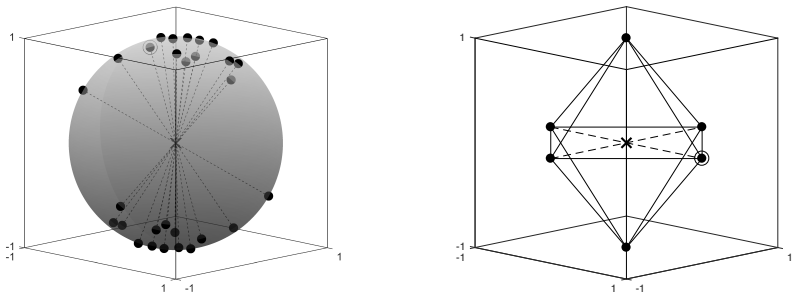


Figure 2.2 – Empirical distributions in $\mathbb{R}^3$ that maximize the ratio between $D_l(\hat{\mathbb{P}}_n, 1)$ and $C_l(\hat{\mathbb{P}}_n, 1)$ for $l = 2$ and $\|\cdot\| = \|\cdot\|_2$ (left panel) as well as $l = 1$ and $\|\cdot\| = \|\cdot\|_1$ (right panel). In both cases, the continuous scenario reduction problem is optimized by the Dirac distribution at $\mathbf{0}$ (marked as ×), whereas the discrete scenario reduction problem is optimized by any of the atoms (such as ∘).

Proposition 2.10. There is $\hat{\mathbb{P}}_n \in \mathscr{P}_{\text{E}}(\mathbb{R}^d, n)$ with $D_2(\hat{\mathbb{P}}_n, m) = \sqrt{2} C_2(\hat{\mathbb{P}}_n, m)$ for all n, m.

Proof. In analogy to the proof of Theorem 2.9, we first show the statement for $m = 1$ (Step 1) and then extend the result to $m > 1$ (Step 2).

Step 1: The first step in the proof of Theorem 2.9 shows that $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ satisfies $D_2(\hat{\mathbb{P}}_n, 1) = \sqrt{2}C_2(\hat{\mathbb{P}}_n, 1)$ if $\sum_{i\in I} \boldsymbol{\xi}_i = \mathbf{0}$ and $\|\boldsymbol{\xi}_1\| = \ldots = \|\boldsymbol{\xi}_n\| = 1$. For an even number $n = 2k$, $k \in \mathbb{N}$, both conditions are satisfied if we place $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_k$ on the surface of the unit ball in $\mathbb{R}^d$ and then choose $\boldsymbol{\xi}_{k+i} = -\boldsymbol{\xi}_i$ for $i = 1, \ldots, k$ (see left panel of Figure 2.2 for an illustration in $\mathbb{R}^3$). Likewise, for an odd number $n = 2k+3$, $k \in \mathbb{N}_0$, we can place $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_k$ on the surface of the unit ball, choose $\boldsymbol{\xi}_{k+i} = -\boldsymbol{\xi}_i$ for $i = 1, \ldots, k$ and fix $\boldsymbol{\xi}_{2k+1} = \mathbf{e}_1$, $\boldsymbol{\xi}_{2k+2} = -\frac{1}{2}\mathbf{e}_1 + \frac{\sqrt{3}}{2}\mathbf{e}_2$ and $\boldsymbol{\xi}_{2k+3} = -\frac{1}{2}\mathbf{e}_1 - \frac{\sqrt{3}}{2}\mathbf{e}_2$.

Step 2: To prove the statement for $m > 1$, we construct an empirical distribution $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ whose atoms satisfy $\mathrm{supp}(\hat{\mathbb{P}}_n) = \Xi_1 \cup \Xi_2$ with $|\Xi_1| = n - m + 1$ and $|\Xi_2| = m - 1$. The atoms $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{n-m+1}$ in Ξ_1 are selected according to the recipe outlined in Step 1, whereas the atoms $\boldsymbol{\xi}_{n-m+2}, \ldots, \boldsymbol{\xi}_n$ in Ξ_2 satisfy $\boldsymbol{\xi}_{n-m+1+i} = (1 + iM)\mathbf{e}_1$, $i = 1, \ldots, m-1$, for any number M satisfying $M > 2\sqrt{n-m+1}$. A direct calculation then shows that the atoms in Ξ_2 are sufficiently far away from those in Ξ_1 as well as from each other so that any optimal partition $\{I_j^\star\}$ to the discrete scenario reduction problem in Theorem 2.8 as well as the continuous scenario reduction problem in Theorem 2.1 consists of the sets $\{i : \boldsymbol{\xi}_i \in \Xi_1\}$ and $\{i\}$, $\boldsymbol{\xi}_i \in \Xi_2$. The result then follows from the fact that either problem accumulates a Wasserstein distance of 0 over the atoms in Ξ_2, whereas the Wasserstein distance of $D_2(\hat{\mathbb{P}}_n, m)$ is a factor of $\sqrt{2}$ bigger than the Wasserstein distance of $C_2(\hat{\mathbb{P}}_n, m)$ over the atoms in Ξ_1 (see Step 1). □

Theorem 2.9 and Proposition 2.10 imply that $\overline{\kappa}_2(n, m) = \sqrt{2}$ for all n and m, that is, the bound is indeed *independent* of both the number of atoms n in the empirical distribution and the number of atoms m in the reduced distribution. We now show that the naïve lower bound of 1 on the approximation ratio is essentially tight.

Proposition 2.11. The lower bound $\underline{\kappa}_2(n, m)$ in (2.17) satisfies $\underline{\kappa}_2(n, m) = 1$ whenever $n \geq 3$ and $m \in \{1, \ldots, n-2\}$, while $\underline{\kappa}_2(n, n-1) = \sqrt{2}$ always.

Proof. We first prove $\underline{\kappa}_2(n, m) = 1$ when $m = 1$ and $n \geq 3$ (Step 1) and when $m \in \{2, \ldots, n-2\}$ (Step 2). Then, we show $\underline{\kappa}_2(n, n-1) = \sqrt{2}$ (Step 3).

Step 1: Choose $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ such that the first $n-1$ atoms $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{n-1}$ are selected according to the recipe outlined in Step 1 in the proof of Proposition 2.10 and $\boldsymbol{\xi}_n = \mathbf{0}$. We thus have $\mathrm{mean}(I) = \mathbf{0}$, and Theorem 2.1 and Remark 2.2 imply that the optimal continuous scenario reduction is given by the Dirac distribution $\delta_{\mathbf{0}}$. Since $\mathbf{0} \in \mathrm{supp}(\hat{\mathbb{P}}_n)$, we have $C_2(\hat{\mathbb{P}}_n, 1) = D_2(\hat{\mathbb{P}}_n, 1)$ and the result follows.

Step 2: To prove the statement for $m > 1$, we proceed as in Step 2 in the proof of Proposition 2.10. In particular, we construct an empirical distribution $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ whose atoms satisfy $\mathrm{supp}(\hat{\mathbb{P}}_n) = \Xi_1 \cup \Xi_2$ with $|\Xi_1| = n - m + 1$ and $|\Xi_2| = m - 1$. The atoms $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{m-m+1}$

in Ξ_1 are selected according to the recipe outlined in Step 1 of this proof, whereas the remaining atoms $\boldsymbol{\xi}_{n-m+1},\ldots,\boldsymbol{\xi}_n$ in Ξ_2 satisfy $\boldsymbol{\xi}_{n-m+1+i} = (1+iM)\mathbf{e}_1$, $i = 1,\ldots,m-1$, for any $M > 2\sqrt{n-m+1}$. A similar argument as in the proof of Proposition 2.10 then shows that $C_2(\hat{\mathbb{P}}_n, m) = D_2(\hat{\mathbb{P}}_n, m)$.

Step 3: Fix any $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$. W.l.o.g., assume that $\{\boldsymbol{\xi}_{n-1}, \boldsymbol{\xi}_n\}$ is the closest pair of atoms in terms of Euclidean distance, and let $d_{\min} = \|\boldsymbol{\xi}_n - \boldsymbol{\xi}_{n-1}\|_2$. One readily verifies that the partition $I_j^\star = \{j\}$, $j = 1,\ldots,n-2$, and $I_{n-1}^\star = \{n-1, n\}$ optimizes both the discrete scenario reduction problem in Theorem 2.8 as well as the continuous scenario reduction problem in Theorem 2.1. We thus have $C_2(\hat{\mathbb{P}}_n, n-1) = \frac{1}{\sqrt{2n}} d_{\min}$ and $D_2(\hat{\mathbb{P}}_n, n-1) = \frac{1}{\sqrt{n}} d_{\min}$, which concludes the proof. □

Hence, for any empirical distribution $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ the type-2 Wasserstein distance between the minimizer of the *discrete* scenario reduction problem and $\hat{\mathbb{P}}_n$ exceeds the Wasserstein distance between the minimizer of the *continuous* scenario reduction problem and $\hat{\mathbb{P}}_n$ by up to 41.4%, and the bound is attainable for any n, m.

2.3.2 Guarantees for the Type-1 Wasserstein Distance

In analogy to Section 2.3.1, we first bound $\overline{\kappa}_1(n,m)$ from above (Theorem 2.12) and below (Proposition 2.13). In contrast to the previous section, we consider an arbitrary norm $\|\cdot\|$, and we adapt the definition of the geometric median accordingly.

Theorem 2.12. The upper bound $\overline{\kappa}_1(n,m)$ in (2.17) satisfies $\overline{\kappa}_1(n,m) \leq 2$ whenever $m \in \{2,\ldots,n-2\}$ as well as $\overline{\kappa}_1(n,1) \leq 2\left(1-\frac{1}{n}\right)$ and $\overline{\kappa}_1(n,n-1) \leq 1$.

Proof. We first prove the statement for $m = 1$ (Step 1) and then extend the result to $m \in \{2,\ldots,n-2\}$ (Step 2) and $m = n-1$ (Step 3).

Step 1: Fix any $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$. As in the proof of Theorem 2.9, we can assume that $\mathrm{gmed}(I) = \mathbf{0}$ and $\frac{1}{n}\sum_{i\in I}\|\boldsymbol{\xi}_i\| = 1$ by re-positioning and scaling the atoms $\boldsymbol{\xi}_i$ appropriately. Theorem 2.1 and Remark 2.2 then imply that for $m = 1$, we have

$$C_1(\hat{\mathbb{P}}_n, 1) \;=\; \frac{1}{n}\sum_{i\in I}\|\boldsymbol{\xi}_i - \mathrm{gmed}(I)\| = 1.$$

Step 1 is thus complete if we can show that $D_1(\hat{\mathbb{P}}_n, 1) \leq 2\left(1-\frac{1}{n}\right)$. Indeed, we have

$$\begin{aligned} D_1(\hat{\mathbb{P}}_n, 1) &= \min_{j\in I} \frac{1}{n}\sum_{i\in I} \|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\| \;=\; \min_{j\in I} \frac{1}{n}\sum_{i\in I\setminus\{j\}} \|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\| \\ &\leq \min_{j\in I} \frac{1}{n}\sum_{i\in I\setminus\{j\}} \big(\|\boldsymbol{\xi}_i\| + \|\boldsymbol{\xi}_j\|\big) \;=\; \min_{j\in I} \frac{1}{n}\left((n-2)\|\boldsymbol{\xi}_j\| + \sum_{i\in I}\|\boldsymbol{\xi}_i\|\right) \\ &= \min_{j\in I} \frac{1}{n}\big((n-2)\|\boldsymbol{\xi}_j\| + n\big) \;=\; 1 + \frac{n-2}{n}\cdot \min_{j\in I}\|\boldsymbol{\xi}_j\| \;\leq\; 2\left(1-\frac{1}{n}\right), \end{aligned}$$

where the two inequalities follow from the triangle inequality and the fact that $\min_{j\in I}\|\boldsymbol{\xi}_j\| \leq \frac{1}{n}\sum_{i\in I}\|\boldsymbol{\xi}_i\| = 1$, respectively.

Step 2: Fix any $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$. Theorem 2.1 and Remark 2.2 then imply that

$$C_1(\hat{\mathbb{P}}_n, m) = \min_{\{I_j\}\in\mathfrak{P}(I,m)} \frac{1}{n}\sum_{j\in J}\sum_{i\in I_j} \|\boldsymbol{\xi}_i - \mathrm{gmed}(I_j)\|.$$

Let $\{I_j^\star\}$ be an optimal partition for this problem. The same arguments as in the proof of Theorem 2.9 show that

$$\begin{aligned} D_1(\hat{\mathbb{P}}_n, m) &\leq \sum_{j\in J} \frac{|I_j^\star|}{n} D_{1,j} \qquad \text{with } D_{1,j} = \min_{j\in I_j^\star} \frac{1}{|I_j^\star|}\sum_{i\in I_j^\star} \|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\| \\ &\leq \sum_{j\in J} \frac{|I_j^\star|}{n} (2\,C_{1,j}) \quad \text{with } C_{1,j} = \frac{1}{|I_j^\star|}\sum_{i\in I_j^\star} \|\boldsymbol{\xi}_i - \mathrm{gmed}(I_j^\star)\|, \end{aligned}$$

and the last expression is equal to $2\,C_1(\hat{\mathbb{P}}_n, m)$ by definition of $C_{1,j}$.

Step 3: For $n = 2$ and $m = n-1 = 1$, Step 1 shows that $\overline{\kappa}_1(2,1) \leq 2\left(1-\frac{1}{2}\right) = 1$. For $n > 2$ and $m = n-1$, the statement can be derived in the same way as the third step in the proof of Proposition 2.11. We omit the details for the sake of brevity. □

Proposition 2.13. There is $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ such that $D_1(\hat{\mathbb{P}}_n, m) = 2\left(1-\frac{m}{n}\right) C_1(\hat{\mathbb{P}}_n, m)$ under the 1-norm for all n divisible by $2m$, all m and all $d \geq \frac{n}{2m}$.

Proof. We first prove the statement for $m = 1$ (Step 1) and then extend the result to $m > 1$ (Step 2). Throughout the proof, we set $k = \frac{n}{2m}$ and consider w.l.o.g. the case where $d = k$.

Step 1: Fix $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ with the atoms $\boldsymbol{\xi}_i = +\mathbf{e}_i$ as well as $\boldsymbol{\xi}_{k+i} = -\mathbf{e}_i$, $i = 1, \ldots, k$. The symmetric placement of the atoms implies that $\mathrm{gmed}(I) = \mathbf{0}$ and hence $C_1(\hat{\mathbb{P}}_n, 1) = 1$. Furthermore, we note that $\|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\|_1 = 2$ for all $i \neq j$, that is, any two atoms are equidistant from another

(see right panel of Figure 2.2 for an illustration in $\mathbb{R}^3$). By Theorem 2.8, any 1-point discrete scenario reduction results in a Wasserstein distance of $2\,\frac{n-1}{n}$ to $\hat{\mathbb{P}}_n$.

Step 2: To prove the statement for $m > 1$, we construct an empirical distribution $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ whose atoms satisfy $\mathrm{supp}(\hat{\mathbb{P}}_n) = \bigcup_{j=1}^m (\Xi_j^+ \cup \Xi_j^-)$ with $|\Xi_j^+| = |\Xi_j^-| = k$, $j = 1,\ldots,m$. The atoms $\boldsymbol{\xi}_{2(j-1)k+1},\ldots,\boldsymbol{\xi}_{2(j-1)k+k}$ in Ξ_j^+ satisfy $\boldsymbol{\xi}_{2(j-1)k+i} = +\mathbf{e}_i + jM\mathbf{e}_1$, $i = 1,\ldots,k$, whereas the atoms $\boldsymbol{\xi}_{2(j-1)k+k+1},\ldots,\boldsymbol{\xi}_{2jk}$ in Ξ_j^- satisfy $\boldsymbol{\xi}_{2(j-1)k+k+i} = -\mathbf{e}_i + jM\mathbf{e}_1$, $i = 1,\ldots,k$, for any number M satisfying $M > 2n+2$. The same arguments as in the proof of Theorem 2.9 show that any optimal partition $\{I_j^\star\}$ to the discrete scenario reduction problem in Theorem 2.8 as well as the continuous scenario reduction problem in Theorem 2.1 consists of the sets indexing the atoms in $\Xi_j^+ \cup \Xi_j^-$, $j = 1,\ldots,m$. Step 1 shows that the continuous scenario reduction problem accumulates a Wasserstein distance of 1 over each set, whereas the discrete scenario reduction problem accumulates a Wasserstein distance of $2\,\frac{2k-1}{2k}$ over each set. The result then follows from the fact that there are m such sets and hence the ratio of the respective overall Wasserstein distances amounts to $m\big(2\frac{2k-1}{2k}\big)/m = 2\big(1-\frac{m}{n}\big)$. □

Theorem 2.12 and Proposition 2.13 imply that $\overline{\kappa}_1(n,m) \in [2(1-m/n),2]$ for all n and $m \in \{2,\ldots,n-2\}$. For the small ratios $m:n$ commonly used in practice, we thus conclude that the bound is *essentially independent* of both the number of atoms n in the empirical distribution and the number of atoms m in the reduced distribution. We close with an analysis of the lower bound $\underline{\kappa}_1(n,m)$.

Proposition 2.14. The lower bound $\underline{\kappa}_1(n,m)$ in (2.17) satisfies $\underline{\kappa}_1(n,m) = 1$ for all n,m.

Proof. The proof widely parallels that of Proposition 2.11, with the difference that the atoms $\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_n$ of the empirical distribution $\hat{\mathbb{P}}_n$ are placed such that a geometric median (as opposed to the mean) of each subset in the optimal partition coincides with one of the atoms in that subset. This allows both continuous and discrete scenario reduction to choose the same support points for the reduced distribution, hence incurring the same Wasserstein distance. Details are omitted for brevity. □

In conclusion, for any $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ the type-1 Wasserstein distance between the minimizer of the *discrete* scenario reduction problem and $\hat{\mathbb{P}}_n$ exceeds the Wasserstein distance between the minimizer of the *continuous* scenario reduction problem and $\hat{\mathbb{P}}_n$ by up to 100%, and this bound is asymptotically attained for decreasing ratios $m:n$.

2.4 Solution Methods for Scenario Reduction Problems

We now review existing solution methods and propose new solution schemes for the discrete and continuous scenario reduction problems. More precisely, we will study two heuristics for discrete and continuous scenario reduction, respectively, that do not come with approximation

guarantees (Section 2.4.1), we will propose a constant-factor approximation scheme for both the discrete and the continuous scenario reduction problem (Section 2.4.2), and we will discuss two exact reformulations of these problems as mixed-integer optimization problems (Section 2.4.3).

In the remainder of this section, we denote by $D_l(\hat{\mathbb{P}}_n, \Xi)$ the type-l Wasserstein distance between $\hat{\mathbb{P}}_n$ and its closest distribution supported on the finite set Ξ. Moreover, for an algorithm providing an upper bound $\overline{D}_l(\hat{\mathbb{P}}_n, m)$ on the discrete scenario reduction problem in $\mathbb{R}^d$, we define the algorithm's *approximation ratio* as the maximum fraction $\overline{D}_l(\hat{\mathbb{P}}_n, m)/D_l(\hat{\mathbb{P}}_n, m)$, where the maximum is taken over all n and m, as well as all empirical distributions $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$.

2.4.1 Heuristics for the Discrete Scenario Reduction Problem

We review in Section 2.4.1 a popular heuristic for the discrete scenario reduction problem due to Dupačová et al. (2003). We will show that despite the simplicity and efficiency of the algorithm, there is no finite upper bound on the algorithm's approximation ratio. Note that this analysis carries over to the fast forward heuristic due to Heitsch and Römisch (2003), which generates the same reduced distribution at a lower computational cost. In Section 2.4.1 we adapt a widely used clustering heuristic to the continuous scenario reduction problem, and we show that this algorithm's approximation ratio cannot be bounded from above either.

Dupačová et al.'s Algorithm

We outline Dupačová et al.'s algorithm for the problem $D_l(\hat{\mathbb{P}}_n, m)$ below.

DUPAČOVÁ ET AL.'S ALGORITHM FOR $D_l(\hat{\mathbb{P}}_n, m)$:

1. Initialize the set of atoms in the reduced set as $R \leftarrow \emptyset$.
2. Select the next atom to be added to the reduced set as
$$\boldsymbol{\zeta} \in \operatorname*{arg\,min}_{\boldsymbol{\zeta} \in \mathrm{supp}(\hat{\mathbb{P}}_n)} D_l(\hat{\mathbb{P}}_n, R \cup \{\boldsymbol{\zeta}\})$$
and update $R \leftarrow R \cup \{\boldsymbol{\zeta}\}$.
3. Repeat Step 2 until $|R| = m$.

Given an empirical distribution $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$, the algorithm iteratively populates the reduced set R containing the atoms of the reduced distribution $\mathbb{Q}$. Each atom $\boldsymbol{\zeta} \in \mathrm{supp}(\hat{\mathbb{P}}_n)$ is selected greedily so as to minimize the Wasserstein distance between $\hat{\mathbb{P}}_n$ and the closest distribution supported on the augmented reduced set $R \cup \{\boldsymbol{\zeta}\}$. After termination, the distribution $\mathbb{Q}$ can be recovered from the reduced set R as follows. Let $\{I_{\boldsymbol{\zeta}}\} \in \mathfrak{P}(I, m)$ be any partition of $\mathrm{supp}(\hat{\mathbb{P}}_n)$

into sets $I_{\boldsymbol{\zeta}}$, $\boldsymbol{\zeta} \in R$, such that $I_{\boldsymbol{\zeta}}$ contains all elements of $\mathrm{supp}(\hat{\mathbb{P}}_n)$ that are closest to $\boldsymbol{\zeta}$ (ties may be broken arbitrarily). Then $\mathbb{Q} = \sum_{\boldsymbol{\zeta} \in R} q_{\boldsymbol{\zeta}} \delta_{\boldsymbol{\zeta}}$, where $q_{\boldsymbol{\zeta}} = |I_{\boldsymbol{\zeta}}|/n$.

Theorem 2.15. For every $d \geq 2$ and $l, p \geq 1$, the approximation ratio of Dupačová et al.'s algorithm is unbounded.

Proof. The proof constructs a specific distribution $\hat{\mathbb{P}}_n$ (Step 1), bounds $D_l(\hat{\mathbb{P}}_n, m)$ from above (Step 2) and bounds the Wasserstein distance between $\hat{\mathbb{P}}_n$ and the output $\mathbb{Q}$ of Dupačová et al.'s algorithm from below (Step 3).

Step 1: Fix $d \geq 2$, $l, p \geq 1$ and $m = 4$, and consider the empirical distribution $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$ with $n = 4z+1$ for some positive integer z as well as $\mathrm{supp}(\hat{\mathbb{P}}_n) = \Xi_1 \cup \cdots \cup \Xi_4 \cup \{\boldsymbol{\xi}_{4z+1}\}$ with $\Xi_j = \{\boldsymbol{\xi}_{(j-1)z+1}, \ldots, \boldsymbol{\xi}_{jz}\}$, $j = 1, \ldots, 4$, and

$$\Xi_1 \subset \mathscr{B}_\epsilon(+\mathbf{e}_1), \quad \Xi_2 \subset \mathscr{B}_\epsilon(-\mathbf{e}_1), \quad \Xi_3 \subset \mathscr{B}_\epsilon(+\mathbf{e}_2), \quad \Xi_4 \subset \mathscr{B}_\epsilon(-\mathbf{e}_2)$$

and $\boldsymbol{\xi}_{4z+1} = \mathbf{0}$, where $\mathscr{B}_\epsilon(\boldsymbol{x}) = \{\boldsymbol{\xi} \in \mathbb{R}^d : \|\boldsymbol{\xi} - \boldsymbol{x}\|_p \leq \epsilon\}$ denotes the ϵ-ball around $\boldsymbol{x}$. Here, $\epsilon > 0$ is small enough so that each atom in Ξ_i is closer to $\mathbf{0}$ than to any atom in any of the other sets Ξ_j. The triangle inequality then implies that

$$\begin{aligned} &\|\boldsymbol{\xi}_i\|_p \in [1-\epsilon, 1+\epsilon] && \forall \boldsymbol{\xi}_i \in \Xi_1, \\ &\|\boldsymbol{\xi}_i - \boldsymbol{\xi}_1\|_p \geq 2-2\epsilon && \forall \boldsymbol{\xi}_i \in \Xi_2, \\ &\|\boldsymbol{\xi}_i - \boldsymbol{\xi}_1\|_p \geq 1-\epsilon && \forall \boldsymbol{\xi}_i \in \Xi_3 \cup \Xi_4. \end{aligned} \tag{2.19}$$

Step 2: By construction, we have that

$$\begin{aligned} D_l(\hat{\mathbb{P}}_n, 4) &\leq d_l\left(\hat{\mathbb{P}}_n, \frac{z+1}{4z+1}\delta_{\boldsymbol{\xi}_z} + \frac{z}{4z+1}\delta_{\boldsymbol{\xi}_{2z}} + \frac{z}{4z+1}\delta_{\boldsymbol{\xi}_{3z}} + \frac{z}{4z+1}\delta_{\boldsymbol{\xi}_{4z}}\right) \\ &\leq \left[\frac{1}{4z+1}\left(\left[\sum_{j=1}^{4}\sum_{\boldsymbol{\xi}_i \in \Xi_j} \|\boldsymbol{\xi}_i - \boldsymbol{\xi}_{jz}\|_p^l\right] + \|\boldsymbol{\xi}_{4z+1} - \boldsymbol{\xi}_z\|_p^l\right)\right]^{1/l} \\ &\leq \left[\frac{1}{4z+1}\left(\left[\sum_{j=1}^{4}\sum_{\boldsymbol{\xi}_i \in \Xi_j} (2\epsilon)^l\right] + (1+\epsilon)^l\right)\right]^{1/l} = \left[\frac{4z2^l\epsilon^l + (1+\epsilon)^l}{4z+1}\right]^{1/l}, \end{aligned}$$

where the first inequality holds because $\boldsymbol{\xi}_z, \boldsymbol{\xi}_{2z}, \boldsymbol{\xi}_{3z}, \boldsymbol{\xi}_{4z} \in \mathrm{supp}(\hat{\mathbb{P}}_n)$, the second inequality holds since moving the atoms in Ξ_j to $\boldsymbol{\xi}_{jz}$, $j = 1, \ldots, 4$, and $\boldsymbol{\xi}_{4z+1}$ to $\boldsymbol{\xi}_z$ represents a feasible transportation plan, and the third inequality is due to (2.19) and the triangle inequality.

Step 3: We first show that for a sufficiently small $\epsilon > 0$, Dupačová et al.'s algorithm adds $\boldsymbol{\xi}_{4z+1} = \mathbf{0}$ to the reduced set R in the first iteration. We then show that under this selection, the

output $\mathbb{Q}$ of Dupačová et al.'s algorithm can be arbitrarily worse than the bound on $D_l(\hat{\mathbb{P}}_n, 4)$ determined in the previous step.

To show the first point, the symmetry inherent in $\mathrm{supp}(\hat{\mathbb{P}}_n)$ implies that it suffices to show that $d_l(\hat{\mathbb{P}}_n, \delta_{\mathbf{0}}) < d_l(\hat{\mathbb{P}}_n, \delta_{\boldsymbol{\xi}_1})$. To this end, we note that

$$d_l^l(\hat{\mathbb{P}}_n, \delta_{\mathbf{0}}) = \frac{1}{4z+1} \sum_{j=1}^{4} \sum_{\boldsymbol{\xi}_i \in \Xi_j} \|\boldsymbol{\xi}_i\|_p^l \leq \frac{1}{4z+1} \sum_{j=1}^{4} \sum_{\boldsymbol{\xi}_i \in \Xi_j} (1+\epsilon)^l = \frac{4z}{4z+1}(1+\epsilon)^l$$

due to equation (2.19), while at the same time

$$\begin{aligned} d_l^l(\hat{\mathbb{P}}_n, \delta_{\boldsymbol{\xi}_1}) &= \frac{1}{4z+1} \sum_{i=2}^{4z+1} \|\boldsymbol{\xi}_i - \boldsymbol{\xi}_1\|_p^l \\ &\geq \frac{1}{4z+1} \sum_{i=z+1}^{4z+1} \|\boldsymbol{\xi}_i - \boldsymbol{\xi}_1\|_p^l \geq \frac{z(2-2\epsilon)^l + (2z+1)(1-\epsilon)^l}{4z+1}. \end{aligned}$$

As ϵ tends to 0, we have that

$$\lim_{\epsilon \to 0} d_l(\hat{\mathbb{P}}_n, \delta_{\mathbf{0}}) \leq \left[\frac{4z}{4z+1}\right]^{1/l} < \left[\frac{z2^l + 2z + 1}{4z+1}\right]^{1/l} \leq \lim_{\epsilon \to 0} d_l(\hat{\mathbb{P}}_n, \delta_{\boldsymbol{\xi}_1}),$$

where the strict inequality is due to $l \geq 1$. As a consequence, we may conclude that there indeed exists an $\epsilon > 0$ such that Dupačová et al.'s algorithm adds $\boldsymbol{\xi}_{4z+1} = \mathbf{0}$ to the reduced set R in the first iteration.

As for the second point, we note that after adding $\boldsymbol{\xi}_{4z+1} = \mathbf{0}$ to the reduced set R, there must be at least one subset Ξ_j, $j \in \{1, \ldots, 4\}$, such that no $\boldsymbol{\xi}_i \in \Xi_j$ is contained in the final reduced set R. Assume w.l.o.g. that this is the case for $j = 1$. We then have

$$d_l(\hat{\mathbb{P}}_n, \mathbb{Q}) \geq \left[\frac{1}{4z+1} \sum_{\boldsymbol{\xi}_i \in \Xi_1} \|\boldsymbol{\xi}_j - \mathbf{0}\|_p\right]^{1/l} \geq \left[\frac{z(1-\epsilon)}{4z+1}\right]^{1/l},$$

and combining this with the result of Step 2, we can conclude that the approximation ratio $d_l(\hat{\mathbb{P}}_n, \mathbb{Q}) / D_l(\hat{\mathbb{P}}_n, 4)$ approaches ∞ as $z \to \infty$ and $z\epsilon^l \to 0$. □

We remark that the algorithm of Dupačová et al. can be improved by adding multiple atoms to the reduced set R in Step 2. Nevertheless, a similar argument as in the proof of Theorem 2.15 shows that the resulting improved algorithm does not allow for a finite upper bound on the approximation ratio either.

k-Means Clustering Algorithm

The k-means clustering algorithm has first been proposed in 1957 for a pulse-code modulation problem (Lloyd 1982), and it has since then become a mainstay clustering approach in many

application domains, for instance market segmentation, recommender systems, and image segmentation as well as feature learning (Jain 2010a). It aims to partition a set of observations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ into m clusters $S_1, \ldots, S_m$ such that the intra-cluster sums of squared distances are minimized. By generalizing the algorithm to arbitrary powers and norms, we can adapt the algorithm to our continuous scenario reduction problem as follows.

k-MEANS CLUSTERING ALGORITHM FOR $C_l(\hat{\mathbb{P}}_n, m)$:

1. Initialize the reduced set $R = \{\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_m\} \subseteq \operatorname{supp}(\hat{\mathbb{P}}_n)$ arbitrarily.
2. Let $\{I_j\} \in \mathfrak{P}(I, m)$ be any partition whose sets I_j, $j \in J$, contain all atoms of $\operatorname{supp}(\hat{\mathbb{P}}_n)$ that are closest to $\boldsymbol{\zeta}_j$ (ties may be broken arbitrarily).
3. For each $j \in J$, update $\boldsymbol{\zeta}_j \leftarrow \arg\min\{\sum_{i \in I_j} \|\boldsymbol{\xi}_i - \boldsymbol{\zeta}\|^l : \boldsymbol{\zeta} \in \mathbb{R}^d\}$.
4. Repeat Steps 2 and 3 until the reduced set R no longer changes.

For the empirical distribution $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$, the algorithm iteratively updates the reduced set R containing the atoms of the reduced distribution $\mathbb{Q}$ through a sequence of assignment (Step 2) and update (Step 3) steps. Step 2 assigns each atom $\boldsymbol{\xi}_i \in \operatorname{supp}(\hat{\mathbb{P}}_n)$ of the empirical distribution to the closest atom in the reduced set, and Step 3 updates each atom in the reduced set so as to minimize the sum of l-th powers of the distances to its assigned atoms from $\operatorname{supp}(\hat{\mathbb{P}}_n)$. After termination, the continuously reduced distribution $\mathbb{Q}$ can be recovered from the reduced set R in the same way as in the previous subsection.

Remark 2.2 implies that for $l = 2$ and $\|\cdot\| = \|\cdot\|_2$, Step 3 reduces to $\boldsymbol{\zeta}_j \leftarrow \frac{1}{|I_j|} \sum_{i \in I_j} \boldsymbol{\xi}_i$, in which case we recover the classical k-means clustering algorithm. Although the algorithm terminates at a local minimum, Dasgupta (2008) has shown that for $l = 2$ and $\|\cdot\| = \|\cdot\|_2$, the solution determined by the algorithm can be arbitrarily suboptimal. We now generalize this finding to generic type-l Wasserstein distances induced by arbitrary p-norms.

Theorem 2.16. If initialized randomly in Step 1, the approximation ratio of the k-means clustering algorithm is unbounded for every $d, l, p \geq 1$ with significant probability.

Proof. In analogy to the proof of Theorem 2.15, we construct a specific distribution $\hat{\mathbb{P}}_n$ (Step 1), bound $D_l(\hat{\mathbb{P}}_n, m)$ from above (Step 2) and bound the Wasserstein distance between $\hat{\mathbb{P}}_n$ and the output $\mathbb{Q}$ of the k-means algorithm from below (Step 3).

Step 1: Fix $d, l, p \geq 1$ and m = 3, and consider the empirical distribution $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, m)$ with $n = 3z + 1$ for some positive integer z as well as $\operatorname{supp}(\hat{\mathbb{P}}_n) = \Xi_1 \cup \Xi_2 \cup \{\boldsymbol{\xi}_{3z+1}\}$ with $\Xi_1 = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{2z}\}$, $\Xi_2 = \{\boldsymbol{\xi}_{2z+1}, \ldots, \boldsymbol{\xi}_{3z}\}$ and

$$\Xi_1 \subset \mathscr{B}_\epsilon(-\mathbf{e}_1), \quad \Xi_2 \subset \mathscr{B}_\epsilon(\mathbf{0}) \quad \text{for} \quad \epsilon \in (0, 1/4), \tag{2.20}$$

as well as $\boldsymbol{\xi}_{3z+1} = \mathbf{e}_1$, where again $\mathscr{B}_\epsilon(\boldsymbol{x}) = \{\boldsymbol{\xi} \in \mathbb{R}^d : \|\boldsymbol{\xi} - \boldsymbol{x}\|_p \leq \epsilon\}$. By construction, the distance between any pair of atoms in Ξ_j is bounded above by $2\epsilon < \frac{1}{2}$, $j = 1,2$, whereas the distance between two atoms from Ξ_1 and Ξ_2 is bounded from below by $1 - 2\epsilon > \frac{1}{2}$.

Step 2: A similar argument as in the proof of Theorem 2.15 shows that

$$\begin{aligned} C_l(\hat{\mathbb{P}}_n, 3) \;&\leq\; d_l\left(\hat{\mathbb{P}}_n, \frac{2k}{3k+1}\delta_{-\mathbf{e}_1} + \frac{k}{3k+1}\delta_{\mathbf{0}} + \frac{1}{3k+1}\delta_{\mathbf{e}_1}\right) \\ &\leq \left[\frac{1}{3z+1}\left(\sum_{\boldsymbol{\xi}_i \in \Xi_1} \|\boldsymbol{\xi}_i + \mathbf{e}_1\|_p^l + \sum_{\boldsymbol{\xi}_i \in \Xi_2} \|\boldsymbol{\xi}_i\|_p^l\right)\right]^{1/l} \;\leq\; \left[\frac{3z\epsilon^l}{3z+1}\right]^{1/l}, \end{aligned}$$

where the last inequality follows from (2.20).

Step 3: We first show that with significant probability, the algorithm chooses a reduced set R containing two atoms from Ξ_1 and one atom from Ξ_2 in the first step. We then show that under this initialization, the output $\mathbb{Q}$ of the algorithm can be arbitrarily worse than the bound on $C_l(\hat{\mathbb{P}}_n, 3)$ determined above.

In view of the first point, we note that the probability of the reduced set R containing two atoms from Ξ_1 and one atom from Ξ_2 after the first step is $\binom{2z}{2}\binom{z}{1}/\binom{3z+1}{3}$ and approaches 44.44% as $z \to \infty$. In the following, we thus assume w.l.o.g. that $R = \{\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2, \boldsymbol{\zeta}_3\}$ with $\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2 \in \Xi_1$ and $\boldsymbol{\zeta}_3 \in \Xi_2$ after the first step.

As for the second point, we note that Step 2 of the algorithm assigns the atoms $\boldsymbol{\xi}_i \in \Xi_1$ to either $\boldsymbol{\zeta}_1$ or $\boldsymbol{\zeta}_2$, whereas the atoms $\boldsymbol{\xi}_i \in \Xi_2 \cup \{\boldsymbol{\xi}_{3z+1}\}$ are assigned to $\boldsymbol{\zeta}_3$. Hence, the update of the reduced set R in the next iteration satisfies $\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2 \in \mathscr{B}_\epsilon(-\mathbf{e}_1)$, whereas $\boldsymbol{\zeta}_3$ is chosen with respect to the set $\Xi_2 \cup \{\boldsymbol{\xi}_{3z+1}\}$. The algorithm then terminates in the third iteration as the reduced set R no longer changes. We thus find that

$$d_l(\hat{\mathbb{P}}_n, \mathbb{Q}) \geq \left[\frac{1}{3z+1}\sum_{\boldsymbol{\xi}_i \notin \Xi_1} \|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_3\|_p^l\right]^{1/l} \geq \left[\frac{1}{3z+1}\left(\|\boldsymbol{\xi}_{3z} - \boldsymbol{\zeta}_3\|_p^l + \|\boldsymbol{\xi}_{3z+1} - \boldsymbol{\zeta}_3\|_p^l\right)\right]^{1/l}.$$

Recall that $\boldsymbol{\xi}_{3z} \in \mathscr{B}_\epsilon(\mathbf{0})$ and $\boldsymbol{\xi}_{3z+1} = \mathbf{e}_1$, which implies that

$$\|\boldsymbol{\xi}_{3z} - \boldsymbol{\zeta}_3\|_p + \|\boldsymbol{\xi}_{3z+1} - \boldsymbol{\zeta}_3\|_p \;\geq\; \|\boldsymbol{\xi}_{3z+1} - \boldsymbol{\xi}_{3z}\|_p \;\geq\; 1 - \epsilon,$$

and that at least one of the two terms $\|\boldsymbol{\xi}_{3z} - \boldsymbol{\zeta}_3\|_p$ or $\|\boldsymbol{\xi}_{3z+1} - \boldsymbol{\zeta}_3\|_p$ is greater than $\frac{1-\epsilon}{2}$. We thus conclude that

$$d_l(\hat{\mathbb{P}}_n, \mathbb{Q}) \;\geq\; \left[\frac{(1-\epsilon)^l}{2^l(3z+1)}\right]^{1/l},$$

which by virtue of Step 2 implies that $d_l(\hat{\mathbb{P}}_n, \mathbb{Q}) / C_l(\hat{\mathbb{P}}_n, 3) \to \infty$ as $\epsilon \to 0$. □

2.4.2 Constant-Factor Approximation for the Scenario Reduction Problem

We now propose a simple approximation scheme for the discrete scenario reduction problem under the type-1 Wasserstein distance whose approximation ratio is bounded from above by 5. We also show that this algorithm gives rise to an approximation scheme for the *continuous* scenario reduction problem with an approximation ratio of 10. To our best knowledge, we describe the first constant-factor approximations for the discrete and continuous scenario reduction problems.

Our algorithm follows from the insight that the discrete scenario reduction problem under the type-1 Wasserstein distance is equivalent to the k-median clustering problem. The k-median clustering problem is a variant of the k-means clustering problem described in Section 2.4.1, where the l-th power of the norm terms is dropped (*i.e.*, $l = 1$). In the following, we adapt a well-known local search algorithm (Arya et al. 2004) to our discrete scenario reduction problem:

LOCAL SEARCH ALGORITHM FOR $D_l(\hat{\mathbb{P}}_n, m)$:

1. Initialize the reduced set $R \subseteq \operatorname{supp}(\hat{\mathbb{P}}_n)$, $|R| = m$, arbitrarily.
2. Select the next exchange to be applied to the reduced set as

   $$(\boldsymbol{\zeta}, \boldsymbol{\zeta}') \in \arg\min \left\{ D_l(\hat{\mathbb{P}}_n, R \cup \{\boldsymbol{\zeta}\} \setminus \{\boldsymbol{\zeta}'\}) : (\boldsymbol{\zeta}, \boldsymbol{\zeta}') \in \left(\operatorname{supp}(\hat{\mathbb{P}}_n) \setminus R\right) \times R \right\},$$

   and update $R \leftarrow R \cup \{\boldsymbol{\zeta}\} \setminus \{\boldsymbol{\zeta}'\}$ if $D_l(\hat{\mathbb{P}}_n, R \cup \{\boldsymbol{\zeta}\} \setminus \{\boldsymbol{\zeta}'\}) < D_l(\hat{\mathbb{P}}_n, R)$.
3. Repeat Step 2 until no further improvement is possible.

For an empirical distribution $\hat{\mathbb{P}}_n \in \mathscr{P}_{\mathrm{E}}(\mathbb{R}^d, n)$, the algorithm constructs a sequence of reduced sets R containing the atoms of the reduced distribution $\mathbb{Q}$. In each iteration, Step 2 selects the exchange $R \cup \{\boldsymbol{\zeta}\} \setminus \{\boldsymbol{\zeta}'\}$, $\boldsymbol{\zeta} \in \operatorname{supp}(\hat{\mathbb{P}}_n)$ and $\boldsymbol{\zeta}' \in R$, that maximally reduces the Wasserstein distance $D_l(\hat{\mathbb{P}}_n, R)$. For performance reasons, this 'best fit' strategy can also be replaced with a 'first fit' strategy which conducts the first exchange $R \cup \{\boldsymbol{\zeta}\} \setminus \{\boldsymbol{\zeta}'\}$ found that leads to a reduction of $D_l(\hat{\mathbb{P}}_n, R)$. After termination, the reduced distribution $\mathbb{Q}$ can be recovered from the reduced set R in the same way as in Section 2.4.1.

It follows from Arya et al. (2004) that the above algorithm (with either 'best fit' or 'first fit') has an approximation ratio of 5 for the discrete scenario reduction problem for all d. We now show that the algorithm also provides solutions to the *continuous* scenario reduction problem with an approximation ratio of at most 10.

Corollary 2.17. The problems $D_l(\hat{\mathbb{P}}_n, m)$ and $C_l(\hat{\mathbb{P}}_n, m)$ are related as follows.

1. Any approximation algorithm for $D_2(\hat{\mathbb{P}}_n, m)$ under the 2-norm with approximation

ratio α gives rise to an approximation algorithm for $C_2(\hat{\mathbb{P}}_n, m)$ under the 2-norm with approximation ratio $\sqrt{2}\alpha$.

2. Any approximation algorithm for $D_1(\hat{\mathbb{P}}_n, m)$ under any norm with approximation ratio α gives rise to an approximation algorithm for $C_1(\hat{\mathbb{P}}_n, m)$ under the same norm with approximation ratio 2α.

Proof. The two statements follow directly from Theorems 2.9 and 2.12, respectively. □

As presented, the local search algorithm is not guaranteed to terminate in polynomial time. This can be remedied by a variant of the algorithm that only accepts exchanges $R \cup \{\boldsymbol{\zeta}\} \setminus \{\boldsymbol{\zeta}'\}$ that reduce the Wasserstein distance $D_l(\hat{\mathbb{P}}_n, R)$ by at least $\epsilon/((n-m)m)$ for some constant $\epsilon > 0$. It follows from Arya et al. (2004) that for any ϵ, this variant terminates in polynomial time and provides a $(5+\epsilon)$-approximation for the discrete scenario reduction problem. The algorithm can also be extended to accommodate multiple swaps in every iteration, which lowers the approximation ratio to $3+\epsilon$ at the expense of additional computations.

We remark that there is a wealth of algorithms for the k-median problem that can be adapted to the discrete scenario reduction problem. For example, Charikar and Li (2012) present a rounding scheme for the k-median problem that gives rise to a polynomial-time algorithm for $D_1(\hat{\mathbb{P}}_n, m)$ with an approximation ratio of 3.25. Likewise, Li and Svensson (2016) introduce a pseudo-algorithm for the k-median problem that offers, to our knowledge, the best currently available constant approximation guarantee of $1+\sqrt{3}+\epsilon$. Finally, Kanungo et al. (2004) propose a local search algorithm for the k-median problem that gives rise to a polynomial-time algorithm for $D_2(\hat{\mathbb{P}}_n, m)$ under the 2-norm with an approximation ratio of $9+\epsilon$. In all three cases, Corollary 2.17 allows us to extend these guarantees to the corresponding versions of the continuous scenario reduction problem.

2.4.3 Mixed-Integer Reformulations of the Discrete and Continuous Scenario Reduction Problems

We first review a well-known mixed-integer linear programming (MILP) reformulation of the discrete scenario reduction problem $D_l(\hat{\mathbb{P}}_n, m)$:

Theorem 2.18. The discrete scenario reduction problem can be formulated as the MILP

$$\begin{array}{rll} D_l^l(\hat{\mathbb{P}}_n, m) = & \min\limits_{\Pi, \boldsymbol{\lambda}} & \dfrac{1}{n}\langle \Pi, \mathbf{D}\rangle \\ & s.t. & \Pi\mathbf{e} = \mathbf{e},\ \Pi \le \mathbf{e}\boldsymbol{\lambda}^\top,\ \boldsymbol{\lambda}^\top\mathbf{e} = m \\ & & \Pi \in \mathbb{R}_+^{n\times n},\ \boldsymbol{\lambda} \in \{0,1\}^n, \end{array} \tag{2.21}$$

with $\mathbf{D} \in \mathbb{S}^n$ and $d_{ij} = \|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\|^l$ encoding the distances among the atoms in $\mathrm{supp}(\hat{\mathbb{P}}_n)$.

Proof. See, *e.g.*, Heitsch and Römisch (2003). □

In problem (2.21), the decision variable π_{ij} determines how much of the probability mass of atom $\boldsymbol{\xi}_i$ in $\hat{\mathbb{P}}_n$ is shifted to the atom $\boldsymbol{\zeta}_j$ in the reduced distribution $\mathbb{Q}$, whereas the decision variable λ_j determines whether the atom $\boldsymbol{\xi}_j \in \operatorname{supp}(\hat{\mathbb{P}}_n)$ is contained in the support of $\mathbb{Q}$. A solution $(\Pi^\star, \boldsymbol{\lambda}^\star)$ to problem (2.21) allows us to recover the reduced distribution via $\mathbb{Q} = \frac{1}{n}\sum_{j=1}^n \mathbf{e}^\top \Pi^\star \mathbf{e}_j \cdot \delta_{\boldsymbol{\xi}_j}$. Problem (2.21) has n binary and n^2 continuous variables as well as n^2+n+1 constraints.

The runtime of any known algorithm for solving the MILP (2.21) is superpolynomial in n and m because the discrete scenario reduction problem is $\mathcal{NP}$-hard. For a fixed m however, the discrete scenario reduction problem can be solved exactly in polynomial time by, for example, exhaustively enumerating $\binom{n}{m}$ possible supports of the reduced distribution. Note that the optimally reduced *uniform* distribution supported on $m = n/2$ points can be determined much more efficiently by solving a matching problem (Stockbridge and Bayraksan 2013). We highlight, however, that our reduced distribution $\mathbb{Q}$ is not necessarily uniform.

We now consider the *continuous* scenario reduction problem. Due to its bilinear objective function, which involves products of transportation weights π_{ij} and the distances $\|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j\|^l$ containing the continuous decision variables $\boldsymbol{\zeta}_j$, this problem may not appear to be amenable to a reformulation as a mixed-integer convex optimization problem. We now show that such a reformulation indeed exists.

Theorem 2.19. The continuous scenario reduction problem can be formulated as the mixed-integer convex optimization problem

$$
\begin{array}{rcll}
C_l^l(\hat{\mathbb{P}}_n, m) \;=\; \min\limits_{\Pi, \boldsymbol{c}, \{\boldsymbol{\zeta}_j\}} & \dfrac{1}{n}\mathbf{e}^\top \boldsymbol{c} & & \\
\text{s.t.} & \Pi\mathbf{e} = \mathbf{e} & & \\
& \|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j\|^l \le c_i + M(1-\pi_{ij}) & \forall\, i \in I,\, \forall\, j \in J & \\
& \Pi \in \{0,1\}^{n\times m},\;\; \boldsymbol{c} \in \mathbb{R}^n_+,\;\; \boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_m \in \mathbb{R}^d, &
\end{array}
\tag{2.22}
$$

where $M = \max_{i,j\in I} \|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\|^l$ denotes the diameter of the support of $\hat{\mathbb{P}}_n$.

In problem (2.22), the decision variable π_{ij} determines whether or not the probability mass of atom $\boldsymbol{\xi}_i$ in the empirical distribution $\hat{\mathbb{P}}_n$ is shifted to the atom $\boldsymbol{\zeta}_j$ in the reduced distribution $\mathbb{Q}$, whereas the decision variable c_i records the cost of moving the atom $\boldsymbol{\xi}_i$ under the transportation plan Π.

Proof of Theorem 2.19. We prove the statement by showing that optimal solutions to problem (2.22) correspond to feasible solutions in problem (2.2) with the same objective function value in their respective problems and vice versa.

Fix a minimizer $(\Pi^\star, \boldsymbol{c}^\star, \boldsymbol{\zeta}_1^\star, \ldots, \boldsymbol{\zeta}_m^\star)$ to problem (2.22), which corresponds to a feasible solution $(\{I_j\}, \boldsymbol{\zeta}_1^\star, \ldots, \boldsymbol{\zeta}_m^\star)$ in problem (2.2) if we set $I_j = \{i \in I : \pi_{ij}^\star = 1\}$ for all $j \in J$. Note that $c_i^\star = \|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j^\star\|^l$ for $j \in J$ and $i \in I_j$. Thus, both solutions adopt the same objective value.

Conversely, fix a minimizer $(\{I_j^\star\}, \boldsymbol{\zeta}_1^\star, \ldots, \boldsymbol{\zeta}_m^\star)$ to problem (2.2). This solution corresponds to a feasible solution $(\Pi, \boldsymbol{c}, \boldsymbol{\zeta}_1^\star, \ldots, \boldsymbol{\zeta}_m^\star)$ to problem (2.22) if we set $\pi_{ij} = 1$ if $i \in I_j^\star$ and $\pi_{ij} = 0$ otherwise for all $j \in J$, as well as $c_i = \|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j^\star\|^l$ for all $i \in I_j$ and $j \in J$. By construction, both solutions adopt the same objective value. □

A solution $(\Pi^\star, \boldsymbol{c}^\star, \boldsymbol{\zeta}_1^\star, \ldots, \boldsymbol{\zeta}_m^\star)$ to problem (2.22) allows us to recover the reduced distribution via $\mathbb{Q} = \frac{1}{n}\sum_{j=1}^m \mathbf{e}^\top \Pi^\star \mathbf{e}_j \cdot \delta_{\boldsymbol{\zeta}_j^\star}$. Problem (2.22) has nm binary and $n + md$ continuous variables as well as $nm + n$ constraints. We now show that (2.22) typically reduces to an MILP or a mixed-integer second-order cone program (MISOCP).

Proposition 2.20. For the type-1 Wasserstein distance induced by $\|\cdot\|_1$ or $\|\cdot\|_\infty$, problem (2.22) reduces to an MILP. For any type-l Wasserstein distance induced by $\|\cdot\|_p$, where $l \geq 1$ and $p \geq 1$ are rational numbers, problem (2.22) reduces to an MISOCP.

Proof. In view of the first statement, we note that $\|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j\|_1 \leq c_i + M(1 - \pi_{ij})$ is satisfied if and only if there is $\boldsymbol{\phi}_{ij} \in \mathbb{R}^d$ such that

$$\boldsymbol{\phi}_{ij} \geq \boldsymbol{\xi}_i - \boldsymbol{\zeta}_j,\ \boldsymbol{\phi}_{ij} \geq \boldsymbol{\zeta}_j - \boldsymbol{\xi}_i \text{ and } \mathbf{e}^\top \boldsymbol{\phi}_{ij} \leq c_i + M(1 - \pi_{ij}).$$

Likewise, $\|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j\|_\infty \leq c_i + M(1 - \pi_{ij})$ holds if and only if there is $\phi_{ij} \in \mathbb{R}$ with

$$\phi_{ij}\mathbf{e} \geq \boldsymbol{\xi}_i - \boldsymbol{\zeta}_j,\ \phi_{ij}\mathbf{e} \geq \boldsymbol{\zeta}_j - \boldsymbol{\xi}_i \text{ and } \phi_{ij} \leq c_i + M(1 - \pi_{ij}).$$

As for the second statement, we note that $\|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j\|_p^l \leq c_i + M(1 - \pi_{ij})$ is satisfied if and only if there is $\phi_{ij} \in \mathbb{R}$ such that

$$\phi_{ij} \geq \|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_j\|_p \text{ and } \phi_{ij}^l \leq c_i + M(1 - \pi_{ij}).$$

For rational $l, p \geq 1$, both inequalities can be expressed through finitely many second-order cone constraints (Alizadeh and Goldfarb 2003, Section 2.3). □

2.5 Numerical Experiments

In the following, we explore the performance of discrete scenario reduction in a color quantization problem (Section 2.5.1) and a stochastic capital budgeting problem (Section 2.5.2).

2.5.1 Color Quantization

Color quantization aims to reduce the color palette of a digital image without compromising its visual appearance. In the standard RGB24 model colors are encoded by vectors of the form $(r, g, b) \in \{0, 1, \ldots, 255\}^3$. This means that the RGB24 model can represent a vast number of

16,777,216 distinct colors. Consequently, color quantization serves primarily as a lossy image compression method.

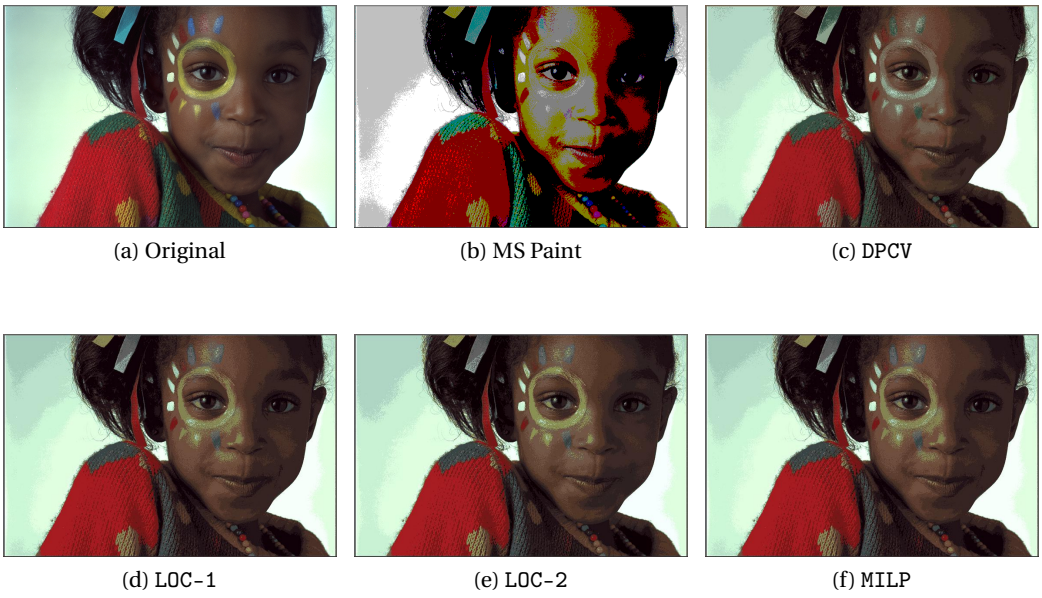
In the following we interpret the color quantization problem as a discrete scenario reduction problem using the type-1 Wasserstein distance induced by the 1-norm on $\mathbb{R}^3$. Thus, we can solve color quantization problems via Dupačová's greedy heuristic, the local search algorithm or the exact MILP reformulation (2.21). In our experiment we aim to compress all 24 pictures from the Kodak Lossless True Color Image Suite (http://r0k.us/graphics/kodak/) to $m = 2^1, \ldots, 2^9$ colors. As the MILP reformulation scales poorly with n, we first reduce each image to $n \lesssim 1{,}024$ colors using the Linux command "convert -colors", which is distributed through ImageMagick (https://www.imagemagick.org). We henceforth refer to the resulting 1,024-color images as the originals.

In all experiments we use an efficient variant of Dupačová's algorithm due to Heitsch and Römisch (2003) (`DPCV`), and we initialize the local search algorithm either with the color palette obtained from Dupačová's algorithm (`LOC-1`) or naïvely with the m most frequent colors of the original image (`LOC-2`). The MILP (2.21) is solved with GUROBI 7.0.1 (`MILP`). All algorithms are implemented in C++, and all experiments are executed on a 3.40GHz i7 CPU machine with 16GB RAM. We report the average and the worst-case runtimes in Table 2.1. Note that `DPCV`, `LOC-1` and `LOC-2` all terminate in less than 14 seconds across all instances, while `MILP` requires substantially more time (the maximum runtime was set to ten hours). Moreover, warmstarting the local search algorithm with the color palette obtained from `DPCV` can significantly reduce the runtimes.

	DPCV	LOC-1	LOC-2	MILP
Average (secs)	2.16	2.60	3.59	1,349.71
Worst-case (secs)	8.21	9.89	13.69	36,120.99

Table 2.1 – Runtimes of different methods for discrete scenario reduction.

As an example, Figure 2.3 shows the image "kodim15.png" as well as the results of different color quantization algorithms for $m = 16$. While the outputs of `LOC-1`, `LOC-2` and `MILP` are almost indistinguishable, the output of `DPCV` has ostensible deficiencies (*e.g.*, it misrepresents the yellow color around the subject's eye). For comparison, we also show the output of the color quantization routine in Microsoft Paint (`MS Paint`). Figure 2.4 visualizes the optimality gaps of `DPCV`, `LOC-1` and `LOC-2` relative to `MILP` (*i.e.*, their respective approximation ratio $-$ 1). Our experiment suggests that the local search algorithm is competitive with `MILP` in terms of output quality but at significantly reduced runtimes. Moreover, the local search algorithm `LOC-1` warmstarted with the color palette obtained from `DPCV` is guaranteed to outperform `DPCV` in terms of optimality gaps.



(a) Original (b) MS Paint (c) DPCV

(d) LOC-1 (e) LOC-2 (f) MILP

Figure 2.3 – Outputs of different color quantization algorithms for image “kodim15.png”.

2.5.2 Capital Budgeting

We follow Hanasusanto et al. (2015b) in considering an investment planning problem of a venture capitalist. With a total budget of B, the venture capitalist wishes to sponsor a subset of N different projects. She may decide to become an early adopter of a project k by investing before observing some risk factors $\boldsymbol{\xi} \in \mathbb{R}^F$, or she may choose to become a laggard by investing after observing $\boldsymbol{\xi}$. To fund a project $k \in \{1,\ldots,N\}$, the venture capitalist needs to pay an uncertain cost of $c_k(\boldsymbol{\xi})$. After providing funding for project k, she earns a random profit of $r_k(\boldsymbol{\xi})$ as an early adopter, or $\theta r_k(\boldsymbol{\xi})$ for some $\theta \in (0,1)$ as a laggard.

Supposing further that the venture capitalist is risk-neutral, we can formulate a profit-maximizing two-stage stochastic program as

$$\begin{array}{ll} \max & \mathbb{E}_{\mathbb{P}}\left(\boldsymbol{r}(\boldsymbol{\xi})^\top(\boldsymbol{x}+\theta\boldsymbol{y}(\boldsymbol{\xi}))\right) \\ \text{s.t.} & \left.\begin{array}{l} \boldsymbol{x},\boldsymbol{y}(\boldsymbol{\xi}) \in \{0,1\}^N \\ \boldsymbol{c}(\boldsymbol{\xi})^\top(\boldsymbol{x}+\boldsymbol{y}(\boldsymbol{\xi})) \le B \\ \boldsymbol{x}+\boldsymbol{y}(\boldsymbol{\xi}) \le \mathbf{1} \end{array}\right\} \mathbb{P}\text{-a.s.}, \end{array} \tag{2.23}$$

where $\mathbb{P}$ denotes the distribution of the risk factors $\boldsymbol{\xi}$. Here, the binary variables x_k and $y_k(\boldsymbol{\xi})$ adopt the value 1 if the venture capitalist decides to invest in project k before and after observing the risk factors $\boldsymbol{\xi}$, respectively.

We solve the capital budgeting problem (2.23) in a data-driven fashion, whereby the unknown true distribution $\mathbb{P}$ of the random vector $\boldsymbol{\xi}$ is replaced with an empirical distribution $\hat{\mathbb{P}}_n$ on

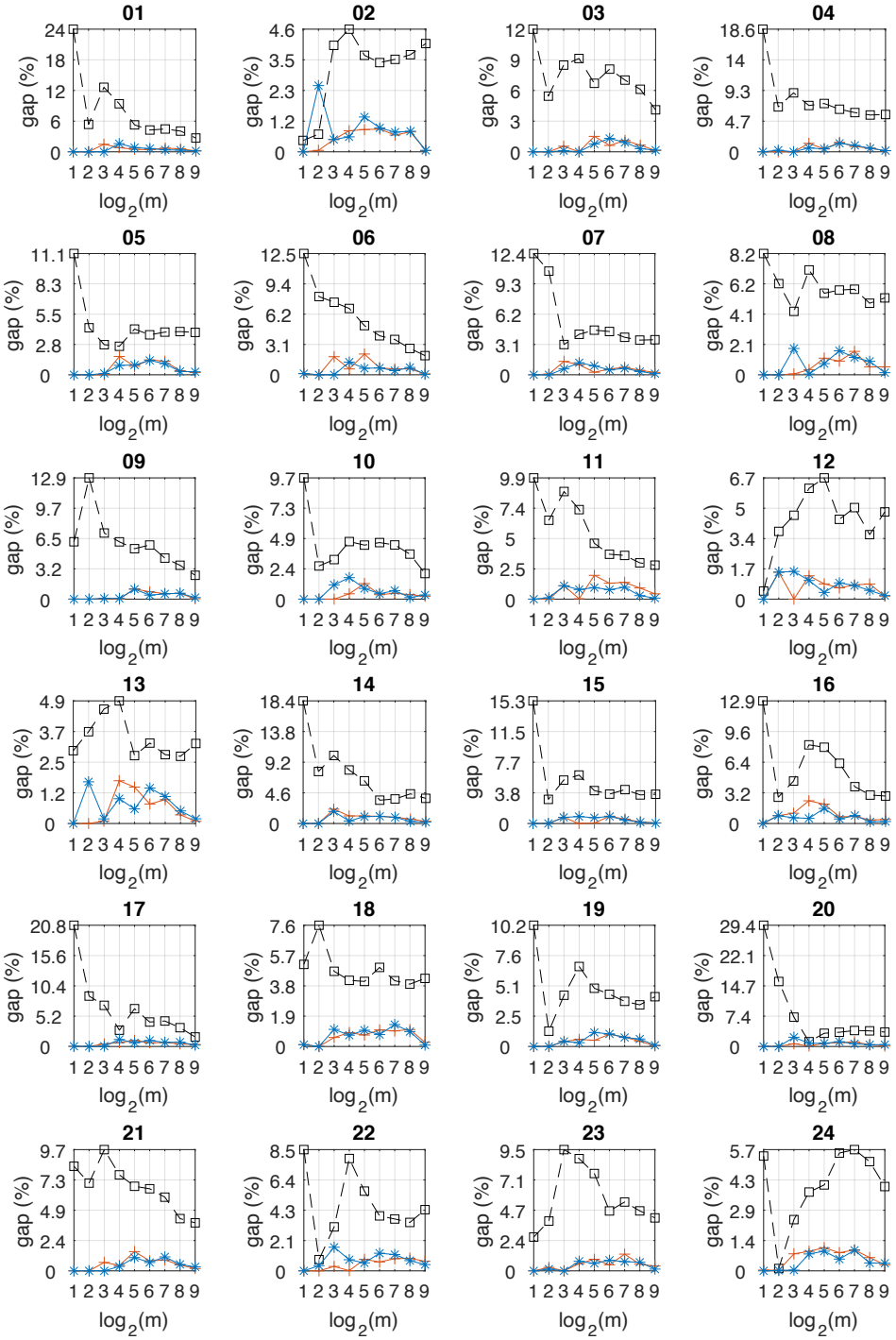


Figure 2.4 – Optimality gaps of DPCV (dashed lines with boxes), LOC-1 (solid lines with pluses) and LOC-2 (solid lines with stars) relative to MILP.

n independent samples from $\mathbb{P}$. In doing so, (2.23) can be equivalently cast as an integer optimization problem with $N(n+1)$ binary variables $\boldsymbol{x}, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n \in \{0,1\}^N$, where $\boldsymbol{y}_i$ captures the second-stage decisions in scenario $\boldsymbol{\xi}_i$.

In our numerical experiments, we consider $N = 10$ projects and $F = 3$ risk factors. Furthermore, we model the cost and profit of project k as

$$c_k(\boldsymbol{\xi}) = (1 + \Phi_k^\top \boldsymbol{\xi}/2) c_k^0 \quad \text{and} \quad r_k(\boldsymbol{\xi}) = (1 + \Psi_k^\top \boldsymbol{\xi}/2) r_k^0,$$

where c_k^0 and r_k^0 denote the nominal cost and the nominal profit of project k, respectively. Similarly, the vectors Φ_k and Ψ_k can be interpreted as the factor-loadings of the nominal cost and profit of project k, respectively.

We repeat the following experiment 100 times. In each round, we independently sample empirical risk factors $\boldsymbol{\xi}_i$ $(1 \le i \le n = 200)$ uniformly from the hypercube $[-1,1]^F$, and the nominal costs $\boldsymbol{c}^0$ from the hypercube $[0,10]^N$. We further set $\boldsymbol{r}^0 = \boldsymbol{c}^0/5$, $B = \mathbf{e}^\top \boldsymbol{c}^0/2$, and $\theta = 0.9$. Following Rubinstein and Kroese (2007), the factor loadings Φ_k and Ψ_k $(1 \le k \le N)$ are sampled uniformly from the probability simplex in $\mathbb{R}^F$. The empirical distribution $\hat{\mathbb{P}}_n$ is then reduced to a distribution $\mathbb{Q}$ supported on $m \in \{20, 40, \ldots, 200\}$ scenarios by using the local search algorithm warmstarted by Dupačová's algorithm (`LOC-1`). We compare the optimal objective values and runtimes of the emerging reduced integer programs with those of the full integer programs in Figure 2.5. All integer programs are solved with GUROBI 7.0.1.

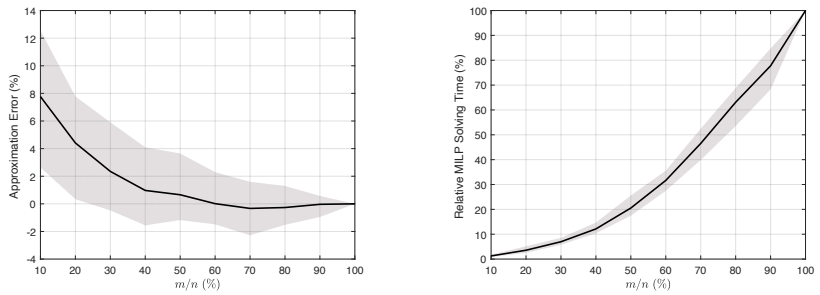


Figure 2.5 – Approximation errors (left panel) and runtimes (right panel) of the reduced capital budgeting problem for different reduction factors m/n. The solid lines represent the averages, while the shaded areas capture the intervals between the 10% and 90% quantiles across 100 independent simulation runs.

Our numerical experiments suggest that scenario reduction may offer a first-order reduction in runtime at the expense of a mere second-order reduction in accuracy.

Appendix: Auxiliary Results

The proof of Theorem 2.3 relies on the following two lemmas.

Lemma 2.21. The semidefinite program (2.6) admits an optimal solution $(\tau, \mathbf{S})$ with $\mathbf{S} = \alpha\mathbb{I} + \beta\mathbf{1}\mathbf{1}^\top$ for some $\alpha, \beta \in \mathbb{R}$.

Proof. Let $(\tau, \mathbf{S}^\star)$ be any optimal solution to (2.6), which exists because (2.6) has a continuous objective function and a compact feasible set, and denote by $\mathfrak{S}$ the set of all permutations of I. For any $\sigma \in \mathfrak{S}$, the permuted solution $(\tau, \mathbf{S}^\sigma)$, with $s^\sigma_{ij} = s^\star_{\sigma(i)\sigma(j)}$ is also optimal in (2.6). Note first that $(\tau, \mathbf{S}^\sigma)$ is feasible in (2.6) because

$$\tau \leq \sum_{j\in J} \frac{1}{|I_j|^2} \sum_{i\in I_j} \left(|I_j|^2 s^\sigma_{ii} - 2|I_j| \sum_{k\in I_j} s^\sigma_{ik} + \sum_{k\in I_j} s^\sigma_{kk} + \sum_{\substack{k,k'\in I_j \\ k\neq k'}} s^\sigma_{kk'} \right)$$

$$\iff \tau \leq \sum_{j\in J} \frac{1}{|I^\sigma_j|^2} \sum_{i\in I^\sigma_j} \left(|I^\sigma_j|^2 s^\star_{ii} - 2|I^\sigma_j| \sum_{k\in I_j} s^\star_{ik} + \sum_{k\in I_j} s^\star_{kk} + \sum_{\substack{k,k'\in I_j \\ k\neq k'}} s^\star_{kk'} \right),$$

where the index sets $I^\sigma_j = \{\sigma(i) : i \in I_j\}$ for $j \in J$ form an m-set partition from within $\mathfrak{P}(I, m)$, and because $\mathbf{S}^\sigma \succeq \mathbf{0}$ and $s^\sigma_{ii} = s^\star_{\sigma(i)\sigma(i)} \leq 1$ for all $i \in I$ by construction. Moreover, it is clear that $(\tau, \mathbf{S}^\sigma)$ and $(\tau, \mathbf{S}^\star)$ share the same objective value in (2.6). Thus, $(\tau, \mathbf{S}^\sigma)$ is optimal in (2.6) for every $\sigma \in \mathfrak{S}$.

The convexity of problem (2.6) implies that $(\tau, \mathbf{S})$ with $\mathbf{S} = \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}} \mathbf{S}^\sigma$ is also optimal in (2.6). The claim follows by noting that $\mathbf{S}$ is invariant under permutations of the coordinates and thus representable as $\alpha\mathbb{I} + \beta\mathbf{1}\mathbf{1}^\top$ for some $\alpha, \beta \in \mathbb{R}$. □

Lemma 2.22. For $\alpha, \beta \in \mathbb{R}$ the eigenvalues of $\mathbf{S} = \alpha\mathbb{I} + \beta\mathbf{1}\mathbf{1}^\top \in \mathbb{S}^n$ are given by $\alpha + n\beta$ (with multiplicity 1) and α (with multiplicity $n-1$).

Proof. Note that $\mathbf{S}$ is a circulant matrix, meaning that each of its rows coincides with the preceding row rotated by one element to the right. Thus, the eigenvalues of $\mathbf{S}$ are given by $\alpha + \beta(1 + \rho_j^1 + \ldots \rho_j^{n-1})$, $j = 0, \ldots, n-1$, where $\rho_j = e^{2\pi i j/n}$ and i denotes the imaginary unit; see, *e.g.*, Gray (2006). For $j = 0$ we then obtain the eigenvalue $\alpha + n\beta$, and for $j = 1, \ldots, n-1$ we obtain the other $n-1$ eigenvalues, all of which equal α because $\sum_{k=0}^{n-1} e^{2\pi i j k/n} = (1 - e^{2\pi i j})/(1 - e^{2\pi i j/n}) = 0$. □

3 Cardinality-Constrained Clustering and Outlier Detection via Conic Optimization

Plain vanilla K-means clustering has proven to be successful in practice, yet it suffers from outlier sensitivity and may produce highly unbalanced clusters. To mitigate both shortcomings, we formulate a joint outlier detection and clustering problem, which assigns a prescribed number of datapoints to an auxiliary outlier cluster and performs cardinality-constrained K-means clustering on the residual dataset, treating the cluster cardinalities as a given input. We cast this problem as a mixed-integer linear program (MILP) that admits tractable semidefinite and linear programming relaxations. We propose deterministic rounding schemes that transform the relaxed solutions to feasible solutions for the MILP. We also prove that these solutions are optimal in the MILP if a cluster separation condition holds.

3.1 Introduction

Clustering aims to partition a set of datapoints into a set of clusters so that datapoints in the same cluster are more similar to one another than to those in other clusters. Among the myriad of clustering approaches from the literature, K-means clustering stands out for its long history dating back to 1957 as well as its impressive performance in various application domains, ranging from market segmentation and recommender systems to image segmentation and feature learning (Jain 2010b).

This paper studies the *cardinality-constrained K-means clustering problem*, which we define as the task of partitioning N datapoints $\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_N \in \mathbb{R}^d$ into K clusters $I_1,\ldots,I_K$ of prescribed sizes $n_1,\ldots,n_K$, with $n_1+\ldots+n_K=N$, so as to minimize the sum of squared intra-cluster distances. We can formalize the cardinality-constrained K-means clustering problem as follows,

$$\begin{array}{ll} \text{minimize} & \sum_{k=1}^{K}\sum_{i\in I_k}\|\boldsymbol{\xi}_i - \frac{1}{n_k}(\sum_{j\in I_k}\boldsymbol{\xi}_j)\|^2 \\ \text{subject to} & (I_1,\ldots,I_K)\in\mathfrak{P}(n_1,\ldots,n_K), \end{array} \tag{3.1}$$

where

$$\mathfrak{P}(n_1,\ldots,n_K) = \left\{(I_1,\ldots,I_K) : |I_k| = n_k \ \ \forall k,\ \ \cup_{k=1}^K I_k = \{1,\ldots,N\},\ \ I_k \cap I_\ell = \emptyset \ \ \forall k \neq \ell\right\}$$

denotes the ordered partitions of the set $\{1,\ldots,N\}$ into K sets of sizes $n_1,\ldots,n_K$, respectively.

Our motivation for studying problem (3.1) is threefold. Firstly, it has been shown by Bennett et al. (2000) and Chen et al. (2006) that the algorithms commonly employed for the *un*constrained K-means clustering problem frequently produce suboptimal solutions where some of the clusters contain very few or even no datapoints. In this context, cardinality constraints can act as a regularizer that avoids local minima of poor quality. Secondly, many application domains require the clusters $I_1,\ldots,I_K$ to be of comparable size. This is the case, among others, in distributed clustering (where different computer clusters should contain similar numbers of network nodes), market segmentation (where each customer segment will subsequently be addressed by a marketing campaign) and document clustering (where topic hierarchies should display a balanced view of the available documents); see Banerjee and Ghosh (2006) and Balcan et al. (2013). Finally, and perhaps most importantly, K-means clustering is highly sensitive to outliers. To illustrate this, consider the dataset in Figure 3.1, which accommodates three clusters as well as three individual outliers. The K-means clustering problem erroneously merges two of the three clusters in order to assign the three outliers to the third cluster (top left graph), whereas a clustering that disregards the three outliers would recover the true clusters and result in a significantly lower objective value (bottom left graph). The cardinality-constrained K-means clustering problem, where the cardinality of each cluster is set to be one third of all datapoints, shows a similar behavior on this dataset (graphs on the right). We will argue below, however, that the cardinality-constrained K-means clustering problem (3.1) offers an intuitive and mathematically rigorous framework to robustify K-means clustering against outliers. A comprehensive and principled treatment of outlier detection methods can be found in the book of Aggarwal (2013).

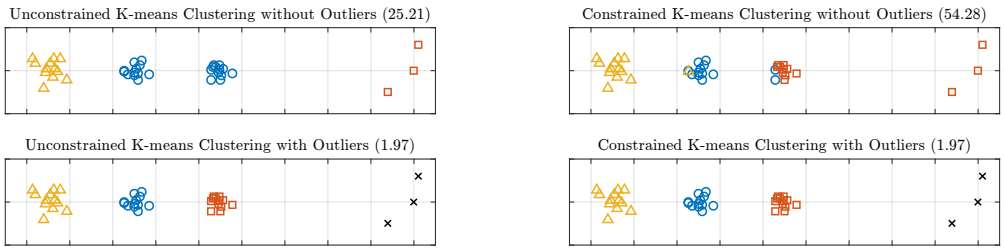


Figure 3.1 – Sensitivity of the (un)constrained K-means clustering problem to outliers. Indicated in parentheses next to the panel titles are the respectively achieved sums of squared intra-cluster distances.

To our best knowledge, to date only two solution approaches have been proposed for problem (3.1). Bennett et al. (2000) combine a classical local search heuristic for the unconstrained

K-means clustering problem due to Lloyd (1982) with the repeated solution of linear assignment problems to solve a variant of problem (3.1) that imposes lower bounds on the cluster sizes $n_1, \ldots, n_K$. Banerjee and Ghosh (2006) solve the balanced version of problem (3.1), where $n_1 = \cdots = n_K$, by sampling a subset of the datapoints, performing a clustering on this subset, and subsequently populating the resulting clusters with the remaining datapoints while adhering to the cardinality constraints. Balanced clustering is also considered by Malinen and Fränti (2014) and Costa et al. (2017). Malinen and Fränti (2014) proceed similarly to Bennett et al. (2000) but take explicit advantage of the Hungarian algorithm to speed up the cluster assignment step within the local search heuristic. Costa et al. (2017) propose a variable neighborhood search heuristic that starts from a random partition of the datapoints into balanced clusters and subsequently searches for better solutions in the neighborhood obtained by an increasing number of datapoint swaps between two clusters. Although all of these heuristics tend to quickly produce solutions of high quality, they are not known to be polynomial-time algorithms, they do not provide bounds on the suboptimality of the identified solutions, and their performance may be sensitive to the choice of the initial solution. Moreover, neither of these local search schemes accommodates for outliers.

In recent years, several conic optimization schemes have been proposed to alleviate the shortcomings of these local search methods for the unconstrained K-means clustering problem (Peng and Wei 2007, Awasthi et al. 2015). Peng and Wei (2007) develop two semidefinite programming relaxations of the unconstrained K-means clustering problem. Their weaker relaxation admits optimal solutions that can be characterized by means of an eigenvalue decomposition. They further use this eigenvalue decomposition to set up a modified K-means clustering problem where the dimensionality of the datapoints is reduced to $K-1$ (provided that their original dimensionality was larger than that). To obtain an upper bound, they solve this K-means clustering problem of reduced dimensionality, which can be done either exactly by enumerating Voronoi partitions, as described in Inaba et al. (1994), or by approximation methods such as those in Hasegawa et al. (1993). Using either approach, the runtime grows polynomially in the number of datapoints N but not in the number of desired clusters K. Hence, this method is primarily suitable for small K. Similar conic approximation schemes have been developed by Elhamifar et al. (2012) and Nellore and Ward (2015) in the context of unconstrained exemplar-based clustering.

Awasthi et al. (2015) and Iguchi et al. (2017) develop probabilistic recovery guarantees for the stronger semidefinite relaxation of Peng and Wei (2007) when the data is generated by a stochastic ball model (*i.e.*, datapoints are drawn randomly from rotation symmetric distributions supported on unit balls). More specifically, they use primal-dual arguments to establish conditions on the cluster separation under which the semidefinite relaxation of Peng and Wei (2007) recovers the underlying clusters with high probability as the number of datapoints N increases. The condition of Awasthi et al. (2015) requires less separation in low dimensions, while the condition of Iguchi et al. (2017) is less restrictive in high dimensions. In addition, Awasthi et al. (2015) consider a linear programming relaxation of the unconstrained K-means clustering problem, and they derive similar recovery guarantees for this relaxation as well.

Two more papers study the recovery guarantees of conic relaxations under a stochastic block model (*i.e.,* the dataset is characterized by a similarity matrix where the expected pairwise similarities of points in the same cluster are higher than those of points in different clusters). Ames (2014) considers the densest K-disjoint-clique problem whose aim is to split a given complete graph into K subgraphs such as to maximize the sum of the average similarities of the resulting subgraphs. K-means clustering can be considered as a specific instance of this broader class of problems. By means of primal-dual arguments, the author derives conditions on the means in the stochastic block model such that his semidefinite relaxation recovers the underlying clusters with high probability as the cardinality of the smallest cluster increases. Vinayak and Hassibi (2016) develop a semidefinite relaxation and regularize it with the trace of the cluster assignment matrix. Using primal-dual arguments they show that, for specific ranges of the regularization parameter, their regularized semidefinite relaxation recovers the true clusters with high probability as the cardinality of the smallest cluster increases. The probabilistic recovery guarantees of Ames (2014) and Vinayak and Hassibi (2016) can also be extended to datasets containing outliers.

	Awasthi et al.	Iguchi et al.	Ames	Vinayak and Hassibi	This Paper
data generating model	stochastic ball	stochastic ball	stochastic block	stochastic block	none/arbitrary
type of relaxation	SDP + LP	SDP	SDP	SDP	SDP + LP
type of guarantee	stochastic	stochastic	stochastic	stochastic	deterministic
guarantee depends on N	yes	yes	yes	yes	no
guarantee depends on d	yes	yes	no	no	no
requires balancedness	yes	yes	no	no	yes
proof technique	primal-dual	primal-dual	primal-dual	primal-dual	valid cuts
access to cardinalities	no	no	no	no	yes
outlier detection	no	no	yes	yes	yes

Table 3.1 – Comparison of Recovery Guarantees for K-means Clustering Relaxations.

In this paper, we propose the first conic optimization scheme for the cardinality-constrained K-means clustering problem (3.1). Our solution approach relies on an exact reformulation of problem (3.1) as an intractable mixed-integer linear program (MILP) to which we add a set of valid cuts before relaxing the resulting model to a tractable semidefinite program (SDP) or linear program (LP). The set of valid cuts is essential in strengthening these relaxations. Both relaxations provide lower bounds on the optimal value of problem (3.1), and they both recover the optimal value of (3.1) whenever a cluster separation condition is met. The latter requires all cluster diameters to be smaller than the distance between any two distinct clusters and, in case of outlier presence, also smaller than the distance between any outlier and any other point. The same condition (in the absence of outliers) was used in Elhamifar et al. (2012) and Awasthi et al. (2015). Our relaxations also give rise to deterministic rounding schemes which produce feasible solutions that are provably optimal in (3.1) whenever the cluster separation condition holds. Table 3.1 compares our recovery guarantees to the results available in the literature. We emphasize that our guarantees are deterministic, that they apply to arbitrary data generating models, that they are dimension-independent, and that they hold for both our SDP and LP

relaxations. Finally, our algorithms extend to instances of (3.1) that are contaminated by outliers and whose cluster cardinalities $n_1, \ldots, n_K$ are not known precisely. We summarize the paper's contributions as follows.

1. We derive a novel MILP reformulation of problem (3.1) that only involves NK binary variables, as opposed to the standard MILP reformulation that contains N^2 binary variables, and whose LP relaxation is at least as tight as the LP relaxation of the standard reformulation.
2. We develop lower bounds which exploit the cardinality information in problem (3.1). Our bounds are tight whenever a cluster separation condition is met. Unlike similar results for other classes of clustering problems, our separation condition is deterministic, model-free and dimension-independent. Furthermore, our proof technique does not rely on the primal-dual argument of SDPs and LPs.
3. We propose deterministic rounding schemes that transform the relaxed solutions to feasible solutions for problem (3.1). The solutions are optimal in (3.1) if the separation condition holds. To our best knowledge, we propose the first tractable solution scheme for problem (3.1) with optimality guarantees.
4. We illustrate that our lower bounds and rounding schemes extend to instances of problem (3.1) that are contaminated by outliers and whose cluster cardinalities are not known precisely.

The remainder of the paper is structured as follows. Section 2 analyzes the cardinality-constrained K-means clustering problem (3.1) and derives the MILP reformulation underlying our solution scheme. Sections 3 and 4 propose and analyze our conic rounding approaches for problem (3.1) in the absence and presence of outliers, respectively. Section 5 presents numerical experiments, and Section 6 gives concluding remarks. Finally, a detailed description of the heuristic proposed by Bennett et al. (2000) for cardinality-constrained K-means clustering is provided in the appendix.

Notation: We denote by $\mathbf{1}$ the vector of all ones and by $\|\cdot\|$ the Euclidean norm. For symmetric square matrices $\mathbf{A}, \mathbf{B} \in \mathbb{S}^N$, the relation $\mathbf{A} \succeq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive semidefinite, while $\mathbf{A} \geq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is elementwise non-negative. The notation $\langle \mathbf{A}, \mathbf{B} \rangle = \mathrm{Tr}(\mathbf{AB})$ represents the trace inner product of $\mathbf{A}$ and $\mathbf{B}$. Furthermore, we use $\mathrm{diag}(\mathbf{A})$ to denote a vector in $\mathbb{R}^N$ whose entries coincide with those of $\mathbf{A}$'s main diagonal. Finally, for a set of N datapoints $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_N$, we use $\mathbf{D} \in \mathbb{S}^N$ to denote the matrix of squared pairwise distances $d_{ij} = \|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\|^2$.

3.2 Problem Formulation and Analysis

We first prove that the clustering problem (3.1) is an instance of a *quadratic assignment problem* and transform (3.1) to an MILP with NK binary variables. Then, we discuss the

complexity of (3.1) and show that an optimal clustering always corresponds to some Voronoi partition of $\mathbb{R}^d$.

Our first result relies on the following auxiliary lemma, which we state without proof.

Lemma 3.1. For any vectors $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n \in \mathbb{R}^d$, we have

$$\sum_{i=1}^{n} \|\boldsymbol{\xi}_i - \tfrac{1}{n}(\textstyle\sum_{j=1}^{n} \boldsymbol{\xi}_j)\|^2 = \dfrac{1}{2n} \displaystyle\sum_{i,j=1}^{n} \|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\|^2.$$

Proof. See Zha et al. (2002, p. 1060). □

Using Lemma 3.1, Costa et al. (2017) notice that the K-means objective can be stated as a sum of quadratic terms. In the following proposition, we elaborate on this insight and prove that problem (3.1) is a specific instance of a quadratic assignment problem.

Proposition 3.2 (Quadratic Assignment Reformulation)**.** The clustering problem (3.1) can be cast as the quadratic assignment problem

$$\underset{\sigma \in \mathfrak{S}^N}{\text{minimize}} \quad \tfrac{1}{2} \langle \mathbf{W}, \mathbf{P}_\sigma \mathbf{D} \mathbf{P}_\sigma^\top \rangle, \tag{3.2}$$

where $\mathbf{W} \in \mathbb{S}^N$ is a block diagonal matrix with blocks $\frac{1}{n_k}\mathbf{1}\mathbf{1}^\top \in \mathbb{S}^{n_k}$, $k = 1, \ldots, K$, $\mathfrak{S}^N$ is the set of permutations of $\{1, \ldots, N\}$, and $\mathbf{P}_\sigma$ is defined through $(\mathbf{P}_\sigma)_{ij} = 1$ if $\sigma(i) = j$; $(\mathbf{P}_\sigma)_{ij} = 0$ otherwise.

Proof. We show that for any feasible solution of (3.1) there exists a feasible solution of (3.2) which attains the same objective value and vice versa. To this end, for any partition $(I_1, \ldots, I_K)$ feasible in (3.1), consider any permutation $\sigma \in \mathfrak{S}^N$ that satisfies $\sigma(\{1 + \sum_{i=1}^{k-1} n_i, \ldots, \sum_{i=1}^{k} n_i\}) = I_k$ for all $k = 1, \ldots, K$, and denote its inverse by σ^{-1}. This permutation is feasible in (3.2), and it achieves the same objective value as $(I_1, \ldots, I_K)$ in (3.1) because

$$\sum_{k=1}^{K} \sum_{i \in I_k} \|\boldsymbol{\xi}_i - \tfrac{1}{n_k}(\textstyle\sum_{j \in I_k} \boldsymbol{\xi}_j)\|^2 = \dfrac{1}{2} \displaystyle\sum_{k=1}^{K} \dfrac{1}{n_k} \sum_{i,j \in I_k} d_{ij} = \dfrac{1}{2} \displaystyle\sum_{k=1}^{K} \dfrac{1}{n_k} \sum_{i,j \in \sigma^{-1}(I_k)} d_{\sigma(i)\sigma(j)} = \dfrac{1}{2} \langle \mathbf{W}, \mathbf{P}_\sigma \mathbf{D} \mathbf{P}_\sigma^\top \rangle,$$

where the first equality is implied by Lemma 3.1, the second equality is a consequence of the definition of σ, and the third equality follows from the definition of $\mathbf{W}$.

Conversely, for any $\sigma \in \mathfrak{S}^N$ feasible in (3.2), consider any partition $(I_1, \ldots, I_K)$ satisfying $I_k = \sigma(\{1 + \sum_{i=1}^{k-1} n_i, \ldots, \sum_{i=1}^{k} n_i\})$ for all $k = 1, \ldots, K$. This partition is feasible in (3.1), and a similar reasoning as before shows that the partition achieves the same objective value as σ in (3.2). □

Generic quadratic assignment problems with N facilities and N locations can be reformulated as MILPs with $\Omega(N^2)$ binary variables via the Kaufmann and Broeckx linearization; see, *e.g.*, Burkard (2013, p. 2741). The LP relaxations of these MILPs are, however, known to be weak, and

give a trivial lower bound of zero; see, *e.g.*, Zhang et al. (2013, Theorem 4.1). In Proposition 3.3 below we show that the intra-cluster permutation symmetry of the datapoints enables us to give an alternative MILP reformulation containing only $NK \ll \Omega(N^2)$ binary variables. We also mention that the related, yet different, cardinality-constrained exemplar-based clustering problem can be formulated as an MILP containing $\Omega(N^2)$ binary variables; see Mulvey and Beck (1984).

Proposition 3.3 (MILP Reformulation)**.** The clustering problem (3.1) is equivalent to the MILP

$$\begin{array}{lll} \textit{minimize} & \frac{1}{2}\sum_{k=1}^{K}\frac{1}{n_k}\sum_{i,j=1}^{N} d_{ij}\eta_{ij}^k & \\ \textit{subject to} & \pi_i^k \in \{0,1\},\ \eta_{ij}^k \in \mathbb{R}_+ & i,j=1,\ldots,N,\ k=1,\ldots,K \\ & \sum_{i=1}^{N}\pi_i^k = n_k & k=1,\ldots,K \\ & \sum_{k=1}^{K}\pi_i^k = 1 & i=1,\ldots,N \\ & \eta_{ij}^k \geq \pi_i^k + \pi_j^k - 1 & i,j=1,\ldots,N,\ k=1,\ldots,K. \end{array} \tag{$\mathscr{P}$}$$

The binary variable π_i^k in the MILP $\mathscr{P}$ satisfies $\pi_i^k = 1$ if $i \in I_k$; $\pi_i^k = 0$ otherwise. At optimality, $\eta_{ij}^k = \max\{\pi_i^k + \pi_j^k - 1, 0\}$ is equal to 1 if $i, j \in I_k$ (*i.e.*, $\pi_i^k = \pi_j^k = 1$) and 0 otherwise.

Proof of Proposition 3.3. At optimality, the decision variables η_{ij}^k in problem $\mathscr{P}$ take the values $\eta_{ij}^k = \max\{\pi_i^k + \pi_j^k - 1, 0\}$. Accordingly, problem $\mathscr{P}$ can equivalently be stated as

$$\begin{array}{lll} \text{minimize} & \frac{1}{2}\sum_{k=1}^{K}\frac{1}{n_k}\sum_{i,j=1}^{N} d_{ij}\max\{\pi_i^k + \pi_j^k - 1, 0\} & \\ \text{subject to} & \pi_i^k \in \{0,1\} & i=1,\ldots,N,\ k=1,\ldots,K \\ & \sum_{i=1}^{N}\pi_i^k = n_k & k=1,\ldots,K \\ & \sum_{k=1}^{K}\pi_i^k = 1 & i=1,\ldots,N. \end{array} \tag{$\mathscr{P}'$}$$

In the following, we show that any feasible solution of (3.1) gives rise to a feasible solution of $\mathscr{P}'$ with the same objective value and vice versa. To this end, consider first a partition $(I_1,\ldots,I_K)$ that is feasible in (3.1). Choosing $\pi_i^k = 1$ if $i \in I_k$ and $\pi_i^k = 0$ otherwise, $k = 1,\ldots,K$, is feasible in $\mathscr{P}'$ and attains the same objective value as $(I_1,\ldots,I_K)$ in (3.1) since

$$\sum_{k=1}^{K}\sum_{i\in I_k}\|\boldsymbol{\xi}_i - \tfrac{1}{n_k}(\textstyle\sum_{j\in I_k}\boldsymbol{\xi}_j)\|^2 = \frac{1}{2}\sum_{k=1}^{K}\frac{1}{n_k}\sum_{i,j\in I_k} d_{ij} = \frac{1}{2}\sum_{k=1}^{K}\frac{1}{n_k}\sum_{i,j=1}^{N} d_{ij}\max\{\pi_i^k + \pi_j^k - 1, 0\}.$$

Here, the first equality is implied by Lemma 3.1, and the second equality follows from the construction of π_i^k. By the same argument, every π_i^k feasible in $\mathscr{P}'$ gives rise to a partition $(I_1,\ldots,I_K)$, $I_k = \{i : \pi_i^k = 1\}$ for $k = 1,\ldots,K$, that is feasible in $\mathscr{P}'$ and that attains the same objective value. □

Remark 3.4. Note that zero is a (trivial) lower bound on the objective value of the LP relaxation

of the MILP $\mathscr{P}$. As a consequence, this LP relaxation is at least as tight as the LP relaxation of the Kaufmann and Broeckx exact MILP formulation of problem (3.2), which always yields the lower bound of zero. It is also possible to construct instances where the LP relaxation of the MILP $\mathscr{P}$ is strictly tighter.

K-means clustering with cardinality constraints is known to be NP-hard as it is a special case of cardinality-constrained p-norm clustering, which was shown to be NP-hard (for any $p > 1$) by Bertoni et al. (2012). The restriction to the Euclidean norm (*i.e.*, $p = 2$), however, allows for a more concise proof, which is given in the following proposition.

Proposition 3.5. *K-means clustering with cardinality constraints is NP-hard even for $K = 2$. Hence, unless $P = NP$, there is no polynomial time algorithm for solving problem (3.1).*

Proof. In analogy to Proposition 3.3, one can show that the unconstrained K-means clustering problem can be formulated as a variant of problem $\mathscr{P}$ that omits the first set of assignment constraints, which require that $\sum_{i=1}^{N} \pi_i^k = n_k$ for all k= $1,\dots,K$, and replaces the (now unconstrained) cardinality n_k in the objective function by the size of I_k, which can be expressed as $\sum_{i=1}^{N} \pi_i^k$. If $K = 2$, we can thus solve the unconstrained K-means clustering problem by solving problem $\mathscr{P}$ for all cluster cardinality combinations $(n_1, n_2) \in \{(1, N-1), (2, N-2), \dots, (\lfloor N/2 \rfloor, \lceil N/2 \rceil)\}$ and selecting the clustering with the lowest objective value. Thus, in this case, if problem $\mathscr{P}$ were polynomial-time solvable, then so would be the unconstrained K-means clustering problem. This, however, would contradict Theorem 1 in Aloise et al. (2009), which shows that the unconstrained K-means clustering problem is NP-hard even for $K = 2$ clusters. □

In the context of balanced clustering, similar hardness results have been established by Pyatkin et al. (2017). Specifically, they prove that the balanced K-means clustering problem is NP-complete for $K \geq 2$ and $\frac{N}{K} \geq 3$ (*i.e.*, the shared cardinality of all clusters is greater than or equal to three). In contrast, if $K \geq 2$ and $\frac{N}{K} = 2$ (*i.e.*, each cluster should contain two points), balanced K-means clustering reduces to a minimum-weight perfect matching problem that can be solved in polynomial-time by different algorithms; see Cook and Rohe (1999, Table I) for a review.

In K-means clustering *without* cardinality constraints, the convex hulls of the optimal clusters do not overlap, and thus each cluster fits within a separate cell of a Voronoi partition of $\mathbb{R}^d$, see *e.g.*, Hasegawa et al. (1993, Theorem 2.1). We demonstrate below that this property is preserved in the presence of cardinality constraints.

Theorem 3.6 (Voronoi Partition)**.** *For every optimal solution to problem (3.1), there exists a Voronoi partition of $\mathbb{R}^d$ such that each cluster is contained in exactly one Voronoi cell.*

Proof. We show that for every optimal clustering $(I_1, \dots, I_K)$ of (3.1) and every $k, \ell \in \{1, \dots, K\}$, $k < \ell$, there exists a hyperplane separating the points in I_k from those in I_ℓ. This in turn

implies the existence of the desired Voronoi partition. Given a cluster I_m for any $m \in \{1,\dots,K\}$, define its cluster center as $\boldsymbol{\zeta}_m = \frac{1}{n_m}\sum_{i\in I_m}\boldsymbol{\xi}_i$, and let $\boldsymbol{h} = \boldsymbol{\zeta}_k - \boldsymbol{\zeta}_\ell$ be the vector that connects the cluster centers of I_k and I_ℓ. The statement holds if $\boldsymbol{h}^\top(\boldsymbol{\xi}_{i_k} - \boldsymbol{\xi}_{i_\ell}) \geq 0$ for all $i_k \in I_k$ and $i_\ell \in I_\ell$ as $\boldsymbol{h}$ itself determines a separating hyperplane for I_k and I_ℓ in that case. We thus assume that $\boldsymbol{h}^\top(\boldsymbol{\xi}_{i_k} - \boldsymbol{\xi}_{i_\ell}) < 0$ for some $i_k \in I_k$ and $i_\ell \in I_\ell$. However, this contradicts the optimality of the clustering $(I_1,\dots,I_K)$ because

$$\begin{aligned}\boldsymbol{h}^\top(\boldsymbol{\xi}_{i_k} - \boldsymbol{\xi}_{i_\ell}) < 0 &\iff (\boldsymbol{\zeta}_k - \boldsymbol{\zeta}_\ell)^\top(\boldsymbol{\xi}_{i_k} - \boldsymbol{\xi}_{i_\ell}) < 0\\ &\iff \boldsymbol{\xi}_{i_k}^\top\boldsymbol{\zeta}_k + \boldsymbol{\xi}_{i_\ell}^\top\boldsymbol{\zeta}_\ell < \boldsymbol{\xi}_{i_k}^\top\boldsymbol{\zeta}_\ell + \boldsymbol{\xi}_{i_\ell}^\top\boldsymbol{\zeta}_k\\ &\iff \|\boldsymbol{\xi}_{i_\ell} - \boldsymbol{\zeta}_k\|^2 + \|\boldsymbol{\xi}_{i_k} - \boldsymbol{\zeta}_\ell\|^2 < \|\boldsymbol{\xi}_{i_k} - \boldsymbol{\zeta}_k\|^2 + \|\boldsymbol{\xi}_{i_\ell} - \boldsymbol{\zeta}_\ell\|^2,\end{aligned}$$

where the last equivalence follows from multiplying both sides of the second inequality with 2 and then completing the squares by adding $\boldsymbol{\xi}_{i_k}^\top\boldsymbol{\xi}_{i_k} + \boldsymbol{\zeta}_k^\top\boldsymbol{\zeta}_k + \boldsymbol{\xi}_{i_\ell}^\top\boldsymbol{\xi}_{i_\ell} + \boldsymbol{\zeta}_\ell^\top\boldsymbol{\zeta}_\ell$ on both sides. Defining $\tilde{I}_k = I_k \cup \{i_\ell\} \setminus \{i_k\}$ and $\tilde{I}_\ell = I_\ell \cup \{i_k\} \setminus \{i_\ell\}$, the above would imply that

$$\begin{aligned}&\sum_{i\in\tilde{I}_k}\|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_k\|^2 + \sum_{i\in\tilde{I}_\ell}\|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_\ell\|^2 + \sum_{\substack{m=1,\dots,K\\ m\notin\{k,\ell\}}}\sum_{i\in I_m}\|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_m\|^2\\ &\quad< \sum_{i\in I_k}\|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_k\|^2 + \sum_{i\in I_\ell}\|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_\ell\|^2 + \sum_{\substack{m=1,\dots,K\\ m\notin\{k,\ell\}}}\sum_{i\in I_m}\|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_m\|^2.\end{aligned}$$

The left-hand side of the above inequality represents an upper bound on the sum of squared intra-cluster distances attained by the clustering $(I_1,\dots,\tilde{I}_k,\dots,\tilde{I}_\ell,\dots,I_K)$ since $\boldsymbol{\zeta}_k$ and $\boldsymbol{\zeta}_\ell$ may not coincide with the minimizers $\frac{1}{n_k}\sum_{i\in\tilde{I}_k}\boldsymbol{\xi}_i$ and $\frac{1}{n_\ell}\sum_{i\in\tilde{I}_\ell}\boldsymbol{\xi}_i$, respectively. Recall that the cluster centers are chosen so as to minimize the sum of the distances from the cluster center to each point in the cluster. We thus conclude that the clustering $(I_1,\dots,\tilde{I}_k,\dots,\tilde{I}_\ell,\dots,I_K)$ attains a strictly lower objective value than $(I_1,\dots,I_K)$ in problem (3.1), which is a contradiction. □

3.3 Cardinality-Constrained Clustering without Outliers

We now relax the intractable MILP $\mathscr{P}$ to tractable conic programs that yield efficiently computable lower and upper bounds on $\mathscr{P}$.

3.3.1 Convex Relaxations and Rounding Algorithm

We first eliminate the η_{ij}^k variables from $\mathscr{P}$ by re-expressing the problem's objective function as

$$\frac{1}{2}\sum_{k=1}^K\frac{1}{n_k}\sum_{i,j=1}^N d_{ij}\eta_{ij}^k = \frac{1}{2}\sum_{k=1}^K\frac{1}{n_k}\sum_{i,j=1}^N d_{ij}\max\{\pi_i^k + \pi_j^k - 1, 0\} = \frac{1}{2}\sum_{k=1}^K\frac{1}{n_k}\sum_{i,j=1}^N d_{ij}\pi_i^k\pi_j^k,$$

where the last equality holds because the variables π_i^k are binary. Next, we apply the variable transformation $x_i^k \leftarrow 2\pi_i^k - 1$, whereby $\mathscr{P}$ simplifies to

$$\begin{array}{lll} \text{minimize} & \frac{1}{8}\sum_{k=1}^K \frac{1}{n_k}\sum_{i,j=1}^N d_{ij}(1+x_i^k)(1+x_j^k) & \\ \text{subject to} & x_i^k \in \{-1,+1\} & i=1,\ldots,N,\ k=1,\ldots,K \\ & \sum_{i=1}^N x_i^k = 2n_k - N & k=1,\ldots,K \\ & \sum_{k=1}^K x_i^k = 2-K & i=1,\ldots,N. \end{array} \tag{3.3}$$

Here, x_i^k takes the value $+1$ if the i-th datapoint is assigned to cluster k and -1 otherwise. Note that the constraints in (3.3) are indeed equivalent to the first two constraints in $\mathscr{P}$, respectively. In Theorem 3.7 below we will show that the reformulation (3.3) of the MILP $\mathscr{P}$ admits the SDP relaxation

$$\begin{array}{lll} \text{minimize} & \frac{1}{8}\left\langle \mathbf{D}, \sum_{k=1}^K \frac{1}{n_k}\left(\mathbf{M}^k + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^k\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^k)^\top\right)\right\rangle & \\ \text{subject to} & (\boldsymbol{x}^k, \mathbf{M}^k) \in \mathscr{C}_{\text{SDP}}(n_k) & k=1,\ldots,K \\ & \sum_{k=1}^K \boldsymbol{x}^k = (2-K)\mathbf{1}, & \end{array} \tag{$\mathscr{R}_{\text{SDP}}$}$$

where, for any $n \in \mathbb{N}$, the convex set $\mathscr{C}_{\text{SDP}}(n) \subset \mathbb{R}^N \times \mathbb{S}^N$ is defined as

$$\mathscr{C}_{\text{SDP}}(n) = \left\{ (\boldsymbol{x}, \mathbf{M}) \in \mathbb{R}^N \times \mathbb{S}^N : \begin{array}{l} \mathbf{1}^\top \boldsymbol{x} = 2n - N,\ \mathbf{M}\mathbf{1} = (2n-N)\boldsymbol{x} \\ \operatorname{diag}(\mathbf{M}) = \mathbf{1},\ \mathbf{M} \succeq \boldsymbol{x}\boldsymbol{x}^\top \\ \mathbf{M} + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}\mathbf{1}^\top + \mathbf{1}\boldsymbol{x}^\top \geq \mathbf{0} \\ \mathbf{M} + \mathbf{1}\mathbf{1}^\top - \boldsymbol{x}\mathbf{1}^\top - \mathbf{1}\boldsymbol{x}^\top \geq \mathbf{0} \\ \mathbf{M} - \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}\mathbf{1}^\top - \mathbf{1}\boldsymbol{x}^\top \leq \mathbf{0} \\ \mathbf{M} - \mathbf{1}\mathbf{1}^\top - \boldsymbol{x}\mathbf{1}^\top + \mathbf{1}\boldsymbol{x}^\top \leq \mathbf{0} \end{array} \right\}.$$

Note that $\mathscr{C}_{\text{SDP}}(n)$ is semidefinite representable because Schur's complement allows us to express the constraint $\mathbf{M} \succeq \boldsymbol{x}\boldsymbol{x}^\top$ as a linear matrix inequality; see, *e.g.*, Boyd and Vandenberghe (2004). Furthermore, we point out that the last four constraints in $\mathscr{C}_{\text{SDP}}(n)$ are also used in the *reformulation-linearization technique* for nonconvex programs, as described by Anstreicher (2009).

We can further relax the above SDP to an LP, henceforth denoted by $\mathscr{R}_{\text{LP}}$, where the constraints $(\boldsymbol{x}^k, \mathbf{M}^k) \in \mathscr{C}_{\text{SDP}}(n_k)$ are replaced with $(\boldsymbol{x}^k, \mathbf{M}^k) \in \mathscr{C}_{\text{LP}}(n_k)$, and where, for any $n \in \mathbb{N}$, the polytope $\mathscr{C}_{\text{LP}}(n)$ is obtained by removing the non-linear constraint $\mathbf{M} \succeq \boldsymbol{x}\boldsymbol{x}^\top$ from $\mathscr{C}_{\text{SDP}}(n)$.

Theorem 3.7 (SDP and LP Relaxations)**.** We have $\min \mathscr{R}_{LP} \leq \min \mathscr{R}_{SDP} \leq \min \mathscr{P}$.

Proof. The inequality $\min \mathscr{R}_{\text{LP}} \leq \min \mathscr{R}_{\text{SDP}}$ is trivially satisfied because $\mathscr{C}_{\text{SDP}}(n)$ is constructed as a subset of $\mathscr{C}_{\text{LP}}(n)$ for every $n \in \mathbb{N}$. To prove the inequality $\min \mathscr{R}_{\text{SDP}} \leq \min \mathscr{P}$, consider any set of binary vectors $\{\boldsymbol{x}^k\}_{k=1}^K$ feasible in (3.3) and define $\mathbf{M}^k = \boldsymbol{x}^k(\boldsymbol{x}^k)^\top$ for $k = 1,\ldots,K$. By construction, the objective value of $\{\boldsymbol{x}^k\}_{k=1}^K$ in (3.3) coincides with that of $\{(\boldsymbol{x}^k, \mathbf{M}^k)\}_{k=1}^K$

in $\mathscr{R}_{\text{SDP}}$. Moreover, the constraints in (3.3) imply that

$$\mathbf{M}^k\mathbf{1} = \boldsymbol{x}^k(\boldsymbol{x}^k)^\top\mathbf{1} = (2n_k - N)\boldsymbol{x}^k, \quad \operatorname{diag}(\mathbf{M}^k) = \mathbf{1}, \quad \mathbf{M}^k \succeq \boldsymbol{x}^k(\boldsymbol{x}^k)^\top$$

and

$$\begin{aligned}
\mathbf{M}^k + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^k\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^k)^\top &= +(\mathbf{1}+\boldsymbol{x}^k)(\mathbf{1}+\boldsymbol{x}^k)^\top \succeq \mathbf{0}\\
\mathbf{M}^k + \mathbf{1}\mathbf{1}^\top - \boldsymbol{x}^k\mathbf{1}^\top - \mathbf{1}(\boldsymbol{x}^k)^\top &= +(\mathbf{1}-\boldsymbol{x}^k)(\mathbf{1}-\boldsymbol{x}^k)^\top \succeq \mathbf{0}\\
\mathbf{M}^k - \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^k\mathbf{1}^\top - \mathbf{1}(\boldsymbol{x}^k)^\top &= -(\mathbf{1}-\boldsymbol{x}^k)(\mathbf{1}+\boldsymbol{x}^k)^\top \preceq \mathbf{0}\\
\mathbf{M}^k - \mathbf{1}\mathbf{1}^\top - \boldsymbol{x}^k\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^k)^\top &= -(\mathbf{1}+\boldsymbol{x}^k)(\mathbf{1}-\boldsymbol{x}^k)^\top \preceq \mathbf{0},
\end{aligned}$$

which ensures that $(\boldsymbol{x}^k, \mathbf{M}^k) \in \mathscr{C}_{\text{SDP}}(n_k)$ for every k. Finally, the constraint $\sum_{k=1}^K \boldsymbol{x}^k = (2-K)\mathbf{1}$ in $\mathscr{R}_{\text{SDP}}$ coincides with the last constraint in (3.3). Thus, $\{(\boldsymbol{x}^k, \mathbf{M}^k)\}_{k=1}^K$ is feasible in $\mathscr{R}_{\text{SDP}}$. The desired inequality now follows because any feasible point in (3.3) corresponds to a feasible point in $\mathscr{R}_{\text{SDP}}$ with the same objective value. Note that the converse implication is generally false. □

Remark 3.8. In the special case when $K = 2$, we can half the number of variables in $\mathscr{R}_{SDP}$ and $\mathscr{R}_{\text{LP}}$ by setting $\boldsymbol{x}^2 = -\boldsymbol{x}^1$ and $\mathbf{M}^2 = \mathbf{M}^1$ without loss of generality.

It is possible to show that $\mathscr{R}_{\text{LP}}$ is at least as tight as the naïve LP relaxation $\mathscr{L}$ of the MILP $\mathscr{P}$, where the integrality constraints are simply ignored. One can also construct instances where $\mathscr{R}_{\text{LP}}$ is strictly tighter than $\mathscr{L}$. We also emphasize that both LP relaxations entail $\mathscr{O}(N^2K)$ variables and $\mathscr{O}(N^2K)$ constraints.

Proposition 3.9. We have $\min \mathscr{R}_{\text{LP}} \geq \min \mathscr{L}$.

Proof. Consider a feasible solution $\{(\boldsymbol{x}^k, \mathbf{M}^k)\}_{k=1}^K$ of $\mathscr{R}_{\text{LP}}$. Its feasibility implies that

$$\text{(a)}\ \sum_{k=1}^K x_i^k = 2-K\ \forall i, \qquad \text{(b)}\ \sum_{i=1}^N x_i^k = 2n_k - N\ \forall k, \qquad \text{(c)}\ m_{ij}^k - x_i^k - x_j^k + 1 \geq 0\ \forall i,j,k.$$

Next, set $\pi_i^k = (x_i^k + 1)/2$ and $\eta_{ij}^k = \frac{1}{4}(m_{ij}^k + x_i^k + x_j^k + 1)$ for all i, j, k. Then,

$$\text{(a')}\ \sum_{k=1}^K \pi_i^k = 1\ \forall i, \qquad \text{(b')}\ \sum_{i=1}^N \pi_i^k = n_k\ \forall k, \qquad \text{(c')}\ \eta_{ij}^k \geq \pi_i^k + \pi_j^k - 1\ \forall i,j,k.$$

Hence, this solution is feasible in $\mathscr{L}$. A direct calculation also reveals that both solutions attain the same objective value in their respective optimization problems. This confirms that $\mathscr{R}_{\text{LP}}$ is a relaxation that is at least as tight as $\mathscr{L}$. □

Next, we develop a rounding algorithm that recovers a feasible clustering (and thus an upper bound on $\mathscr{P}$) from an optimal solution of the relaxed problem $\mathscr{R}_{\text{SDP}}$ or $\mathscr{R}_{\text{LP}}$; see Algorithm 1.

Recall that the continuous variables $\boldsymbol{x}^k = (x_1^k, \ldots, x_N^k)^\top$ in $\mathscr{R}_{\text{SDP}}$ and $\mathscr{R}_{\text{LP}}$ correspond to the binary variables in (3.3) with identical names. This correspondence motivates us to solve a

Algorithm 1 Rounding algorithm for cardinality-constrained clustering

1: **Input:** $\mathscr{I}_1 = \{1,\dots,N\}$ (data indices), $n_k \in \mathbb{N},\ k = 1,\dots,K$ (cluster sizes).
2: Solve $\mathscr{R}_{\mathrm{SDP}}$ or $\mathscr{R}_{\mathrm{LP}}$ for the datapoints $\boldsymbol{\xi}_i$, $i \in \mathscr{I}_1$, and record the optimal $\boldsymbol{x}^1,\dots,\boldsymbol{x}^K \in \mathbb{R}^N$.
3: Solve the linear assignment problem

$$\Pi' \in \operatorname*{argmax}_{\Pi} \left\{ \sum_{i=1}^N \sum_{k=1}^K \pi_i^k x_i^k :\ \pi_i^k \in \{0,1\},\ \sum_{i=1}^N \pi_i^k = n_k\ \ \forall k,\ \sum_{k=1}^K \pi_i^k = 1\ \ \forall i \right\}.$$

4: Set $I'_k \leftarrow \{i : (\pi')_i^k = 1\}$ for all $k = 1,\dots,K$.
5: Set $\boldsymbol{\zeta}_k \leftarrow \frac{1}{n_k} \sum_{i \in I'_k} \boldsymbol{\xi}_i$ for all $k = 1,\dots,K$.
6: Solve the linear assignment problem

$$\Pi^\star \in \operatorname*{argmin}_{\Pi} \left\{ \sum_{i=1}^N \sum_{k=1}^K \pi_i^k \|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_k\|^2 :\ \pi_i^k \in \{0,1\},\ \sum_{i=1}^N \pi_i^k = n_k\ \ \forall k,\ \sum_{k=1}^K \pi_i^k = 1\ \ \forall i \right\}.$$

7: Set $I_k \leftarrow \{i : (\pi^\star)_i^k = 1\}$ for all $k = 1,\dots,K$.
8: **Output:** $I_1,\dots,I_K$.

linear assignment problem in Step 3 of Algorithm 1, which seeks a matrix $\mathbf{\Pi} \in \{0,1\}^{N\times K}$ with $\pi_i^k \approx \frac{1}{2}(x_i^k + 1)$ for all i and k subject to the prescribed cardinality constraints. Note that even though this assignment problem constitutes an MILP, it can be solved in polynomial time because its constraint matrix is totally unimodular, implying that its LP relaxation is exact. Alternatively, one may solve the assignment problem using the Hungarian algorithm; see, *e.g.*, Burkard et al. (2009).

Note that Steps 5–7 of Algorithm 1 are reminiscent of a *single* iteration of Lloyd's algorithm for cardinality-constrained K-means clustering as described by Bennett et al. (2000). Specifically, Step 5 calculates the cluster centers $\boldsymbol{\zeta}_k$, while Steps 6 and 7 reassign each point to the nearest center while adhering to the cardinality constraints. Algorithm 1 thus follows just one step of Lloyd's algorithm initialized with an optimizer of $\mathscr{R}_{\mathrm{SDP}}$ or $\mathscr{R}_{\mathrm{LP}}$. This refinement step ensures that the output clustering is compatible with a Voronoi partition of $\mathbb{R}^d$, which is desirable in view of Theorem 3.6.

3.3.2 Tighter Relaxations for Balanced Clustering

The computational burden of solving $\mathscr{R}_{\mathrm{SDP}}$ and $\mathscr{R}_{\mathrm{LP}}$ grows with K. We show in this section that if all clusters share the same size n (*i.e.*, $n_k = n$ for all k), then $\mathscr{R}_{\mathrm{SDP}}$ can be replaced by

$$\begin{array}{ll} \text{minimize} & \frac{1}{8n} \langle \mathbf{D}, \mathbf{M}^1 + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^1\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^1)^\top + (K-1)\left(\mathbf{M} + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}\mathbf{1}^\top + \mathbf{1}\boldsymbol{x}^\top\right)\rangle \\ \text{subject to} & (\boldsymbol{x}^1, \mathbf{M}^1), (\boldsymbol{x}, \mathbf{M}) \in \mathscr{C}_{\mathrm{SDP}}(n), \quad \boldsymbol{x}^1 + (K-1)\boldsymbol{x} = (2-K)\mathbf{1}, \quad x_1^1 = 1, \end{array} \qquad (\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}})$$

whose size no longer scales with K. Similarly, $\mathscr{R}_{\mathrm{LP}}$ simplifies to the LP $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$ obtained from $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ by replacing $\mathscr{C}_{\mathrm{SDP}}(n)$ with $\mathscr{C}_{\mathrm{LP}}(n)$. This is a manifestation of how symmetry can be

exploited to simplify convex programs, a phenomenon which is studied in a more general setting by Gatermann and Parrilo (2004).

Corollary 3.10 (Relaxations for Balanced Clustering)**.** We have $\min \mathscr{R}^b_{LP} \leq \min \mathscr{R}^b_{SDP} \leq \min \mathscr{P}$.

Proof. The inequality $\min \mathscr{R}^{\mathrm{b}}_{\mathrm{LP}} \leq \min \mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ is trivially satisfied. To prove the inequality $\min \mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}} \leq \min \mathscr{P}$, we first add the symmetry breaking constraint $x^1_1 = 1$ to the MILP $\mathscr{P}$. Note that this constraint does not increase the optimal value of $\mathscr{P}$. It just requires that the cluster containing the datapoint $\boldsymbol{\xi}_1$ should be assigned the number $k = 1$. This choice is unrestrictive because all clusters have the same size. By repeating the reasoning that led to Theorem 3.7, the MILP $\mathscr{P}$ can then be relaxed to a variant of the SDP $\mathscr{R}_{\mathrm{SDP}}$ that includes the (linear) symmetry breaking constraint $x^1_1 = 1$. Note that the constraints and the objective function of the resulting SDP are invariant under permutations of the cluster indices $k = 2, \ldots, K$ because $n_k = n$ for all k. Note also that the constraints are not invariant under permutations involving $k = 1$ due to the symmetry breaking constraint. Next, consider any feasible solution $\{(\boldsymbol{x}^k, \mathbf{M}^k)\}^K_{k=1}$ of this SDP, and define

$$\boldsymbol{x} = \frac{1}{K-1} \sum_{k=2}^{K} \boldsymbol{x}^k \quad \text{and} \quad \mathbf{M} = \frac{1}{K-1} \sum_{k=2}^{K} \mathbf{M}^k.$$

Moreover, construct a permutation-symmetric solution $\{(\boldsymbol{x}^k_{\mathrm{s}}, \mathbf{M}^k_{\mathrm{s}})\}^K_{k=1}$ by setting

$$\begin{aligned} &\boldsymbol{x}^1_{\mathrm{s}} = \boldsymbol{x}^1, \quad \boldsymbol{x}^k_{\mathrm{s}} = \boldsymbol{x} \quad \forall k = 2, \ldots, K, \\ &\mathbf{M}^1_{\mathrm{s}} = \mathbf{M}^1, \quad \mathbf{M}^k_{\mathrm{s}} = \mathbf{M} \quad \forall k = 2, \ldots, K. \end{aligned}$$

By the convexity and permutation symmetry of the SDP, the symmetrized solution $\{(\boldsymbol{x}^k_{\mathrm{s}}, \mathbf{M}^k_{\mathrm{s}})\}^K_{k=1}$ is also feasible in the SDP and attains the same objective value as $\{(\boldsymbol{x}^k, \mathbf{M}^k)\}^K_{k=1}$. Moreover, as the choice of $\{(\boldsymbol{x}^k, \mathbf{M}^k)\}^K_{k=1}$ was arbitrary, we may indeed restrict attention to symmetrized solutions with $\boldsymbol{x}^k = \boldsymbol{x}^\ell$ and $\mathbf{M}^k = \mathbf{M}^\ell$ for all $k, \ell \in \{2, \ldots, K\}$ without increasing the objective value of the SDP. Therefore, the simplified SDP relaxation $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ provides a lower bound on $\mathscr{P}$. □

If $n_k = n$ for all k, then the SDP and LP relaxations from Section 3.3.1 admit an optimal solution where both $\boldsymbol{x}^k$ and $\mathbf{M}^k$ are independent of k, in which case Algorithm 1 performs poorly. This motivates the improved relaxations $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ and $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$ involving the symmetry breaking constraint $x^1_1 = 1$, which ensures that—without loss of generality—the cluster harboring the first datapoint $\boldsymbol{\xi}_1$ is indexed by $k = 1$. As the symmetry between clusters $2, \ldots, K$ persists and because any additional symmetry breaking constraint would be restrictive, the optimal solutions of $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ and $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$ only facilitate a reliable recovery of cluster 1. To recover *all* clusters, however, we can solve $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ or $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$ $K-1$ times over the yet unassigned datapoints, see Algorithm 2. The resulting clustering could be improved by appending one iteration of Lloyd's algorithm (akin to Steps 5–7 in Algorithm 1).

In contrast, the naïve relaxation $\mathscr{L}$ of $\mathscr{P}$ becomes significantly weaker when all cardinalities

are equal. To see this, we note that a solution $\pi_i^k = 1/K$ and $\eta_{ij}^k = 0$ for all $i, j = 1, \ldots, N$ and for all $k = 1, \ldots, K$ is feasible in $\mathscr{L}$ (*i.e.*, it satisfies all constraints in problem $\mathscr{P}$ except the integrality constraints which are imposed on π_i^k) whenever $K \geq 2$. Hence, the optimal objective value of $\mathscr{L}$ is zero. This could be avoided by adding a symmetry breaking constraint $\pi_1^1 = 1$ to problem $\mathscr{L}$ to ensure that the cluster containing the first datapoint $\boldsymbol{\xi}_1$ is indexed by $k = 1$. However, the improvement appears to be marginal.

Algorithm 2 Rounding algorithm for balanced clustering

Input: $\mathscr{I}_1 = \{1, \ldots, N\}$ (data indices), $n \in \mathbb{N}$ (cluster size), $K = N/n \in \mathbb{N}$ (# clusters).
for $k = 1, \ldots, K-1$ **do**
  Solve $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ or $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$ for the datapoints $\boldsymbol{\xi}_i$, $i \in \mathscr{I}_k$, and record the optimal $\boldsymbol{x}^1 \in \mathbb{R}^{|\mathscr{I}_k|}$.
  Determine a bijection $\rho : \{1, \ldots, |\mathscr{I}_k|\} \to \mathscr{I}_k$ such that $x^1_{\rho(1)} \geq x^1_{\rho(2)} \geq \cdots \geq x^1_{\rho(|\mathscr{I}_k|)}$.
  Set $I_k \leftarrow \{\rho(1), \ldots, \rho(n)\}$ and $\mathscr{I}_{k+1} \leftarrow \mathscr{I}_k \backslash I_k$.
Set $I_K \leftarrow \mathscr{I}_K$.
Output: $I_1, \ldots, I_K$.

3.3.3 Comparison to existing SDP Relaxations

We now compare $\mathscr{R}_{\mathrm{SDP}}$ and $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ with existing SDP relaxations from the literature. First, we report the various SDP relaxations proposed by Peng and Wei (2007) and Awasthi et al. (2015). Then, we establish that two of them are equivalent. Finally, we show that $\mathscr{R}_{\mathrm{SDP}}$ and $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ are relaxations that are at least as tight as their corresponding counterparts from the literature. The numerical experiments in Section 3.5 provide evidence that this relation can also be strict.

Peng and Wei (2007) suggest two different SDP relaxations for the *un*constrained K-means clustering problem and an SDP relaxation for the balanced K-means clustering problem. All of them involve a Gram matrix $\mathbf{W} \in \mathbb{S}^N$ with entries $w_{ij} = \boldsymbol{\xi}_i^\top \boldsymbol{\xi}_j$. Their stronger relaxation for the unconstrained K-means clustering problem takes the form

$$\begin{array}{ll} \text{minimize} & \langle \mathbf{W}, \mathbb{I} - \mathbf{Z} \rangle \\ \text{subject to} & \mathbf{Z} \in \mathbb{S}^N \\ & \mathbf{Z} \geq \mathbf{0},\ \mathbf{Z} \geq \mathbf{0},\ \mathbf{Z}\mathbf{1} = \mathbf{1},\ \mathrm{Tr}(\mathbf{Z}) = K, \end{array} \qquad (\mathscr{PW}_1)$$

where $\mathbb{I}$ denotes the identity matrix of dimension N. Note that the constraints $\mathbf{Z} \geq \mathbf{0}$ and $\mathbf{Z}\mathbf{1} = \mathbf{1}$ ensure that $\mathbf{Z}$ is a stochastic matrix, and hence all of its eigenvalues lie between 0 and 1. Thus, further relaxing the non-negativity constraints leads to the following weaker relaxation,

$$\begin{array}{ll} \text{minimize} & \langle \mathbf{W}, \mathbb{I} - \mathbf{Z} \rangle \\ \text{subject to} & \mathbf{Z} \in \mathbb{S}^N \\ & \mathbb{I} \geq \mathbf{Z} \geq \mathbf{0},\ \mathbf{Z}\mathbf{1} = \mathbf{1},\ \mathrm{Tr}(\mathbf{Z}) = K. \end{array} \qquad (\mathscr{PW}_2)$$

Peng and Wei (2007) also demonstrate that $\mathscr{PW}_2$ essentially reduces to an eigenvalue problem, which implies that one can solve $\mathscr{PW}_2$ in $\mathscr{O}(KN^2)$ time; see Golub and Loan (1996). Their SDP

relaxation for the *balanced* K-means clustering problem is similar to $\mathscr{PW}_1$ and takes the form

$$\begin{array}{ll} \text{minimize} & \langle \mathbf{W}, \mathbb{I} - \mathbf{Z} \rangle \\ \text{subject to} & \mathbf{Z} \in \mathbb{S}^N \\ & \mathbf{Z} \geq \mathbf{0},\ \mathbf{0} \leq \mathbf{Z} \leq (K/N)\mathbf{1}\mathbf{1}^\top,\ \mathbf{Z}\mathbf{1} = \mathbf{1},\ \mathrm{Tr}(\mathbf{Z}) = K. \end{array} \tag{$\mathscr{PW}_1^{\mathrm{b}}$}$$

Awasthi et al. (2015) suggest another SDP relaxation for the unconstrained K-means clustering problem, based on the same matrix of squared pairwise distances $\mathbf{D}$ considered in this paper,

$$\begin{array}{ll} \text{minimize} & \langle \mathbf{D}, \mathbf{Z} \rangle \\ \text{subject to} & \mathbf{Z} \in \mathbb{S}^N \\ & \mathbf{Z} \geq \mathbf{0},\ \mathbf{Z} \geq \mathbf{0},\ \mathbf{Z}\mathbf{1} = \mathbf{1},\ \mathrm{Tr}(\mathbf{Z}) = K. \end{array} \tag{$\mathscr{A}$}$$

The following observation asserts that the stronger relaxation $\mathscr{PW}_1$ of Peng and Wei (2007) and the relaxation $\mathscr{A}$ of Awasthi et al. (2015) are actually equivalent.

Observation 1 *The problems $\mathscr{PW}_1$ and $\mathscr{A}$ are equivalent.*

Proof. Begin by expressing the objective of $\mathscr{PW}_1$ in terms of the pairwise distance matrix $\mathbf{D}$,

$$\begin{aligned} \langle \mathbf{W}, \mathbb{I} - \mathbf{Z} \rangle &= \frac{1}{2}\Big[2\langle \mathbf{W}, \mathbb{I} \rangle - \langle 2\mathbf{W}, \mathbf{Z} \rangle - \langle \mathbf{D}, \mathbf{Z} \rangle\Big] + \frac{1}{2}\langle \mathbf{D}, \mathbf{Z} \rangle \\ &= \frac{1}{2}\Big[2\langle \mathbf{W}, \mathbb{I} \rangle - \langle 2\mathbf{W} + \mathbf{D}, \mathbf{Z} \rangle\Big] + \frac{1}{2}\langle \mathbf{D}, \mathbf{Z} \rangle \\ &\overset{\text{(a)}}{=} \frac{1}{2}\Big[2\langle \mathbf{W}, \mathbb{I} \rangle - \langle \mathbf{1}\,\mathrm{diag}(\mathbf{W})^\top + \mathrm{diag}(\mathbf{W})\,\mathbf{1}^\top, \mathbf{Z} \rangle\Big] + \frac{1}{2}\langle \mathbf{D}, \mathbf{Z} \rangle \\ &= \frac{1}{2}\Big[2\langle \mathbf{W}, \mathbb{I} \rangle - \langle \mathbf{1}\,\mathrm{diag}(\mathbf{W})^\top, \mathbf{Z} \rangle - \langle \mathrm{diag}(\mathbf{W})\,\mathbf{1}^\top, \mathbf{Z} \rangle\Big] + \frac{1}{2}\langle \mathbf{D}, \mathbf{Z} \rangle \\ &= \frac{1}{2}\Big[2\mathrm{Tr}\big(\mathbf{W}\big) - \mathrm{Tr}\big(\mathbf{Z}\mathbf{1}\,\mathrm{diag}(\mathbf{W})^\top\big) - \mathrm{Tr}\big(\mathrm{diag}(\mathbf{W})\,\mathbf{1}^\top\mathbf{Z}\big)\Big] + \frac{1}{2}\langle \mathbf{D}, \mathbf{Z} \rangle \\ &\overset{\text{(b)}}{=} \frac{1}{2}\Big[2\mathrm{Tr}\big(\mathbf{W}\big) - \mathrm{Tr}\big(\mathbf{1}\,\mathrm{diag}(\mathbf{W})^\top\big) - \mathrm{Tr}\big(\mathrm{diag}(\mathbf{W})\,\mathbf{1}^\top\big)\Big] + \frac{1}{2}\langle \mathbf{D}, \mathbf{Z} \rangle \\ &= \frac{1}{2}\Big[2\big(\mathbf{1}^\top\mathrm{diag}(\mathbf{W})\big) - \mathbf{1}^\top\mathrm{diag}(\mathbf{W}) - \mathbf{1}^\top\mathrm{diag}(\mathbf{W})\Big] + \frac{1}{2}\langle \mathbf{D}, \mathbf{Z} \rangle \\ &= \frac{1}{2}\langle \mathbf{D}, \mathbf{Z} \rangle. \end{aligned} \tag{3.4}$$

Here, (a) follows from the observation that the ij-th element of the matrix $2\mathbf{W} + \mathbf{D}$ can be written as $2\boldsymbol{\xi}_i^\top\boldsymbol{\xi}_j + \|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\|^2 = \|\boldsymbol{\xi}_i\|^2 + \|\boldsymbol{\xi}_j\|^2$, and (b) uses the insights that $\mathbf{Z}\mathbf{1} = \mathbf{1}$ and $\mathbf{1}^\top\mathbf{Z} = \mathbf{1}^\top$. Comparing $\mathscr{PW}_1$ and $\mathscr{A}$, identity (3.4) shows that the two relaxations are equivalent because their objective functions are the same (up to a factor two) while they share the same feasible

set. □

Next, we establish that $\mathscr{R}_{\text{SDP}}$ is at least as tight a relaxation of the cardinality-constrained K-means clustering problem (3.1) as the stronger relaxation $\mathscr{PW}_1$ of Peng and Wei (2007).

Proposition 3.11. We have $\min \mathscr{R}_{\text{SDP}} \geq \min \mathscr{PW}_1$.

Note that, through Observation 1, Proposition 3.11 also implies that $\mathscr{R}_{\text{SDP}}$ is at least as tight as the relaxation $\mathscr{A}$ of Awasthi et al. (2015).

Proof of Proposition 3.11. To prove that $\mathscr{R}_{\text{SDP}}$ is at least as tight a relaxation as $\mathscr{PW}_1$, we will argue that for every feasible solution $\{(\boldsymbol{x}^k, \mathbf{M}^k)\}_{k=1}^K$ of $\mathscr{R}_{\text{SDP}}$ one can construct a solution

$$\overline{\mathbf{Z}} = \frac{1}{4}\sum_{k=1}^{K} \frac{1}{n_k}\left(\mathbf{M}^k + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^k\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^k)^\top\right)$$

which is feasible in $\mathscr{PW}_1$ and achieves the same objective value. We first verify the feasibility of the proposed solution $\overline{\mathbf{Z}}$. Note that $\overline{\mathbf{Z}}$ is symmetric by construction. Next, we can directly verify that $\overline{\mathbf{Z}}$ is positive semidefinite since

$$\begin{array}{lll}
\overline{\mathbf{Z}} \geq 0 \quad \Longleftarrow \quad \mathbf{M}^k + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^k\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^k)^\top \geq \mathbf{0} & & \forall k = 1,\ldots,K \\
\quad \Longleftrightarrow \quad \boldsymbol{v}^\top\left(\mathbf{M}^k + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^k\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^k)^\top\right)\boldsymbol{v} \geq 0 & \forall \boldsymbol{v} \in \mathbb{R}^N & \forall k = 1,\ldots,K \\
\quad \Longleftarrow \quad \boldsymbol{v}^\top\left(\boldsymbol{x}^k(\boldsymbol{x}^k)^\top + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^k\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^k)^\top\right)\boldsymbol{v} \geq 0 & \forall \boldsymbol{v} \in \mathbb{R}^N & \forall k = 1,\ldots,K \\
\quad \Longleftrightarrow \quad \left(\boldsymbol{v}^\top\boldsymbol{x}^k\right)^2 + \left(\boldsymbol{v}^\top\mathbf{1}\right)^2 + 2\left(\boldsymbol{v}^\top\boldsymbol{x}^k\right)\left(\boldsymbol{v}^\top\mathbf{1}\right) \geq 0 & \forall \boldsymbol{v} \in \mathbb{R}^N & \forall k = 1,\ldots,K \\
\quad \Longleftrightarrow \quad \left(\boldsymbol{v}^\top\boldsymbol{x}^k + \boldsymbol{v}^\top\mathbf{1}\right)^2 \geq 0 & \forall \boldsymbol{v} \in \mathbb{R}^N & \forall k = 1,\ldots,K,
\end{array}$$

where the third implication is due to the definition of $\mathscr{C}_{\text{SDP}}(n_k)$, which requires that $\mathbf{M}^k \geq \boldsymbol{x}^k(\boldsymbol{x}^k)^\top$. The last statement holds trivially because any quadratic form is non-negative. Next, we can ensure the element-wise non-negativity of $\overline{\mathbf{Z}}$, again through the definition of $\mathscr{C}_{\text{SDP}}(n_k)$:

$$\overline{\mathbf{Z}} \geq \mathbf{0} \quad \Longleftarrow \quad \mathbf{M}^k + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^k\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^k)^\top \geq \mathbf{0} \qquad \forall k = 1,\ldots,K.$$

Furthermore, combining the definition of $\mathscr{C}_{\text{SDP}}(n_k)$ and the constraint $\sum_{k=1}^K \boldsymbol{x}^k = (2-K)\mathbf{1}$ of $\mathscr{R}_{\text{SDP}}$, we can see that each row of $\overline{\mathbf{Z}}$ indeed sums up to one:

$$\begin{aligned}
\overline{\mathbf{Z}}\mathbf{1} &= \frac{1}{4}\sum_{k=1}^{K}\frac{1}{n_k}(\mathbf{M}^k\mathbf{1} + \mathbf{1}\mathbf{1}^\top\mathbf{1} + \boldsymbol{x}^k\mathbf{1}^\top\mathbf{1} + \mathbf{1}(\boldsymbol{x}^k)^\top\mathbf{1}) \\
&= \frac{1}{4}\sum_{k=1}^{K}\frac{1}{n_k}((2n_k - N)\boldsymbol{x}^k + N\mathbf{1} + N\boldsymbol{x}^k + (2n_k - N)\mathbf{1}) \\
&= \frac{1}{2}\sum_{k=1}^{K}(\boldsymbol{x}^k + \mathbf{1}) = \mathbf{1}.
\end{aligned}$$

Finally, the trace of $\overline{\mathbf{Z}}$ is uniquely determined as follows:

$$\begin{aligned}\mathrm{Tr}(\overline{\mathbf{Z}}) &= \frac{1}{4}\sum_{k=1}^{K}\frac{1}{n_k}\mathrm{Tr}\big(\mathbf{M}^k + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^k\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^k)^\top\big)\\ &= \frac{1}{4}\sum_{k=1}^{K}\frac{1}{n_k}\big(2N + 2(\mathbf{1}^\top\boldsymbol{x}^k)\big)\\ &= \frac{1}{4}\sum_{k=1}^{K}\frac{1}{n_k}\big(2N + 2(2n_k - N)\big) = K.\end{aligned}$$

Thus, $\overline{\mathbf{Z}}$ is feasible in $\mathscr{PW}_1$, and it remains to prove that it achieves the same objective value as the original solution $\{(\boldsymbol{x}^k,\mathbf{M}^k)\}_{k=1}^K$ in $\mathscr{R}_{\mathrm{SDP}}$. Invoking relation (3.4), it is easy to see that

$$\big\langle \mathbf{W}, \mathbb{I} - \overline{\mathbf{Z}}\big\rangle = \frac{1}{2}\big\langle \mathbf{D}, \overline{\mathbf{Z}}\big\rangle = \frac{1}{8}\Big\langle \mathbf{D}, \sum_{k=1}^{K}\frac{1}{n_k}(\mathbf{M}^k + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^k\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^k)^\top)\Big\rangle.$$

The proof thus concludes. □

Finally, we assert that $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ is at least as tight a relaxation of the balanced K-means clustering problem as the corresponding relaxation $\mathscr{PW}^{\mathrm{b}}_1$ of Peng and Wei (2007).

Proposition 3.12. We have $\min \mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}} \geq \min \mathscr{PW}^{\mathrm{b}}_1$.

Proof. To show that $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ is at least as tight a relaxation as $\mathscr{PW}^{\mathrm{b}}_1$, we will again argue that for every feasible solution $\{(\boldsymbol{x}^1,\mathbf{M}^1),(\boldsymbol{x},\mathbf{M})\}$ of $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ one can construct a solution

$$\overline{\mathbf{Z}} = \frac{K}{4N}\big((\mathbf{M}^1 + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^1\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^1)^\top) + (K-1)(\mathbf{M} + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}\mathbf{1}^\top + \mathbf{1}\boldsymbol{x}^\top)\big)$$

that is feasible in $\mathscr{PW}^{\mathrm{b}}_1$ and achieves the same objective value. Following similar steps as in the proof of Proposition 3.11, one can verify that $\overline{\mathbf{Z}}$ indeed satisfies $\overline{\mathbf{Z}} \geq \mathbf{0}$, $\overline{\mathbf{Z}} \geq \mathbf{0}$, $\overline{\mathbf{Z}}\mathbf{1} = \mathbf{1}$ and $\mathrm{Tr}(\overline{\mathbf{Z}}) = K$. In order to see that $\overline{\mathbf{Z}} \leq (K/N)\mathbf{1}\mathbf{1}^\top$, note from the definition of $\mathscr{C}_{\mathrm{SDP}}(n)$ (where $n = N/K$ denotes the shared cardinality of all clusters) that

$$\begin{aligned}&2(\mathbf{M}^1 - \mathbf{1}\mathbf{1}^\top) = (\mathbf{M}^1 - \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^1\mathbf{1}^\top - \mathbf{1}(\boldsymbol{x}^1)^\top) + (\mathbf{M}^1 - \mathbf{1}\mathbf{1}^\top - \boldsymbol{x}^1\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^1)^\top) \leq \mathbf{0} \implies \mathbf{M}^1 \leq \mathbf{1}\mathbf{1}^\top,\\ &2(\mathbf{M} - \mathbf{1}\mathbf{1}^\top) = (\mathbf{M} - \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}\mathbf{1}^\top - \mathbf{1}\boldsymbol{x}^\top) + (\mathbf{M} - \mathbf{1}\mathbf{1}^\top - \boldsymbol{x}\mathbf{1}^\top + \mathbf{1}\boldsymbol{x}^\top) \leq \mathbf{0} \implies \mathbf{M} \leq \mathbf{1}\mathbf{1}^\top.\end{aligned}$$

Using this insight and the constraint $\boldsymbol{x}^1 + (K-1)\boldsymbol{x} = (2-K)\mathbf{1}$ of $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$, any arbitrary element $\overline{z}_{ij}$ of $\overline{\mathbf{Z}}$ can be bounded above as desired,

$$\begin{aligned}\overline{z}_{ij} &= \frac{K}{4N}\Big((m^1_{ij} + 1 + x^1_i + x^1_j) + (K-1)(m_{ij} + 1 + x_i + x_j)\Big)\\ &\leq \frac{K}{4N}\Big(2K + x^1_i + (K-1)x_i + x^1_j + (K-1)x_j\Big) = \frac{K}{N}.\end{aligned}$$

Finally, a direct calculation reveals that the objective of $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ evaluated at $\{(\boldsymbol{x}^1,\mathbf{M}^1),(\boldsymbol{x},\mathbf{M})\}$

coincides with the objective of $\mathscr{PW}_1^{\mathrm{b}}$ evaluated at $\overline{\mathbf{Z}}$, which from (3.4) is equal to $\frac{1}{2}\langle \mathbf{D}, \overline{\mathbf{Z}} \rangle$. Hence, $\mathscr{R}_{\mathrm{SDP}}^{\mathrm{b}}$ is at least as tight a relaxation as $\mathscr{PW}_1^{\mathrm{b}}$, and the proof concludes. □

Note that while Propositions 3.11 and 3.12 demonstrate that our SDP relaxations $\mathscr{R}_{\mathrm{SDP}}$ and $\mathscr{R}_{\mathrm{SDP}}^{\mathrm{b}}$ are at least as tight as their respective counterparts by Peng and Wei (2007), similar tightness results cannot be established for our LP relaxations. Indeed, our numerical experiments based on real-world datasets in Section 3.5 show that both $\mathscr{R}_{\mathrm{LP}}$ and $\mathscr{R}_{\mathrm{LP}}^{\mathrm{b}}$ can be strictly weaker than $\mathscr{PW}_1$ and $\mathscr{PW}_1^{\mathrm{b}}$, respectively. Furthermore, it is possible to construct artificial datasets on which even $\mathscr{PW}_2$ outperforms $\mathscr{R}_{\mathrm{LP}}$ and $\mathscr{R}_{\mathrm{LP}}^{\mathrm{b}}$.

3.3.4 Perfect Recovery Guarantees

We now demonstrate that the relaxations of Section 3.3.2 are tight and that Algorithm 2 finds the optimal clustering if the clusters are perfectly separated in the sense of the following assumption.

***(S)* Perfect Separation:** There exists a balanced partition $(J_1, \ldots, J_K)$ of $\{1, \ldots, N\}$ where each cluster $k = 1, \ldots, K$ has the same cardinality $|J_k| = N/K \in \mathbb{N}$, and

$$\max_{1 \leq k \leq K} \max_{i,j \in J_k} d_{ij} < \min_{1 \leq k_1 < k_2 \leq K} \min_{i \in J_{k_1},\, j \in J_{k_2}} d_{ij}.$$

Assumption **(S)** implies that the dataset admits the natural balanced clustering $(J_1, \ldots, J_K)$, and that the largest cluster diameter (*i.e.*, $\max_{1 \leq k \leq K} \max_{i,j \in J_k} d_{ij}$) is smaller than the smallest distance between any two distinct clusters (*i.e.*, $\min_{1 \leq k_1 < k_2 \leq K} \min_{i \in J_{k_1}, j \in J_{k_2}} d_{ij}$).

Theorem 3.13. If Assumption **(S)** holds, then the optimal values of $\mathscr{R}_{LP}^{b}$ and $\mathscr{P}$ coincide. Moreover, the clustering $(J_1, \ldots, J_K)$ is optimal in $\mathscr{P}$ and is recovered by Algorithm 2.

Put simply, Theorem 3.13 states that for datasets whose hidden classes are balanced and well separated, Algorithm 2 will succeed in recovering this hidden, provably optimal clustering.

Proof of Theorem 3.13. Throughout the proof we assume without loss of generality that the clustering $(J_1, \ldots, J_K)$ from Assumption **(S)** satisfies $1 \in J_1$, that is, the cluster containing the datapoint $\boldsymbol{\xi}_1$ is assigned the number $k = 1$. The proof now proceeds in two steps. In the first step, we show that the optimal values of the LP $\mathscr{R}_{\mathrm{LP}}^{\mathrm{b}}$ and the MILP $\mathscr{P}$ are equal and that they both coincide with the sum of squared intra-cluster distances of the clustering $(J_1, \ldots, J_K)$, which amounts to

$$\frac{1}{2n} \sum_{k=1}^{K} \sum_{i,j \in J_k} d_{ij}.$$

In the second step we demonstrate that the output $(I_1, \dots, I_K)$ of Algorithm 2 coincides with the optimal clustering $(J_1, \dots, J_K)$ from Assumption **(S)**. As the algorithm uses the same procedure K times to recover the clusters one by one, it is actually sufficient to show that the first iteration of the algorithm correctly identifies the first cluster, that is, it suffices to prove that $I_1 = J_1$.

Step 1: For any feasible solution $(\boldsymbol{x}^1, \boldsymbol{x}, \mathbf{M}^1, \mathbf{M})$ of $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$, we define $\mathbf{H}, \mathbf{W} \in \mathbb{S}^N$ through

$$\mathbf{H} = \mathbf{M}^1 + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^1\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^1)^\top \quad \text{and} \quad \mathbf{W} = \mathbf{M} + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}\mathbf{1}^\top + \mathbf{1}\boldsymbol{x}^\top. \tag{3.5}$$

From the definition of $\mathscr{C}_{\mathrm{LP}}(n)$ it is clear that $\mathbf{H}, \mathbf{W} \geq \mathbf{0}$. Moreover, we also have that

$$\begin{aligned}\sum_{i \neq j} h_{ij} &= \sum_{i \neq j} m^1_{ij} + N(N-1) + 2(N-1)(\boldsymbol{x}^1)^\top \mathbf{1} \\ &= (2n-N)^2 - N + N(N-1) + 2(N-1)(2n-N) = 4n(n-1).\end{aligned}$$

A similar calculation for $\mathbf{W}$ reveals that $\sum_{i \neq j} w_{ij} = 4n(n-1)$. Next, we consider the objective function of $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$, which can be rewritten in terms of $\mathbf{W}$ and $\mathbf{H}$ as

$$\frac{1}{8n} \langle \mathbf{D}, \mathbf{H} + (K-1)\mathbf{W} \rangle = \frac{1}{8n} \sum_{i \neq j} d_{ij}(h_{ij} + (K-1)w_{ij}). \tag{3.6}$$

The sum on the right-hand side can be viewed as a weighted average of the squared distances d_{ij} with non-negative weights $h_{ij} + (K-1)w_{ij}$, where the total weight is given by

$$\sum_{i \neq j} \big(h_{ij} + (K-1)w_{ij}\big) = 4Kn(n-1).$$

Furthermore each weight $h_{ij} + (K-1)w_{ij}$ is bounded above by 4 because

$$\begin{aligned}h_{ij} + (K-1)w_{ij} &= (m^1_{ij} + 1 + x^1_i + x^1_j) + (K-1)(m_{ij} + 1 + x_i + x_j) \\ &\leq 2K + (x^1_i + (K-1)x_i) + (x^1_j + (K-1)x_j) = 4,\end{aligned} \tag{3.7}$$

where the inequality holds because $\mathbf{M}^1, \mathbf{M} \leq \mathbf{1}\mathbf{1}^\top$ (which we know from the proof of Proposition 3.12) and the last equality follows from the constraint $\boldsymbol{x}^1 + (K-1)\boldsymbol{x} = (2-K)\mathbf{1}$ in $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$.

Hence, the sum on the right-hand side of (3.6) assigns each squared distance d_{ij} with $i \neq j$ a weight of at most 4, while the total weight equals $4Kn(n-1)$. A lower bound on the sum is thus obtained by assigning a weight of 4 to the $Kn(n-1)$ smallest values d_{ij} with $i \neq j$. Thus, we have

$$\begin{aligned}\frac{1}{8n} \langle \mathbf{D}, \mathbf{H} + (K-1)\mathbf{W} \rangle &\geq \frac{1}{2n} \big\{\text{sum of the } Kn(n-1) \text{ smallest entries of } d_{ij} \text{ with } i \neq j\big\} \\ &= \frac{1}{2n} \sum_{k=1}^{K} \sum_{i,j \in J_k} d_{ij},\end{aligned} \tag{3.8}$$

where the last equality follows from Assumption **(S)**. By Lemma 3.1, the right-hand side of (3.8) represents the objective value of the clustering $(J_1,\dots,J_K)$ in the MILP $\mathscr{P}$. Thus, $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$ provides an upper bound on $\mathscr{P}$. By Corollary 3.10, $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$ also provides a lower bound on $\mathscr{P}$. We may thus conclude that the LP relaxation $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$ is tight and, as a consequence, that the clustering $(J_1,\dots,J_K)$ is indeed optimal in $\mathscr{P}$.

Step 2: As the inequality in (3.8) is tight, any optimal solution to $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$ satisfies $h_{ij}+(K-1)w_{ij}=4$ whenever $i\neq j$ and $i,j\in J_k$ for some $k=1,\dots,K$ (*i.e.*, whenever the datapoints $\boldsymbol{\xi}_i$ and $\boldsymbol{\xi}_j$ belong to the same cluster). We will use this insight to show that Algorithm 2 outputs $I_1=J_1$.

For any $i\in J_1$, the above reasoning and our convention that $1\in J_1$ imply that $h_{1i}+(K-1)w_{1i}=4$. This in turn implies via (3.7) that $m^1_{1i}=m_{1i}=1$ for all $i\in J_1$.

From the definition of $\mathscr{C}_{\mathrm{LP}}(n)$, we know that

$$2(\mathbf{M}^1+\mathbf{1}\mathbf{1}^\top)=(\mathbf{M}^1+\mathbf{1}\mathbf{1}^\top+\boldsymbol{x}^1\mathbf{1}^\top+\mathbf{1}(\boldsymbol{x}^1)^\top)+(\mathbf{M}^1+\mathbf{1}\mathbf{1}^\top-\boldsymbol{x}^1\mathbf{1}^\top-\mathbf{1}(\boldsymbol{x}^1)^\top)\geq\mathbf{0}\implies\mathbf{M}^1\geq-\mathbf{1}\mathbf{1}^\top.$$

This allows us to conclude that

$$2n-N=\sum_{i=1}^N m^1_{1i}=\sum_{i\in J_1}m^1_{1i}+\sum_{i\notin J_1}m^1_{1i}\geq n+(N-n)(-1)=2n-N,$$

where the first equality holds because $\mathbf{M}^1\mathbf{1}=(2n-N)\boldsymbol{x}^1$, which is one of the constraints in $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$, and because of our convention that $x^1_1=1$. Hence, the above inequality must be satisfied as an equality, which in turn implies that $m^1_{1i}=-1$ for all $i\notin J_1$.

For any $i\notin J_1$, the $1i$-th entry of the matrix inequality $\mathbf{M}^1+\mathbf{1}\mathbf{1}^\top-\boldsymbol{x}^1\mathbf{1}^\top-\mathbf{1}(\boldsymbol{x}^1)^\top\geq\mathbf{0}$ from the definition of $\mathscr{C}_{\mathrm{LP}}(n)$ can be expressed as

$$0\leq m^1_{1i}+1-x^1_1-x^1_i\quad\forall i=1,\dots,N\implies x^1_i\leq-1,$$

where the implication holds because $m^1_{1i}=-1$ for $i\notin J_1$ and because $x^1_1=1$ due to the symmetry breaking constraint in $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$. Similarly, for any $i\in J_1$, the ii-th entry of the matrix inequality $\mathbf{M}^1+\mathbf{1}\mathbf{1}^\top-\boldsymbol{x}^1\mathbf{1}^\top-\mathbf{1}(\boldsymbol{x}^1)^\top\geq\mathbf{0}$ can be rewritten as

$$0\leq m^1_{ii}+1-2x^1_i\quad\forall i=1,\dots,N\implies x^1_i\leq1,$$

where the implication follows from the constraint $\mathrm{diag}(\mathbf{M}^1)=\mathbf{1}$ in $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$.

As $x^1_i\leq1$ for all $i\in J_1$ and $x^1_i\leq-1$ for all $i\notin J_1$, the equality constraint $\mathbf{1}^\top\boldsymbol{x}^1=2n-N$ from the definition of $\mathscr{C}_{\mathrm{LP}}(n)$ can only be satisfied if $x^1_i=1$ for all $i\in J_1$ and $x^1_i=-1$ for all $i\notin J_1$. Since Algorithm 2 constructs I_1 as the index set of the n largest entries of the vector $\boldsymbol{x}^1$, we conclude that it must output $I_1=J_1$ and the proof completes. □

Theorem 3.13 implies via Corollary 3.10 that the optimal values of $\mathscr{R}^{\mathrm{b}}_{\mathrm{SDP}}$ and $\mathscr{P}$ are also equal. Thus, both the LP and the SDP relaxation lead to perfect recovery.

In the related literature, Assumption **(S)** has previously been used by Elhamifar et al. (2012) to show that the natural clustering can be recovered in the context of unconstrained exemplar-based clustering whenever a regularization parameter is chosen appropriately. In contrast, our formulation does not rely on regularization parameters. Likewise, Theorem 3.13 is reminiscent of Theorem 9 by Awasthi et al. (2015) which formalizes the recovery properties of their LP relaxation for the unconstrained K-means clustering problem. Awasthi et al. (2015) assume, however, that the datapoints are drawn independently from a mixture of K isotropic distributions and provide a probabilistic recovery guarantee that improves with N and deteriorates with d. In contrast, our recovery guarantee for constrained clustering is deterministic, model-free and dimension-independent. If Assumption **(S)** holds, simpler algorithms than Algorithm 1 and 2 can be designed to recover the true clusters. For instance, a simple threshold approach (*i.e.*, assigning datapoints to the same cluster whenever the distance between them falls below a given threshold) would be able to recover the true clusters whenever Assumption **(S)** holds. It seems unlikely, however, that such approaches would perform well in a setting where Assumption **(S)** is not satisfied. In fact, Awasthi et al. (2015) show that their LP relaxation fails to recover the true clusters with high probability if Assumption **(S)** is violated. In contrast, the numerical experiments of Section 3.5 suggest that Algorithms 1 and 2 perform well even if Assumption **(S)** is violated.

Remark 3.14. To our best knowledge, there is no perfect recovery result for the cardinality-constrained K-means clustering algorithm by Bennett et al. (2000), see Appendix, whose performance depends critically on its initialization. To see that it can be trapped in a local optimum, consider the $N = 4$ two-dimensional datapoints $\boldsymbol{\xi}_1 = (0,0)$, $\boldsymbol{\xi}_2 = (a,0)$, $\boldsymbol{\xi}_3 = (a,b)$ and $\boldsymbol{\xi}_4 = (0,b)$ with $0 < a < b$, and assume that we seek two balanced clusters. If the algorithm is initialized with the clustering $\{\{1,4\},\{2,3\}\}$, then this clustering remains unchanged, and the algorithm terminates and reports a suboptimal solution with relative optimality gap $b^2/a^2 - 1$. In contrast, as Assumption **(S)** holds, Algorithm 2 recovers the optimal clustering $\{\{1,2\},\{3,4\}\}$ by Theorem 3.13.

3.4 Cardinality-Constrained Clustering with Outliers

If the dataset is corrupted by outliers, then the optimal value of (3.1) may be high, indicating that the dataset admits no natural clustering. Note that the bounds from Section 3.3 could still be tight, *i.e.*, it is thinkable that the optimal clustering is far from 'ideal' even if it can be found with Algorithm 2. If we gradually remove datapoints that are expensive to assign to any cluster, however, we should eventually discover an 'ideal' low-cost clustering. In the extreme case, if we omit all but K datapoints, then the optimal value of (3.1) drops to zero, and Algorithm 2 detects the optimal clustering due to Theorem 3.13.

We now show that the results of Section 3.3 (particularly Proposition 3.3 and Theorem 3.7)

extend to situations where n_0 datapoints must be assigned to an auxiliary *outlier cluster* indexed by $k = 0$ ($\sum_{k=0}^{K} n_k = N$), and where neither the distances between outliers and retained datapoints nor the distances between different outliers contribute to the objective function. In fact, we could equivalently postulate that each of the n_0 outliers forms a trivial singleton cluster. The use of cardinality constraints in integrated clustering and outlier detection has previously been considered by Chawla and Gionis (2013) in the context of local search heuristics. Inspired by this work, we henceforth minimize the sum of squared intra-cluster distances of the $N - n_0$ non-outlier datapoints. We first prove that the joint outlier detection and cardinality-constrained clustering problem admits an exact MILP reformulation.

Proposition 3.15 (MILP Reformulation)**.** The joint outlier detection and cardinality-constrained clustering problem is equivalent to the MILP

$$\begin{array}{lll} \textit{minimize} & \frac{1}{2}\sum_{k=1}^{K}\frac{1}{n_k}\sum_{i,j=1}^{N} d_{ij}\eta_{ij}^k & \\ \textit{subject to} & \pi_i^k \in \{0,1\},\ \eta_{ij}^k \in \mathbb{R}_+ & i,j = 1,\ldots,N,\ k = 0,\ldots,K \\ & \sum_{i=1}^{N}\pi_i^k = n_k & k = 0,\ldots,K \\ & \sum_{k=0}^{K}\pi_i^k = 1 & i = 1,\ldots,N \\ & \eta_{ij}^k \geq \pi_i^k + \pi_j^k - 1 & i,j = 1,\ldots,N,\ k = 0,\ldots,K. \end{array} \qquad (\mathscr{P}^{\mathrm{o}})$$

Proof. This is an immediate extension of Proposition 3.3 to account for the outlier cluster. □

In analogy to Section 3.3.1, one can demonstrate that the MILP $\mathscr{P}^{\mathrm{o}}$ admits the SDP relaxation

$$\begin{array}{ll} \text{minimize} & \frac{1}{8}\left\langle \mathbf{D}, \sum_{k=1}^{K}\frac{1}{n_k}\left(\mathbf{M}^k + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}^k\mathbf{1}^\top + \mathbf{1}(\boldsymbol{x}^k)^\top\right)\right\rangle \\ \text{subject to} & (\boldsymbol{x}^k, \mathbf{M}^k) \in \mathscr{C}_{\mathrm{SDP}}(n_k) \quad k = 0,\ldots,K \\ & \sum_{k=0}^{K}\boldsymbol{x}^k = (1-K)\mathbf{1}. \end{array} \qquad (\mathscr{R}^{\mathrm{o}}_{\mathrm{SDP}})$$

Moreover, $\mathscr{R}^{\mathrm{o}}_{\mathrm{SDP}}$ can be further relaxed to an LP, henceforth denoted by $\mathscr{R}^{\mathrm{o}}_{\mathrm{LP}}$, by replacing the semidefinite representable set $\mathscr{C}_{\mathrm{SDP}}(n_k)$ in $\mathscr{R}^{\mathrm{o}}_{\mathrm{SDP}}$ with the polytope $\mathscr{C}_{\mathrm{LP}}(n_k)$ for all $k = 0,\ldots,K$.

Theorem 3.16 (SDP and LP Relaxations)**.** We have $\min \mathscr{R}^o_{LP} \leq \min \mathscr{R}^o_{SDP} \leq \min \mathscr{P}^o$.

Proof. This result generalizes Theorem 3.7 to account for the additional outlier cluster. As it requires no fundamentally new ideas, the proof is omitted for brevity. □

The relaxations $\mathscr{R}^{\mathrm{o}}_{\mathrm{SDP}}$ and $\mathscr{R}^{\mathrm{o}}_{\mathrm{LP}}$ not only provide a lower bound on $\mathscr{P}^{\mathrm{o}}$, but they also give rise to a rounding algorithm that recovers a feasible clustering and thus an upper bound on $\mathscr{P}^{\mathrm{o}}$; see Algorithm 3. Note that this procedure calls the outlier-unaware Algorithm 1 as a subroutine.

If all normal clusters are equally sized, *i.e.*, $n_k = n$ for $k = 1,\ldots,K$, then $\mathscr{R}^{\mathrm{o}}_{\mathrm{SDP}}$ can be replaced by

Algorithm 3 Rounding algorithm for joint outlier detection and cardinality-constrained clustering

1: **Input:** $\mathscr{I}_0 = \{1,\dots,N\}$ (data indices), $n_k \in \mathbb{N},\ k = 0,\dots,K$ (cluster sizes).
2: Solve $\mathscr{R}^{\mathrm{o}}_{\mathrm{SDP}}$ or $\mathscr{R}^{\mathrm{o}}_{\mathrm{LP}}$ for the datapoints $\boldsymbol{\xi}_i$, $i \in \mathscr{I}_0$, and record the optimal $\boldsymbol{x}^0 \in \mathbb{R}^N$.
3: Determine a bijection $\rho : \mathscr{I}_0 \to \mathscr{I}_0$ such that $x^0_{\rho(1)} \geq x^0_{\rho(2)} \geq \cdots \geq x^0_{\rho(N)}$.
4: Set $I_0 \leftarrow \{\rho(1),\dots,\rho(n_0)\}$ and $\mathscr{I}_1 \leftarrow \mathscr{I}_0 \setminus I_0$.
5: Call Algorithm 1 with input $(\mathscr{I}_1, \{n_k\}_{k=1}^K)$ to obtain $I_1,\dots,I_K$.
6: **Output:** $I_0,\dots,I_K$.

$$\begin{array}{ll} \text{minimize} & \frac{K}{8n}\,\langle \mathbf{D}, \mathbf{M} + \mathbf{1}\mathbf{1}^\top + \boldsymbol{x}\mathbf{1}^\top + \mathbf{1}\boldsymbol{x}^\top \rangle \\ \text{subject to} & (\boldsymbol{x},\mathbf{M}) \in \mathscr{C}_{\mathrm{SDP}}(n), \quad (\boldsymbol{x}^0,\mathbf{M}^0) \in \mathscr{C}_{\mathrm{SDP}}(n_0), \quad K\boldsymbol{x} + \boldsymbol{x}^0 = (1-K)\mathbf{1}, \end{array} \qquad (\mathscr{R}^{\mathrm{ob}}_{\mathrm{SDP}})$$

whose size no longer scales with K. Similarly, $\mathscr{R}^{\mathrm{o}}_{\mathrm{LP}}$ simplifies to the LP $\mathscr{R}^{\mathrm{ob}}_{\mathrm{LP}}$ obtained from $\mathscr{R}^{\mathrm{ob}}_{\mathrm{SDP}}$ by replacing $\mathscr{C}_{\mathrm{SDP}}(n)$ and $\mathscr{C}_{\mathrm{SDP}}(n_0)$ with $\mathscr{C}_{\mathrm{LP}}(n)$ and $\mathscr{C}_{\mathrm{LP}}(n_0)$, respectively. Note that the cardinality $n_0 = N - Kn$ may differ from n.

Corollary 3.17 (Relaxations for Balanced Clustering)**.** We have $\min \mathscr{R}^{ob}_{LP} \leq \min \mathscr{R}^{ob}_{SDP} \leq \min \mathscr{P}^{o}$.

Proof. This follows from a marginal modification of the argument that led to Corollary 3.10. □

If the normal clusters are required to be balanced, then Algorithm 3 should be modified as follows. First, in Step 2 the relaxations $\mathscr{R}^{\mathrm{ob}}_{\mathrm{SDP}}$ or $\mathscr{R}^{\mathrm{ob}}_{\mathrm{LP}}$ can be solved instead of $\mathscr{R}^{\mathrm{o}}_{\mathrm{SDP}}$ or $\mathscr{R}^{\mathrm{o}}_{\mathrm{LP}}$, respectively. Moreover, in Step 5 Algorithm 2 must be called as a subroutine instead of Algorithm 1.

In the presence of outliers, the perfect recovery result from Theorem 3.13 remains valid if the following perfect separation condition is met, which can be viewed as a generalization of Assumption **(S)**.

***(S')* Perfect Separation:** There exists a partition $(J_0, J_1, \dots, J_K)$ of $\{1,\dots,N\}$ where each normal cluster $k = 1,\dots,K$ has the same cardinality $|J_k| = (N - n_0)/K \in \mathbb{N}$, while

$$\max_{1\leq k\leq K} \max_{i,j\in J_k} d_{ij} < \min_{1\leq k_1<k_2\leq K} \min_{i\in J_{k_1}, j\in J_{k_2}} d_{ij} \quad \text{and} \quad \max_{1\leq k\leq K} \max_{i,j\in J_k} d_{ij} < \min_{i\in J_0,\, j\in\{1,\dots,N\}\setminus\{i\}} d_{ij}.$$

Assumption **(S')** implies that the dataset admits the natural outlier cluster J_0 and the natural normal clusters $(J_1,\dots,J_K)$. It also postulates that the diameter of each normal cluster is strictly smaller than (i) the distance between any two distinct normal clusters and (ii) the distance between any outlier and any other datapoint. Under this condition, Algorithm 3 correctly identifies the optimal clustering.

Theorem 3.18. If Assumption **(S')** holds, then the optimal values of $\mathscr{R}^{ob}_{LP}$ and $\mathscr{P}^o$ coincide. Moreover, the clustering $(J_0,\dots,J_K)$ is optimal in $\mathscr{P}^o$ and is recovered by Algorithm 3.

Proof. The proof parallels that of Theorem 3.13 and can be divided into two steps. In the first step we show that the LP relaxation $\mathscr{R}^{\text{ob}}_{\text{LP}}$ for balanced clustering and outlier detection is tight, and in the second step we demonstrate that Algorithm 3 correctly identifies the clusters $(J_0,\dots,J_K)$. As for the second step, it suffices to prove that the algorithm correctly identifies the outlier cluster J_0. Indeed, once the outliers are removed, the residual dataset satisfies Assumption **(S)**, and Theorem 3.13 guarantees that the normal clusters $(J_1,\dots,J_K)$ are correctly identified with Algorithm 2.

As a preliminary, note that $(\boldsymbol{x},\mathbf{M}) \in \mathscr{C}_{\text{LP}}(n)$ implies

$$\begin{aligned}\operatorname{diag}(\mathbf{M}+\mathbf{1}\mathbf{1}^\top+\boldsymbol{x}\mathbf{1}^\top+\mathbf{1}\boldsymbol{x}^\top)\geq\mathbf{0} &\implies \boldsymbol{x}\geq-\mathbf{1},\\ \operatorname{diag}(\mathbf{M}+\mathbf{1}\mathbf{1}^\top-\boldsymbol{x}\mathbf{1}^\top-\mathbf{1}\boldsymbol{x}^\top)\geq\mathbf{0} &\implies \boldsymbol{x}\leq+\mathbf{1},\end{aligned}$$

where the implications use $\operatorname{diag}(\mathbf{M})=\mathbf{1}$. Similarly, $(\boldsymbol{x}^0,\mathbf{M}^0)\in\mathscr{C}_{\text{LP}}(n_0)$ implies $-\mathbf{1}\leq\boldsymbol{x}^0\leq+\mathbf{1}$.

Step 1: For any feasible solution $(\boldsymbol{x}^0,\boldsymbol{x},\mathbf{M}^0,\mathbf{M})$ of $\mathscr{R}^{\text{ob}}_{\text{LP}}$, introduce the auxiliary matrix $\mathbf{H}=\mathbf{M}+\mathbf{1}\mathbf{1}^\top+\mathbf{1}\boldsymbol{x}^\top+\boldsymbol{x}\mathbf{1}^\top$. Recall from the proof of Theorem 3.13 that $\mathbf{H}\geq\mathbf{0}$ and

$$\sum_{i\neq j} h_{ij} = 4n(n-1).$$

The constraint $K\boldsymbol{x}+\boldsymbol{x}^0=(1-K)\mathbf{1}$ from $\mathscr{R}^{\text{ob}}_{\text{LP}}$ ensures via the inequality $-\mathbf{1}\leq\boldsymbol{x}^0$ that $\boldsymbol{x}\leq(\frac{2}{K}-1)\mathbf{1}$. Recalling from the proof of Theorem 3.13 that $\mathbf{M}\leq\mathbf{1}\mathbf{1}^\top$, we then find

$$h_{ij}=m_{ij}+1+x_i+x_j\leq 1+1+\left(\frac{2}{K}-1\right)+\left(\frac{2}{K}-1\right)=\frac{4}{K}\quad\forall i,j=1,\dots,N. \tag{3.9}$$

Similar arguments as in the proof of Theorem 3.13 reveal that the objective function of the joint outlier detection and (balanced) clustering problem $\mathscr{R}^{\text{ob}}_{\text{LP}}$ can be expressed as

$$\begin{aligned}\frac{K}{8n}\langle\mathbf{D},\mathbf{H}\rangle &\geq \frac{1}{2n}\left\{\text{sum of the } Kn(n-1) \text{ smallest entries of } d_{ij} \text{ with } i\neq j\right\}\\ &= \frac{1}{2n}\sum_{k=1}^{K}\sum_{i,j\in J_k} d_{ij},\end{aligned} \tag{3.10}$$

where the equality follows from Assumption **(S')**. By Lemma 3.1, the right-hand side of (3.10) represents the objective value of the clustering $(J_0,\dots,J_K)$ in the MILP $\mathscr{P}^{\text{o}}$. Thus, $\mathscr{R}^{\text{ob}}_{\text{LP}}$ provides an upper bound on $\mathscr{P}^{\text{o}}$. By Corollary 3.17, $\mathscr{R}^{\text{ob}}_{\text{LP}}$ also provides a lower bound on $\mathscr{P}^{\text{o}}$. We may thus conclude that the LP relaxation $\mathscr{R}^{\text{ob}}_{\text{LP}}$ is tight and, as a consequence, that the clustering $(J_0,\dots,J_K)$ is indeed optimal in $\mathscr{P}^{\text{o}}$.

Step 2: As the inequality in (3.10) is tight, any optimal solution to $\mathscr{R}^{\text{ob}}_{\text{LP}}$ satisfies $h_{ij} = \frac{4}{K}$ whenever $i \neq j$ and $i, j \in J_k$ for some $k = 1, \ldots, K$ (*i.e.*, whenever $\boldsymbol{\xi}_i$ and $\boldsymbol{\xi}_j$ are *not* outliers and belong to the same cluster). This in turn implies via (3.9) that $x_i = \frac{2}{K} - 1$ for all $i \in \cup_{k=1}^{K} J_k$. Furthermore, the constraint $\mathbf{1}^\top \boldsymbol{x} = 2n - N$ from $\mathscr{C}_{\text{LP}}(n)$ implies

$$2n - N = \sum_{k=1}^{K} \sum_{i \in J_k} x_i + \sum_{i \in J_0} x_i \geq Kn\left(\frac{2}{K} - 1\right) + \sum_{i \in J_0} (-1) = 2n - N,$$

where the inequality holds because $-\mathbf{1} \leq \boldsymbol{x}$. Thus, the above inequality must in fact hold as an equality, which implies that $x_i = -1$ for all $i \in J_0$. The constraint $K\boldsymbol{x} + \boldsymbol{x}^0 = (1 - K)\mathbf{1}$ from $\mathscr{R}^{\text{ob}}_{\text{LP}}$ further implies that $x^0_i = -1$ for all $i \in \cup_{k=1}^{K} J_k$ and $x^0_i = +1$ for all $i \in J_0$.

Since Algorithm 3 constructs I_0 as the index set of the $n_0 = N - Kn$ largest entries of the vector $\boldsymbol{x}^0$, we conclude that it must output $I_0 = J_0$, and the proof completes. □

Remark 3.19 (Unknown Cluster Cardinalities)**.** The joint outlier detection and cardinality-constrained clustering problem $\mathscr{P}^o$ can also be used when the number of outliers is not precisely known and only an estimate of the relative size (as opposed to the exact cardinality) of the clusters is available. To this end, we solve $\mathscr{P}^o$ for different values of n_0, respectively assigning the remaining $N - n_0$ datapoints to clusters whose relative sizes respect the available estimates. The value $n_0^\star$ representing the most reasonable number of outliers to be removed from the dataset can then be determined using the elbow method; see, *e.g.*, Gareth et al. (2017, Chapter 10).

As an illustration, consider again the dataset depicted in Figure 3.1 which showcases the crux of outlier detection in the context of cardinality-constrained clustering. In Section 1, we inadvertently assumed to have the knowledge that the dataset under consideration was contaminated by three outliers. To demonstrate the practical usefulness of our approach, we will now employ the elbow method to determine the number of outliers n_0 without making any assumptions about the dataset. As elucidated in Remark 4, the ideal value of n_0 can be determined by solving problem $\mathscr{P}^o$ repeatedly. However, as $\mathscr{P}^o$ constitutes an intractable optimization problem, we solve its convex relaxations $\mathscr{R}^{\text{ob}}_{\text{LP}}$ and $\mathscr{R}^{\text{ob}}_{\text{SDP}}$ instead and plot the resulting objective values in logarithmic scale in Figure 3.2. It becomes apparent that $n_0^\star = 3$ is most appropriate as it marks the transition from the initially steep decline pattern of the objective value to a substantially flatter decline pattern. Note that n_0 needs to be a multiple of $K = 3$ to allow for balanced clustering.

3.5 Numerical Experiments

We now investigate the performance of our algorithms on synthetic as well as real-world clustering problems with and without outliers. All LPs and SDPs are solved with CPLEX 12.7.1 and MOSEK 8.0, respectively, using the YALMIP interface on a 3.40GHz i7 computer with 16GB RAM.

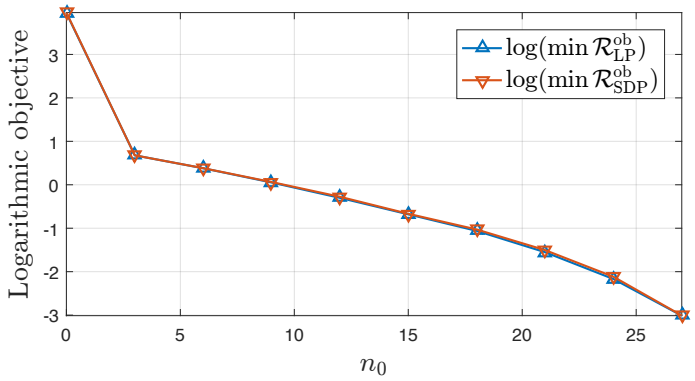


Figure 3.2 – Elbow plot for the dataset depicted in Figure 3.1.

3.5.1 Cardinality-Constrained K-Means Clustering (Real-World Data)

We compare the performance of our algorithms from Section 3.3 with the algorithm of Bennett et al. (2000), see Appendix, and with the two SDP relaxations proposed by Peng and Wei (2007) on the classification datasets of the UCI Machine Learning Repository (http://archive.ics.uci.edu/ml/) with 150–300 datapoints, up to 200 continuous attributes and no missing values. Table 3.2 reports the main characteristics of these datasets. In our experiments, we set the cluster cardinalities to the numbers of true class occurrences in each dataset. It should be emphasized that, in contrast to the other methods, and with exception of the two balanced datasets, the SDP relaxations of Peng and Wei (2007) do not have access to the cluster cardinalities. They should thus be seen as a baseline for the performance of the other methods. Furthermore, we remark that all datasets severely violate Assumption **(S)**. Indeed, the ratios of largest cluster diameter to smallest distance between clusters (when the clusters are determined by the true labels) vary from 7 to 149, while they should be smaller than one in order to satisfy Assumption **(S)**. Also, only two datasets actually entail balanced clusters.

Table 3.3 reports the lower bounds provided by $\mathscr{R}_{\text{LP}}/\mathscr{R}^{\text{b}}_{\text{LP}}$ and $\mathscr{R}_{\text{SDP}}/\mathscr{R}^{\text{b}}_{\text{SDP}}$ (LB), the upper bounds from Algorithms 1 and 2 (UB), the objective value of the best of 10 runs of the algorithm of Bennett et al. (UB), randomly initialized by the cluster centers produced by the *K-means++ algorithm* of Arthur and Vassilvitskii (2007), the coefficient of variation across these 10 runs (CV), the respective lower bounds (LB) obtained from the SDP relaxations $\mathscr{PW}_1/\mathscr{PW}^{\text{b}}_1$ and $\mathscr{PW}_2$ of Peng and Wei (2007), and the solution times for each of these methods. The latter was limited to a maximum of three hours, and in one case (namely, "Glass Identification"), $\mathscr{R}_{\text{SDP}}$ did not terminate within this limit. The "–" signs in Table 3.3 indicate this occurrence.

The obtained lower bounds of $\mathscr{R}_{\text{SDP}}/\mathscr{R}^{\text{b}}_{\text{SDP}}$ allow us to certify that the algorithm of Bennett et al. (2000) provides nearly optimal solutions in almost all instances. Also, both Algorithms 1 and 2 are competitive in terms of solution quality with the algorithm of Bennett et al. (2000) while providing rigorous error bounds. Moreover, as expected in view of Propositions 3.11 and 3.12, for all datasets $\mathscr{R}_{\text{SDP}}/\mathscr{R}^{\text{b}}_{\text{SDP}}$ yield better lower bounds than the SDP relaxations $\mathscr{PW}_1/\mathscr{PW}^{\text{b}}_1$

and $\mathscr{PW}_2$ of Peng and Wei (2007). The lower bounds obtained from $\mathscr{R}_{\text{LP}}/\mathscr{R}^{\text{b}}_{\text{LP}}$ are competitive with those provided by the relaxations $\mathscr{PW}_1/\mathscr{PW}^{\text{b}}_1$, and they are always better than the lower bounds provided by their relaxation $\mathscr{PW}_2$. It should be mentioned, however, that one can construct instances where the situation is reversed, *i.e.*, both $\mathscr{PW}_1/\mathscr{PW}^{\text{b}}_1$ and $\mathscr{PW}_2$ are tighter than $\mathscr{R}_{\text{LP}}/\mathscr{R}^{\text{b}}_{\text{LP}}$. Peng and Wei (2007) also suggest a procedure to compute a feasible clustering (and thus upper bounds) for the unconstrained K-means clustering problem. However, this procedure relies on an enumeration of all possible Voronoi partitions, which is impractical for $K \geq 3$; see Inaba et al. (1994). Furthermore, it is not clear how to impose cardinality constraints in this setting.

ID	Dataset Name	N (# datapoints)	d (# dimensions)	K (# clusters)	n_k (cardinalities)	balanced
1	Iris	150	4	3	50, 50, 50	yes
2	Seeds	210	7	3	70, 70, 70	yes
3	Planning Relax	182	12	2	130, 52	no
4	Connectionist Bench	208	60	2	111, 97	no
5	Urban Land Cover	168	147	9	23, 29, 14, 15, 17, 25, 16, 14, 15	no
6	Parkinsons	195	22	2	48, 147	no
7	Glass Identification	214	9	6	70, 76, 17, 13, 9, 29	no

Table 3.2 – Overview of the main dataset characteristics.

	$\mathscr{R}_{\text{LP}}/\mathscr{R}^{\text{b}}_{\text{LP}}$			$\mathscr{R}_{\text{SDP}}/\mathscr{R}^{\text{b}}_{\text{SDP}}$			Bennett et al.			$\mathscr{PW}_1/\mathscr{PW}^{\text{b}}_1$		$\mathscr{PW}_2$	
ID	UB	LB	time [s]	UB	LB	time [s]	UB	CV [%]	time [s]	LB	time [s]	LB	time [s]
1	81.4	78.8	17	81.4	81.4	584	81.4	0.0	6	81.4	154	15.2	0.02
2	620.7	539.0	46	605.6	605.6	3,823	605.6	0.0	7	604.5	1,320	19.0	0.03
3	325.9	297.0	24	315.7	315.7	2,637	315.8	0.3	9	299.0	510	273.7	0.02
4	312.6	259.1	49	280.6	280.1	3,638	280.6	0.4	6	270.0	1,376	246.2	0.04
5	3.61e9	3.17e9	2,241	3.54e9	3.44e9	10,754	3.64e9	9.2	13	2.05e9	460	1.94e8	0.02
6	1.36e6	1.36e6	22	1.36e6	1.36e6	2,000	1.36e6	15.1	7	1.11e6	777	6.31e5	0.02
7	469.0	377.2	232	–	–	–	438.2	28.4	13	321.9	1,500	23.8	0.03

Table 3.3 – Performance of $\mathscr{R}_{\text{LP}}$, $\mathscr{R}_{\text{SDP}}$, Bennett et al., and Peng and Wei. The "–" signs indicate that the problem instance could not be solved within a time limit of three hours.

Specifically, in the context of the two balanced datasets (*i.e.*, "Iris" and "Seeds"), we can enrich the preceding comparison with the heuristics proposed by Costa et al. (2017) and Malinen and Fränti (2014). As for the variable neighborhood search method of Costa et al. (2017), we were provided with the executables of the *C++* implementation used in that paper. For the "Iris" dataset, the best objective value out of 10 independent runs of this method was 81.4 (which is provably optimal thanks to the lower bounds provided by $\mathscr{R}_{\text{SDP}}$ and $\mathscr{PW}^{\text{b}}_1$) and the time to execute all runs was 0.12 seconds. For the "Seeds" dataset, the best objective value out of 10 independent runs was 605.6 (again, provably optimal in view of the lower bound provided by $\mathscr{R}_{\text{SDP}}$) and the overall runtime was 0.53 seconds. The algorithm of Malinen and Fränti (2014) follows the same steps as the one of Bennett et al. (2000) with the improvement that the cluster assignment step is solved by the Hungarian algorithm, which provides better runtime guarantees and typically solves faster than interior-point methods for LPs. For this reason,

the upper bounds of Malinen and Fränti (2014) for the "Iris" and "Seeds" dataset coincide with those of Bennett et al. (2000) while their algorithm can be expected to terminate faster. A direct comparison of the time complexity of these two methods can be found in Malinen and Fränti (2014).

3.5.2 Cardinality-Constrained K-Means Clustering (Synthetic Data)

We now randomly generate partitions of 10, 20 and 70 datapoints in $\mathbb{R}^2$ that are drawn from uniform distributions over $K = 3$ unit balls centered at $\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2$ and $\boldsymbol{\zeta}_3$, respectively, such that $\|\boldsymbol{\zeta}_1 - \boldsymbol{\zeta}_2\| = \|\boldsymbol{\zeta}_1 - \boldsymbol{\zeta}_3\| = \|\boldsymbol{\zeta}_2 - \boldsymbol{\zeta}_3\| = \delta$. Theorem 3.13 shows that $\mathscr{R}^{\mathrm{b}}_{\mathrm{LP}}$ is tight and that Algorithm 2 can recover the true clusters whenever $n_1 = n_2 = n_3$ and $\delta \geq 4$. Figure 3.3 demonstrates that in practice, perfect recovery is often achieved by Algorithm 1 even if $\delta \ll 4$ and $n_1 \neq n_2 \neq n_3$. We also note that $\mathscr{R}_{\mathrm{SDP}}$ outperforms $\mathscr{R}_{\mathrm{LP}}$ when δ is small, and that the algorithm of Bennett et al. frequently fails to determine the optimal solution even if it is run 10 times. In line with the results from the real-world datasets, $\mathscr{R}_{\mathrm{SDP}}$ and $\mathscr{R}_{\mathrm{LP}}$ are tighter than the stronger SDP relaxation of Peng and Wei (2007). Furthermore, it can be shown that in this setting the weaker relaxation of Peng and Wei (2007) always yields the trivial lower bound of zero. The average runtimes are 7s ($\mathscr{R}_{\mathrm{LP}}$), 106s ($\mathscr{R}_{\mathrm{SDP}}$), 11s (Bennett et al.) and 15.6s (Peng and Wei).

3.5.3 Outlier Detection

We use $\mathscr{R}^{\mathrm{o}}_{\mathrm{LP}}$ and Algorithm 3 to classify the *Breast Cancer Wisconsin (Diagnostic)* dataset. The dataset has $d = 30$ numerical features, which we standardize using a Z-score transformation, and it contains 357 benign and 212 malignant cases of breast cancer. We interpret the malignant cases as outliers and thus set $K = 1$. Figure 3.4 reports the prediction accuracy as well as the false positives (benign cancers classified as malignant) and false negatives (malignant cancers classified as benign) as we increase the number of outliers n_0 from 0 to 400. The figure shows that while setting $n_0 \approx 212$ (the true number of malignant cancers) maximizes the prediction accuracy, any choice $n_0 \in [156, 280]$ leads to a competitive prediction accuracy above 80%. Thus, even rough estimates of the number of malignant cancer datapoints can lead to cancer predictors of decent quality. The average runtime is 286s, and the optimality gap is consistently below 3.23% for all values of n_0.

3.6 Conclusion

Clustering is a hard combinatorial optimization problem. For decades, it has almost exclusively been addressed by heuristic approaches. Many of these heuristics have proven to be very successful in practice as they often provide solutions of high, or at least satisfactory, quality within attractive runtimes. The common drawback of these methods is that there is typically no way of certifying the optimality of the provided solutions nor to give guaranteed bounds on their suboptimality.

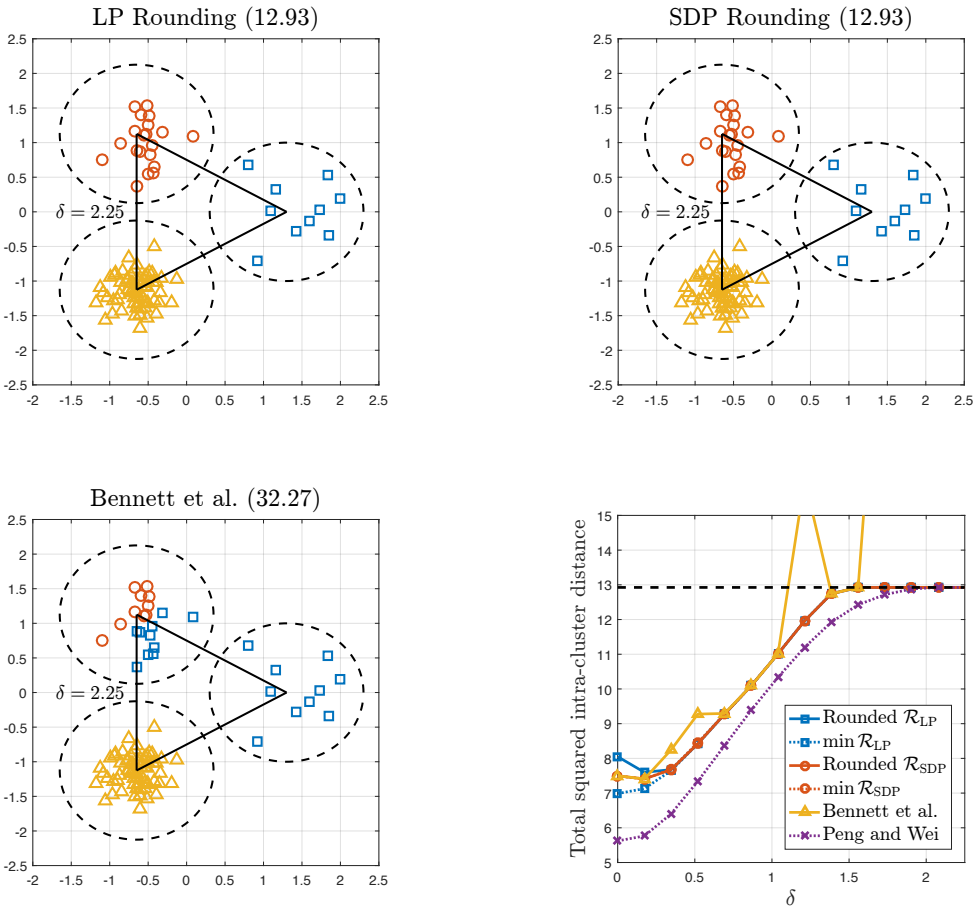


Figure 3.3 – Comparison between different algorithms for (cardinality-constrained) K-means clustering for 100 datapoints where the cardinalities are given by $(n_1, n_2, n_3) = (10, 20, 70)$. Indicated in parentheses next to the panel titles are the respectively achieved sums of squared intra-cluster distances.

Maybe precisely because of this shortcoming, more recently, convex optimization approaches have been proposed for solving relaxed versions of the clustering problem. These conic programs are polynomial-time solvable and offer bounds on the suboptimality of a given solution. Furthermore, the solutions of these conic relaxations can be "rounded" to obtain actually feasible solutions to the original clustering problem, which results in a new class of heuristic methods.

The results presented in this paper follow precisely this recent paradigm. Combined, conic relaxations and (rounding) heuristics offer solutions to the clustering problem together with a-posteriori guarantees on their optimality. Naturally, one would also wish for attractive a-priori guarantees on the performance of these combined methods. The conditions required to derive such a-priori guarantees are still quite restrictive, but the strong performance of these methods on practical instances makes us confident that this is a promising avenue for

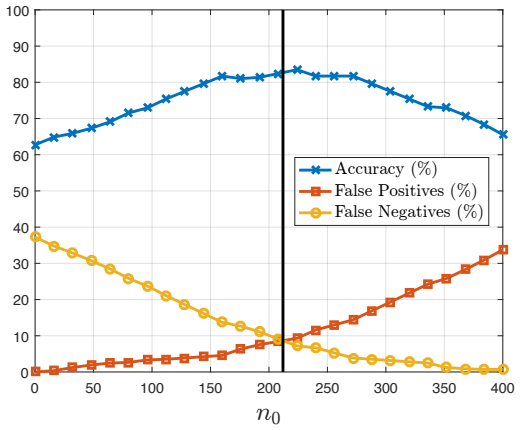


Figure 3.4 – Outlier detection for breast cancer diagnosis.

future research.

Appendix: Algorithm of Bennett et al. (2000)

The algorithm of Bennett et al. (2000) is designed for a variant of problem (3.1), where only lower bounds on the clusters' cardinalities are imposed. This algorithm has a natural extension to our cardinality-constrained clustering problem (3.1) as follows.

Algorithm 4 Algorithm of Bennett et al. for cardinality-constrained clustering

1: **Input:** $\mathscr{I}_1 = \{1,\ldots,N\}$ (data indices), $n_k \in \mathbb{N}$, $k = 1,\ldots,K$ (cluster sizes).
2: Generate the cluster centers $\boldsymbol{\zeta}_1,\ldots,\boldsymbol{\zeta}_K \in \mathbb{R}^d$.
3: Solve the linear assignment problem

$$\boldsymbol{\Pi}^\star \in \operatorname*{argmin}_{\Pi} \left\{ \sum_{i=1}^{N}\sum_{k=1}^{K} \pi_i^k \|\boldsymbol{\xi}_i - \boldsymbol{\zeta}_k\|^2 : \pi_i^k \in \{0,1\},\ \sum_{i=1}^{N} \pi_i^k = n_k\ \forall k,\ \sum_{k=1}^{K} \pi_i^k = 1\ \forall i \right\}.$$

4: Set $I_k \leftarrow \{i : (\pi^\star)_i^k = 1\}$ for all $k = 1,\ldots,K$.
5: Set $\boldsymbol{\zeta}_k \leftarrow \frac{1}{n_k}\sum_{i\in I_k} \boldsymbol{\xi}_i$ for all $k = 1,\ldots,K$.
6: Repeat Steps 3–5 until there are no more changes in $\boldsymbol{\zeta}_1,\ldots,\boldsymbol{\zeta}_K$.
7: **Output:** $I_1,\ldots,I_K$.

Algorithm 4 adapts a classical local search heuristic for the unconstrained K-means clustering problem due to Lloyd (1982) to problem (3.1). At initialization, it generates random cluster centers $\boldsymbol{\zeta}_k$, $k = 1,\ldots,K$. Each subsequent iteration of the algorithm consists of two steps. The first step assigns every datapoint $\boldsymbol{\xi}_i$ to the nearest cluster center while adhering to the prescribed cluster cardinalities, whereas the second step replaces each center $\boldsymbol{\zeta}_k$ with the mean of the datapoints that have been assigned to cluster k. The algorithm terminates when the cluster centers $\boldsymbol{\zeta}_1,\ldots,\boldsymbol{\zeta}_K$ no longer change.

4 A Planner-Trader Decomposition for Multi-Market Hydro Scheduling

Peak/off-peak spreads on European electricity spot markets are eroding due to the ongoing nuclear phaseout and the steady growth in photovoltaic capacity. The reduced profitability of peak/off-peak arbitrage thus forces hydropower producers to participate in the reserve markets. We propose a two-layer stochastic programming framework for the optimal operation of a fleet of interconnected hydropower plants that sell energy on both the spot and the reserve markets. The outer layer problem (the *planner's problem*) optimizes end-of-day reservoir filling levels over one year, whereas the inner layer problem (the *trader's problem*) selects optimal hourly market bids within each day. Using an information restriction whereby the planner prescribes the end-of-day reservoir targets one day in advance, we prove that the trader's problem simplifies from an infinite-dimensional stochastic program with 25 stages to a finite two-stage stochastic program with only two scenarios. Substituting this reformulation back into the outer layer and approximating the reservoir targets by affine decision rules then allows us to simplify the planner's problem from an infinite-dimensional stochastic program with 365 stages to a two-stage stochastic program that can conveniently be solved via the sample average approximation. Numerical experiments based on a cascade in the Salzburg region of Austria demonstrate the effectiveness of the suggested framework.

4.1 Introduction

Electricity from renewable sources, *e.g.*, wind, geothermal, solar and hydropower, has seen its share growing in European electricity markets in recent years. The increase of renewable energies is resulting in numerous environmental and economic benefits. However, electricity generation from some of these sources, especially wind and solar, is intermittent and difficult to forecast because of its reliance on weather and sunlight conditions. Hence, there is a growing need to invest in power plants with storage capacities that can produce or consume electricity on a short notice. For example, pumped-storage hydropower plants are capable of buffering short-term fluctuations in demand and supply because of their storage capabilities and negligible start-up times.

In terms of wholesale electricity markets, most generation companies in Switzerland, France, Germany, Luxembourg and Austria participate in the European Energy Exchange (EEX), which is one of the largest electricity markets in central Europe.[I] Among the markets offered by EEX, the European Power Exchange Spot (EPEX SPOT) is an exchange for power spot trading. It consists of different forward markets, with the main component being the day-ahead market. Generation companies operating pumped-storage hydropower plants typically trade in this market. In the remainder of the paper, we use the terminology *spot market* to refer to this day-ahead market. Traditionally, pumped-storage hydropower plants benefit from participating in the spot market by releasing the water downstream for electricity generation at peak times and by pumping the water upstream during off-peak periods for future generation ('*buy low and sell high*'). In doing so, the generation companies exploit the spreads between peak and off-peak electricity prices to make immediate profits. However, these spreads are eroding since 2008 (Mayer 2014) for two main reasons: (i) the phaseout of nuclear power plants throughout Western Europe and (ii) the rapid growth in photovoltaic capacity; see Morris and Pehnt (2015) as well as Wirth (2016). Nuclear power plants are important sources of base load power. As a result, their withdrawal from the electricity markets increases the base load electricity prices. On the other hand, the growth in photovoltaic capacity increases the amount of electricity supply during daytime, which significantly overlaps with the weekdays' peak hours and thus reduces the peak electricity prices.

As the spot market is a day-ahead market, electricity supply and demand are settled on the day before delivery. In practice, supply and demand cannot be matched ex ante for many reasons. Examples include operational outages, withdrawal of power plants due to maintenance and sudden changes in demand. While small mismatches between supply and demand are usually corrected by trading in intraday markets, larger mismatches need to be handled separately. Moreover, since wind and sunlight conditions can change abruptly and are difficult to predict with high accuracy, wind and solar production is highly volatile. Thus, as the penetration of renewable energy sources increases, the resulting fluctuations can be large, and they cannot be absorbed completely in the spot and intraday markets (Conejo et al. 2010b).

To maintain the frequency of the electricity grid at 50Hz, the imbalances between demand and supply have to be diminished. To achieve this, the transmission system operators procure ancillary services (in this case, control energy) in advance on separate markets.[II] These markets have different names in different countries, such as balancing markets, reserve markets, regulation markets, and control markets. To avoid terminological confusion, we will consistently use the term *reserve market* in the remainder of the paper. For a succinct overview of how the reserve markets work and what role the hydropower producers play in these markets, we refer interested readers to Beck and Scherer (2015) as well as Hirth and Ziegenhagen (2015).

[I] https://www.eex.com

[II] The European Network of Transmission System Operators for Electricity (ENTSOE) publishes the list of transmission system operators in Europe, available at https://www.entsoe.eu.

When electricity demand and supply differ in a particular control area, the transmission system operator can request reserve market participants to increase or decrease their electricity output up to a certain capacity limit agreed a day in advance. For generation companies, the benefit of trading in the reserve markets is two-fold. First, they collect capacity prices for every unit of power made available, regardless of whether the reserve capacities are activated or not. Second, they also earn activation prices for every unit of energy that is actually requested by the transmission system operator in order to iron out unforeseen supply-demand mismatches. Since the amount of energy requested can be positive (upward regulation) or negative (downward regulation), we distinguish between reserve-up and reserve-down markets, respectively. The opportunity to participate in the reserve markets should ease the pressure on hydropower producers who are struggling to recover original profitability on the spot market because of the eroding peak/off-peak spreads.

The aim of this paper is to develop a stochastic programming framework for maximizing the total revenues of a hydropower producer trading simultaneously in both the spot and reserve markets. The resulting optimization model is computationally challenging because it involves a large number of decision stages as well as significant uncertainty in electricity prices, natural inflows into the reservoirs and reserve market activations. For example, if the planning horizon spans one year, electricity is traded daily, and the reserve markets are operated hourly, then the number of decision stages already exceeds a few thousands. Furthermore, in a system with multiple connected reservoirs, a coordinated water release and pumping policy is required because the water released from an upstream reservoir contributes to the inflows of its downstream reservoir(s).

As pointed out by Shapiro and Nemirovski (2005), multistage stochastic programs *'generically are computationally intractable already when medium-accuracy solutions are sought.'* It would appear hopeless for us to directly solve the formulated stochastic program. Inspired by Pritchard et al. (2005), we decompose the problem temporally, which in our case is achieved by subdividing the planning horizon into days. Prior to the beginning of each day, the generation company sets end-of-day water level targets for all reservoirs. These end-of-day reservoir targets are obtained by solving a yearly reservoir management problem (the *planner's problem*), which takes the form of an "inter-day" stochastic program. To gain tractability, we solve this stochastic program by linear decision rules techniques (Ben-Tal et al. 2004a, Shapiro and Nemirovski 2005, Kuhn et al. 2011). For predetermined end-of-day reservoir targets, the generation company then solves another stochastic program to determine optimal bids for both the spot and reserve markets (the *trader's problem*), while respecting the reservoir targets prescribed by the planner's problem, robustly across all possible reserve market activations. We show that, for fixed end-of-day reservoir targets, this "intra-day" stochastic program can equivalently be reduced to a tractable linear program.

The contributions of this paper may be summarized as follows.

(i) We propose a bi-layer stochastic programming framework for maximizing the revenue

of a hydropower producer who trades in the spot and reserve markets and whose power plant has arbitrary topology. Our model accounts for uncertainty in electricity spot prices, capacity prices, activation prices, natural inflows and reserve market activations. We assume that the hydropower producer wishes to be robust with respect to the reserve capacity activations, as non-compliance would incur high penalties or lead to an exclusion from the reserve markets.

(ii) Using an information restriction whereby the hydropower producer prescribes the end-of-day reservoir targets one day in advance, we prove that the inner layer of our stochastic program (the *trader's problem*) simplifies from an infinite-dimensional stochastic program with 25 stages to a finite two-stage stochastic program with only two scenarios.

(iii) By approximating the daily reservoir targets through affine decision rules, we simplify the outer layer of our stochastic program (the *planner's problem*) from an infinite-dimensional stochastic program with 365 stages to a two-stage stochastic program that can conveniently be solved via the sample average approximation.

(iv) We apply our decomposition scheme to a cascade of three reservoirs operating in the control area of the Austrian Power Grid AG (APG). Our experimental results suggest that hydropower producers can significantly increase their revenues by participating in the reserve markets.

Löhndorf et al. (2013) study the revenue maximizing operating policy of a hydropower producer that controls multiple connected reservoirs. The model has later been extended to the operation of wind farms by Wozabal and Rameseder (2020). Both models differ from ours in the following key aspects: *(i)* they only consider bidding on the spot market, whereas we allow for the simultaneous bidding on the spot and reserve markets; *(ii)* they optimize the producer's entire supply function, whereas we restrict ourselves to a single point on this function; and *(iii)* they use stochastic dual dynamic programming and approximate dynamic programming to solve the resulting problems, whereas we employ a combination of complexity reduction techniques and affine decision rules. Gauvin et al. (2017, 2018) combine robust optimization with affine decision rules to solve hydropower production models that minimize the risk of floods. To this end, Gauvin et al. (2017) develop a model that minimizes the conditional value-at-risk of the total weighted floods over the planning horizon, subject to limited water spilling capacities and water flow delays between reservoirs. Gauvin et al. (2018) minimize the expected value of the penalized aggregated flooding, assuming that the random inflows are governed by a linear time series. Carpentier et al. (2019), finally, study an energy storage problem where the degradation of batteries (controlled by a battery replacement policy) is modeled as a slow dynamics component and the amount of stored energy (controlled by a charging/discharging strategy) is modeled as a fast dynamic component, respectively. Our approach differs from theirs in that *(i)* they only consider bidding on the spot market, whereas we allow for the simultaneous bidding on the spot and reserve markets; and *(ii)* they use approximate dual dynamic programming to solve the resulting problem, whereas we employ a combination of complexity reduction techniques and affine decision rules.

The remainder of the paper unfolds as follows. Section 4.2 develops two variants of the bidding problem faced by a hydropower producer, leading to two stochastic programs that will be referred to as the *individual bidding model* and the *collective bidding model*. Section 4.3 transforms both formulations to equivalent bi-layer stochastic programs. Sections 4.4 and 4.5 present reductions for the inner layers of these bi-layer stochastic programs. Section 4.6 solves the resulting simplified bi-layer stochastic programs using a decision rule architecture. Finally, Section 4.7 applies the proposed method to a cascade of hydropower plants situated in the Salzburg region of Austria.

Notation: We denote by $\mathbf{0}$ and $\mathbf{1}$ the appropriately sized vectors of all zeros and all ones, respectively. The Hadamard product is denoted by "$\circ$" and refers to the element-wise vector multiplication. For any $x \in \mathbb{R}$, we set $x^+ = \max\{x, 0\}$ and $x^- = \max\{-x, 0\}$ such that $x = x^+ - x^-$. All random objects are defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ consisting of a sample space Ω, a σ-algebra $\mathscr{F} \subseteq 2^\Omega$ of events and a probability measure $\mathbb{P}$ on $\mathscr{F}$. For a σ-algebra $\mathscr{G} \subseteq \mathscr{F}$, we denote by $\mathscr{L}^k(\mathscr{G})$ the set of all integrable, $\mathscr{G}$-measurable functions $g : \Omega \to \mathbb{R}^k$, and if $k = 1$, we simply write $\mathscr{L}(\mathscr{G})$.

4.2 Individual and Collective Bidding Models

We consider a hydropower generation company that operates a cascade of reservoirs and trades electricity in both the spot and the secondary reserve market. We represent the topology of the interconnected reservoirs by a directed acyclic graph with a set of nodes $\mathscr{R}$ representing the reservoirs and a set of arcs $\mathscr{A} \subseteq \mathscr{R} \times \mathscr{R}$ representing the hydraulic connections between the reservoirs. We denote the cardinalities of $\mathscr{R}$ and $\mathscr{A}$ by R and A, respectively. A tuple (r, r') is an arc in $\mathscr{A}$ if r is an upstream reservoir of r' (and hence r' is a downstream reservoir of r). Without loss of generality, we assume that each arc is equipped with a generator that converts kinetic energy of water flowing downstream into electric energy and with a pump that uses electric energy to lift water upstream. Note that some of these generators and pumps may have zero capacity, which models situations in which these devices are actually absent. Without loss of generality, we further assume that the graph has a unique sink node (denoted by $\otimes$ in schematic representations and indexed by R). This dummy reservoir has infinite volume and models the discharge into a river at the end of the reservoir cascade. Accordingly, arcs discharging into this dummy reservoir may be equipped with generators, but not with pumps. In line with real hydropower plants, we assume that every reservoir (except the aforementioned dummy reservoir) has at least one outgoing arc along which water can be discharged whenever required. The topology of the cascade can be encoded conveniently through an $R \times A$ incidence matrix $\mathbf{M}$, where for each $(r, a) \in \mathscr{R} \times \mathscr{A}$ we have that

$$m_{r,a} = \begin{cases} -1 & \text{if arc } a \text{ leaves reservoir } r, \\ +1 & \text{if arc } a \text{ enters reservoir } r, \\ 0 & \text{otherwise.} \end{cases}$$

The planning horizon should span at least one year in order to account for the seasonality of electricity prices and water inflows. We partition the planning horizon into days indexed by $d \in \mathscr{D} := \{1, \ldots, D\}$, as well as into hours indexed by $t \in \mathscr{T} := \{1, \ldots, T\}$. This means that T is divisible by D, that is, the planning horizon is assumed to accommodate an integral number of days. We further denote by $H = T/D$ the number of hours per day, but we emphasize that all subsequent results remain valid for any $H \in \mathbb{N}$, *e.g.*, if we partition each day into 15 minute intervals. The present time is modeled by a fictitious hour $t = 0$ on day $d = 0$. It is followed by hour $t = 1$ on day $d = 1$, that is, the first hour of the planning horizon. For any $t, t' \in \{0\} \cup \mathscr{T}$ with $t \leq t'$, we denote the set of hours between and including t and t' as $[t, t'] = \{t, t+1, \ldots, t'\}$. Note in particular that $[t, t] = \{t\}$. Furthermore, we denote by $d(t)$ the day containing hour t and by $\mathscr{T}(d)$ the set of all hours in day d. Finally, we introduce the following functions of $d \in \mathscr{D}$ and $t \in \mathscr{T}$:

$$\uparrow(d) = \text{ last hour of day } d, \qquad \uparrow(t) = \text{ last hour of day } d(t),$$
$$\downarrow(d) = \text{ first hour of day } d, \qquad \downarrow(t) = \text{ first hour of day } d(t),$$
$$\Downarrow(d) = \text{ last hour of day } d-1, \qquad \Downarrow(t) = \text{ last hour of day } d(t)-1.$$

By slight abuse of notation, the operators $\uparrow(\cdot)$, $\downarrow(\cdot)$ and $\Downarrow(\cdot)$ are defined both on $\mathscr{D}$ and $\mathscr{T}$. The correct interpretation of these operators will always be clear from the context. Observe also that the two $\Downarrow(\cdot)$ operators are not strictly needed because they can be expressed in terms of $\downarrow(\cdot)$ via the identities $\Downarrow(d) = \uparrow(d-1)$ and $\Downarrow(t) = \uparrow(d(t)-1)$. However, they are useful to simplify notation.

Each reservoir is characterized by—potentially time-dependent—lower and upper bounds on its filling level. The time dependence of these bounds may reflect changing safety margins that are imposed to account for the seasonality of inflows, scheduled maintenance work or environmental regulations. For each hour $t \in \mathscr{T}$, we collect the lower and upper reservoir bounds in the vectors $\underline{\boldsymbol{w}}_t \in \mathbb{R}_+^R$ [m^3] and $\overline{\boldsymbol{w}}_t \in (\mathbb{R}_+ \cup \{+\infty\})^R$ [m^3], respectively. The bounds for the dummy reservoir are simply given by $\underline{w}_{t,R} = 0$ and $\overline{w}_{t,R} = +\infty$. Similarly, the generator and the pump installed along a given arc are characterized by upper bounds on the generation and the pumping flow, respectively, (the lower bounds are trivially zero) and by their efficiencies, which may also depend on time. For any $t \in \mathscr{T}$, we collect the upper bounds on the hourly water flows through the generators and pumps in the vectors $\overline{\boldsymbol{g}}_t \in \mathbb{R}_+^A$ [m^3] and $\overline{\boldsymbol{p}}_t \in \mathbb{R}_+^A$ [m^3], respectively. Also, for any $t \in \mathscr{T}$, we group the generator efficiencies in the vector $\boldsymbol{\eta}_t \in \mathbb{R}_+^A$ [MWh/m^3] and the *inverse* pump efficiencies in the vector $\boldsymbol{\zeta}_t \in \mathbb{R}_+^A$ [MWh/m^3]. The laws of thermodynamics imply that $\boldsymbol{\eta}_t < \boldsymbol{\zeta}_t$ for all $t \in \mathscr{T}$.

The decision problem of the hydropower generation company is affected by three types of risk factors, namely fluctuating market prices, uncertain reserve activations as well as natural inflows into the reservoirs caused by unpredictable meteorological phenomena. We assume that these risk factors are exogenous (*i.e.*, the company has no means to impact them) and that they have finite expectation. Specifically, we denote the spot prices by $\pi_t^{\mathrm{s}} \in \mathbb{R}_+$ [\$/MWh] and the capacity prices for up- and down-regulation by $\pi_t^{\mathrm{u}} \in \mathbb{R}_+$ [\$/MWh] and $\pi_t^{\mathrm{v}} \in \mathbb{R}_+$ [\$/MWh],

respectively. We assume that these prices may change from one hour to the next but are revealed one day in advance at time $\Downarrow(t)$. This is an approximation to reality, where the markets are cleared during the day. We make it to avoid clutter. In case of a call-off on the reserve market, the producer will receive a financial compensation equal to the activation price $\psi_t^{\mathrm{u}} \in \mathbb{R}_+$ [\$/MWh] or $\psi_t^{\mathrm{v}} \in \mathbb{R}_+$ [\$/MWh] multiplied with the amount of energy requested for up- or down-regulation, respectively. The activation prices are revealed in real time. The uncertain reserve activations, which are also revealed in real time, are modeled by Bernoulli random variables ρ_t^{u} and ρ_t^{v}, which evaluate to 0 if there is no call-off and to 1 if there is a call-off on the reserve-up or the reserve-down market, respectively. We assume that there is never a simultaneous call-off on the reserve-up and the reserve-down market, that is, ρ_t^{u} and ρ_t^{v} cannot evaluate to 1 at the same time. This assumption is reasonable as long as the power generation company sells electricity in a single bidding zone that may either face an over-supply or an under-supply of electricity, but not both at the same time. We further assume that the support of the reserve activations at time t, conditional on the history up to time $t-1$, is independent of any past information. In other words, regardless of the past, there is always a strictly positive probability that the tuple $(\rho_t^{\mathrm{u}}, \rho_t^{\mathrm{v}})$ may take any of the values in $\{(0,0),(1,0),(0,1)\}$. Moreover, we assume that the reserve activations are serially independent and independent of all other exogenous uncertainties. This is justified because reserves should only be activated in case of an unexpected event. Finally, we denote the natural inflows into the various reservoirs by the vector $\boldsymbol{\phi}_t \in \mathbb{R}_+^R$ [m^3], and we assume that they are revealed a day in advance at time $\Downarrow(t)$ thanks to an accurate forecast.

We gather all random variables whose values are revealed in hour $t \in \{0\} \cup \mathscr{T}$ in the vector

$$\boldsymbol{\xi}_t = \begin{cases} (\{(\pi_\tau^{\mathrm{s}}, \pi_\tau^{\mathrm{u}}, \pi_\tau^{\mathrm{v}}, \boldsymbol{\phi}_\tau)\}_{\tau=1}^{H}) & \text{if } t = 0, \\ (\psi_t^{\mathrm{u}}, \psi_t^{\mathrm{v}}, \rho_t^{\mathrm{u}}, \rho_t^{\mathrm{v}}, \{(\pi_\tau^{\mathrm{s}}, \pi_\tau^{\mathrm{u}}, \pi_\tau^{\mathrm{v}}, \boldsymbol{\phi}_\tau)\}_{\tau=t+1}^{t+H}) & \text{if } t \in \Downarrow(\mathscr{T}) \setminus 0, \\ (\psi_t^{\mathrm{u}}, \psi_t^{\mathrm{v}}, \rho_t^{\mathrm{u}}, \rho_t^{\mathrm{v}}) & \text{otherwise.} \end{cases}$$

Similarly, we gather all random variables whose values are revealed within the interval $[t, t']$ in the vector $\boldsymbol{\xi}_{[t,t']} = (\boldsymbol{\xi}_t, \ldots, \boldsymbol{\xi}_{t'})$. Furthermore, we capture the information revealed during this time interval by the σ-algebra $\mathscr{F}_{[t,t']} = \sigma(\boldsymbol{\xi}_{[t,t']})$, *i.e.*, the σ-algebra generated by $\boldsymbol{\xi}_{[t,t']}$. For ease of notation, when $t = 0$, we simply write $\boldsymbol{\xi}_{[t']}$ and $\mathscr{F}_{[t']}$, respectively.

The hydropower generation company must decide each day how much energy to offer on the different markets in every hour of the following day. This in turn necessitates a strategy for operating the reservoir system that is guaranteed to honor all market commitments. There are two complementary paradigms for bidding on the different markets, which we refer to as *individual bidding* and as *collective bidding*. To understand their difference, recall that energy is produced by the generators and consumed by the pumps installed along the arcs. Under the individual bidding paradigm, the company places an individual bid for every arc of its reservoir system. We gather these individual (arc-specific) bids placed on the spot market, the reserve-up market and the reserve-down market in the vectors $\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t \in$

$\mathscr{L}^A(\mathscr{F}_{[\Downarrow(t)]})$ [MWh], respectively. Under the collective bidding paradigm, on the other hand, the hydropower company places a single bid on each market but does not specify which arc will be used to produce (or consume) the promised amount of energy. Collective bidding is non-inferior to individual bidding because it increases the producer's flexibility in choosing the generation and pumping decisions. By a slight abuse of notation, we denote the collective bids placed on the spot market, the reserve-up market and the reserve-down market by $s_t, u_t, v_t \in \mathscr{L}(\mathscr{F}_{[\Downarrow(t)]})$ [MWh], respectively. The two bidding paradigms will give rise to two different, but intimately related optimization models, whose connections will be scrutinized in detail. In particular, the analysis of the (simpler) individual bidding model provides the basis for the analysis of the collective bidding model. Overall, however, we believe that the collective bidding model reflects the operation of real power plants more faithfully and may thus more useful in practice. Both models involve the water flows along the arcs as well as the reservoir filling levels as real-time operational decisions. Specifically, we denote the generation flows, *i.e.*, the amounts of water running through the turbines in each hour, by $\boldsymbol{g}_t \in \mathscr{L}^A(\mathscr{F}_{[t]})$ [m^3] and the consumption flows, *i.e.*, the amounts of water running through the pumps in each hour, by $\boldsymbol{p}_t \in \mathscr{L}^A(\mathscr{F}_{[t]})$ [m^3]. Furthermore, we assume that one can spill unlimited amounts of water $\boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[t]})$ [m^3] along all arcs from the upstream reservoirs to the downstream reservoirs. The reservoir filling levels at the end of each hour are denoted by $\boldsymbol{w}_t \in \mathscr{L}^R(\mathscr{F}_{[t]})$ [m^3].

We assume that the hydropower generation company is risk-neutral but aims to satisfy all constraints robustly—especially those related to reserve market commitments. Thus, its objective is to maximize the expected cumulative revenues across the entire planning horizon while maintaining its ability to respond to all possible call-offs on the reserve markets. In the context of the individual bidding model, the objective function thus takes the form

$$\sum_{t\in\mathscr{T}} \mathbb{E}\big[\pi_t^{\mathrm{s}}\mathbf{1}^\top \boldsymbol{s}_t + (\pi_t^{\mathrm{u}} + \rho_t^{\mathrm{u}}\psi_t^{\mathrm{u}})\mathbf{1}^\top \boldsymbol{u}_t + (\pi_t^{\mathrm{v}} + \rho_t^{\mathrm{v}}\psi_t^{\mathrm{v}})\mathbf{1}^\top \boldsymbol{v}_t\big].$$

The non-anticipativity constraints to be respected are

$$\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t \in \mathscr{L}^A(\mathscr{F}_{[\Downarrow(t)]}),\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[t]}),\ \boldsymbol{w}_t \in \mathscr{L}^R(\mathscr{F}_{[t]}) \qquad \forall\, t \in \mathscr{T}.$$

While spot market bids have no sign restrictions (*i.e.*, the company can both buy or sell electricity in the spot market), bids in the reserve markets as well as the operational flow decisions must be non-negative. Also, the flow decisions must obey the capacity limits of the respective pipes. Furthermore, the reservoir levels must stay within the respective bounds. Thus, we require

$$\mathbf{0} \le \boldsymbol{u}_t,\ \mathbf{0} \le \boldsymbol{v}_t,\ \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t,\ \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_t \le \overline{\boldsymbol{w}}_t \qquad \forall\, t \in \mathscr{T},\ \mathbb{P}\text{-a.s.}$$

In addition, the hydropower company is obliged to produce and/or consume energy according to its market commitments. If the same system operator is responsible for both the spot and the reserve markets, then the market commitments are enforced through the following

constraint

$$\boldsymbol{s}_t + \rho_t^{\mathrm{u}} \boldsymbol{u}_t - \rho_t^{\mathrm{v}} \boldsymbol{v}_t = \boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t \qquad \forall t \in \mathscr{T},\ \mathbb{P}\text{-a.s.}$$

Taking into account the natural inflows, the reservoir filling levels obey the dynamic equation

$$\boldsymbol{w}_t = \boldsymbol{w}_{t-1} + \boldsymbol{\phi}_t + \mathbf{M}(\boldsymbol{g}_t - \boldsymbol{p}_t + \boldsymbol{z}_t) \qquad \forall t \in \mathscr{T},\ \mathbb{P}\text{-a.s.}$$

Finally, in order to prevent the eventual depletion of the reservoirs at the end of the planning horizon, the hydropower company imposes the following terminal water level constraint

$$\boldsymbol{w}_0 \leq \boldsymbol{w}_T \qquad \mathbb{P}\text{-a.s.}$$

In summary, if the company decides to adopt the individual bidding model, it must solve the stochastic optimization problem

$$\begin{array}{lll}
\sup & \sum_{t\in\mathscr{T}} \mathbb{E}\big[\pi_t^{\mathrm{s}} \mathbf{1}^\top \boldsymbol{s}_t + (\pi_t^{\mathrm{u}} + \rho_t^{\mathrm{u}} \psi_t^{\mathrm{u}}) \mathbf{1}^\top \boldsymbol{u}_t + (\pi_t^{\mathrm{v}} + \rho_t^{\mathrm{v}} \psi_t^{\mathrm{v}}) \mathbf{1}^\top \boldsymbol{v}_t\big] & \\
\text{s.t.} & \boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t \in \mathscr{L}^A(\mathscr{F}_{[\Downarrow(t)]}),\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[t]}),\ \boldsymbol{w}_t \in \mathscr{L}^R(\mathscr{F}_{[t]}) & \forall t \in \mathscr{T} \\
& \mathbf{0} \leq \boldsymbol{u}_t,\ \mathbf{0} \leq \boldsymbol{v}_t,\ \mathbf{0} \leq \boldsymbol{g}_t \leq \overline{\boldsymbol{g}}_t,\ \mathbf{0} \leq \boldsymbol{p}_t \leq \overline{\boldsymbol{p}}_t,\ \mathbf{0} \leq \boldsymbol{z}_t & \forall t \in \mathscr{T},\ \mathbb{P}\text{-a.s.} \\
& \boldsymbol{s}_t + \rho_t^{\mathrm{u}} \boldsymbol{u}_t - \rho_t^{\mathrm{v}} \boldsymbol{v}_t = \boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t & \forall t \in \mathscr{T},\ \mathbb{P}\text{-a.s.} \\
& \boldsymbol{w}_t = \boldsymbol{w}_{t-1} + \boldsymbol{\phi}_t + \mathbf{M}(\boldsymbol{g}_t - \boldsymbol{p}_t + \boldsymbol{z}_t) & \forall t \in \mathscr{T},\ \mathbb{P}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_t \leq \boldsymbol{w}_t \leq \overline{\boldsymbol{w}}_t,\ \boldsymbol{w}_0 \leq \boldsymbol{w}_T & \forall t \in \mathscr{T},\ \mathbb{P}\text{-a.s.}
\end{array} \tag{I}$$

Conversely, in the context of the collective bidding model, the objective function takes the form

$$\sum_{t\in\mathscr{T}} \mathbb{E}\big[\pi_t^{\mathrm{s}} s_t + (\pi_t^{\mathrm{u}} + \rho_t^{\mathrm{u}} \psi_t^{\mathrm{u}}) u_t + (\pi_t^{\mathrm{v}} + \rho_t^{\mathrm{v}} \psi_t^{\mathrm{v}}) v_t\big],$$

and the non-anticipativity constraints to be respected are

$$s_t, u_t, v_t \in \mathscr{L}(\mathscr{F}_{[\Downarrow(t)]}),\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[t]}),\ \boldsymbol{w}_t \in \mathscr{L}^R(\mathscr{F}_{[t]}) \qquad \forall t \in \mathscr{T}.$$

As in the individual bidding model, all decision variables must reside within their bounds, that is,

$$0 \leq u_t,\ 0 \leq v_t,\ \mathbf{0} \leq \boldsymbol{g}_t \leq \overline{\boldsymbol{g}}_t,\ \mathbf{0} \leq \boldsymbol{p}_t \leq \overline{\boldsymbol{p}}_t,\ \mathbf{0} \leq \boldsymbol{z}_t,\ \underline{\boldsymbol{w}}_t \leq \boldsymbol{w}_t \leq \overline{\boldsymbol{w}}_t \qquad \forall t \in \mathscr{T},\ \mathbb{P}\text{-a.s.}$$

As the hydropower company now only places aggregate bids, the market commitments are enforced through a scalar (instead of a vector-valued) constraint of the form

$$s_t + \rho_t^{\mathrm{u}} u_t - \rho_t^{\mathrm{v}} v_t = \boldsymbol{\eta}_t^\top \boldsymbol{g}_t - \boldsymbol{\zeta}_t^\top \boldsymbol{p}_t \qquad \forall t \in \mathscr{T},\ \mathbb{P}\text{-a.s.}$$

The reservoir dynamics and termination constraints remain unchanged. Thus, if the company decides to adopt the collective bidding model, it must solve the stochastic optimization

problem

$$
\begin{array}{lll}
\sup & \sum_{t\in\mathscr{T}} \mathbb{E}\big[\pi^{\mathrm{s}}_t s_t + (\pi^{\mathrm{u}}_t + \rho^{\mathrm{u}}_t \psi^{\mathrm{u}}_t) u_t + (\pi^{\mathrm{v}}_t + \rho^{\mathrm{v}}_t \psi^{\mathrm{v}}_t) v_t\big] & \\
\text{s.t.} & s_t, u_t, v_t \in \mathscr{L}(\mathscr{F}_{[\Downarrow(t)]}),\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[t]}),\ \boldsymbol{w}_t \in \mathscr{L}^R(\mathscr{F}_{[t]}) & \forall\, t \in \mathscr{T} \\
& 0 \le u_t,\ 0 \le v_t,\ \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t & \forall\, t \in \mathscr{T},\ \mathbb{P}\text{-a.s.} \\
& s_t + \rho^{\mathrm{u}}_t u_t - \rho^{\mathrm{v}}_t v_t = \boldsymbol{\eta}_t^\top \boldsymbol{g}_t - \boldsymbol{\zeta}_t^\top \boldsymbol{p}_t & \forall\, t \in \mathscr{T},\ \mathbb{P}\text{-a.s.} \\
& \boldsymbol{w}_t = \boldsymbol{w}_{t-1} + \boldsymbol{\phi}_t + \mathbf{M}(\boldsymbol{g}_t - \boldsymbol{p}_t + \boldsymbol{z}_t) & \forall\, t \in \mathscr{T},\ \mathbb{P}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_t \le \overline{\boldsymbol{w}}_t,\ \boldsymbol{w}_0 \le \boldsymbol{w}_T & \forall\, t \in \mathscr{T},\ \mathbb{P}\text{-a.s.}
\end{array}
\tag{C}
$$

4.3 Planner-Trader Decomposition

We will demonstrate here that both problem (I) and problem (C) can be decomposed into subproblems that lend themselves to further simplification. To this end, we first establish a basic property of the topology matrix $\mathbf{M}$, which will be crucial for the derivations below.

Lemma 4.1. For any $\boldsymbol{h} \in \mathbb{R}^{R-1}$, $\boldsymbol{h} \le \mathbf{0}$, there exists $\boldsymbol{z} \in \mathbb{R}^A$, $\boldsymbol{z} \ge \mathbf{0}$, such that $\mathbf{M}\boldsymbol{z} = [\boldsymbol{h}^\top, -\mathbf{1}^\top \boldsymbol{h}]^\top$.

As every arc $a \in \mathscr{A}$ connects an upstream and a downstream reservoir, every column of $\mathbf{M}$ has exactly two non-zero elements that are equal to $+1$ and -1, respectively. Hence, $\mathbf{1}^\top \mathbf{M}\boldsymbol{z} = \mathbf{0}^\top \boldsymbol{z} = 0$ for any $\boldsymbol{z} \in \mathbb{R}^A$, that is, not all elements of the $\mathbf{M}\boldsymbol{z}$ can have the same sign. Lemma 4.1 strengthens this elementary result by asserting that we can always find $\boldsymbol{z} \ge \mathbf{0}$ such that the first $R-1$ components of $\mathbf{M}\boldsymbol{z}$ match a prescribed non-positive vector, while only the last component of $\mathbf{M}\boldsymbol{z}$ is non-negative.

Proof of Lemma 4.1. We recall that the dummy reservoir R represents the unique sink of our cascade topology $(\mathscr{R}, \mathscr{A})$. For any reservoir $r \in \mathscr{R} \setminus \{R\}$, there thus exists a simple directed path from r to R with arcs $\mathscr{A}^r \subseteq \mathscr{A}$. For each such reservoir r, choose a vector $\boldsymbol{z}^r \in \mathbb{R}^A_+$ such that $z^r_a = -h_r$ if $a \in \mathscr{A}^r$; $= 0$ otherwise. It follows from the definition of the incidence matrix $\mathbf{M}$ that

$$
[\mathbf{M}\boldsymbol{z}^r]_i \;=\; \sum_{a\in\mathscr{A}} m_{i,a} z^r_a \;=\; \sum_{a\in\mathscr{A}^r} m_{i,a} z^r_a \;=\; \begin{cases} +h_r & \text{if } i = r \\ -h_r & \text{if } i = R \\ 0 & \text{otherwise} \end{cases} \qquad \forall\, i \in \mathscr{R}.
$$

Indeed, there is a single arc $a \in \mathscr{A}^r$ that leaves $i = r$, which implies that $m_{i,a}(-h_r) = (-1)(-h_r)$. Likewise, there is a unique arc $a \in \mathscr{A}^r$ that enters $i = R$, which implies that $m_{i,a}(-h_r) = (+1)(-h_r)$. For intermediate nodes on the path from r to R, the terms $(-1)(-h_r)$ and $(+1)(-h_r)$ corresponding to the entering and the leaving arc cancel each other out. If i is not contained in the path from r to R, finally, then $m_{i,a}$ is zero for all $a \in \mathscr{A}^r$. We thus conclude that $\mathbf{M}\boldsymbol{z}^r = h_r \mathbf{e}_r - h_r \mathbf{e}_R$, where $\{\mathbf{e}_i\}_{i=1}^R$ denotes the canonical basis for $\mathbb{R}^R$, and hence

$$
\mathbf{M}\boldsymbol{z} \;=\; \mathbf{M} \sum_{r=1}^{R-1} \boldsymbol{z}^r \;=\; \sum_{r=1}^{R-1} (h_r \mathbf{e}_r - h_r \mathbf{e}_R) \;=\; [\boldsymbol{h}^\top, -\mathbf{1}^\top \boldsymbol{h}]^\top
$$

for $\boldsymbol{z} = \sum_{r=1}^{R-1} \boldsymbol{z}^r$. Since $\boldsymbol{z} \geq \mathbf{0}$ by construction, this concludes the proof. □

We are now ready to reformulate the individual bidding model (I) as a coupled bi-layer stochastic program, where the outer layer problem (the *planner's problem*) optimizes over the reservoir filling levels with a daily granularity, whereas the inner layer problem (the *trader's problem*) optimizes over the bidding decisions with an hourly granularity. To this end, recall that the decision variables $\{\boldsymbol{w}_t\}_{t\in\mathscr{T}}$ in problem (I) represent the end-of-hour filling levels of the reservoirs. In order to reformulate (I), we will overload notation and denote by $\{\boldsymbol{w}_d\}_{d\in\mathscr{D}}$, $\{\overline{\boldsymbol{w}}_d\}_{d\in\mathscr{D}}$ and $\{\underline{\boldsymbol{w}}_d\}_{d\in\mathscr{D}}$ the end-of-day reservoir filling levels and their lower and upper bounds, respectively, where $\boldsymbol{w}_d = \boldsymbol{w}_{\uparrow(d)}$, $\underline{\boldsymbol{w}}_d = \underline{\boldsymbol{w}}_{\uparrow(d)}$ and $\overline{\boldsymbol{w}}_d = \overline{\boldsymbol{w}}_{\uparrow(d)}$ for all $d \in \mathscr{D}$. Next, we use the reservoir balance constraints to re-express all end-of-hour filling levels $\{\boldsymbol{w}_t\}_{t\in\mathscr{T}}$ in terms of the end-of-day filling levels $\{\boldsymbol{w}_d\}_{d\in\mathscr{D}}$, the hourly flow decisions $\{\boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t\}_{t\in\mathscr{T}}$ and the natural inflows $\{\boldsymbol{\phi}_t\}_{t\in\mathscr{T}}$. Specifically, for every hour $t \in \mathscr{T}(d)$ on day $d \in \mathscr{D}$ we obtain

$$\boldsymbol{w}_t = \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau).$$

Using this relation, the individual bidding model (I) can be recast equivalently as

$$\begin{array}{lll}
\sup & \sum_{t\in\mathscr{T}} \mathbb{E}\left[\pi_t^{\mathrm{s}} \mathbf{1}^\top \boldsymbol{s}_t + (\pi_t^{\mathrm{u}} + \rho_t^{\mathrm{u}}\psi_t^{\mathrm{u}})\mathbf{1}^\top \boldsymbol{u}_t + (\pi_t^{\mathrm{v}} + \rho_t^{\mathrm{v}}\psi_t^{\mathrm{v}})\mathbf{1}^\top \boldsymbol{v}_t\right] & \\
\text{s.t.} & \boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t \in \mathscr{L}^A(\mathscr{F}_{[\downarrow(t)]}),\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[t]}),\ \boldsymbol{w}_d \in \mathscr{L}^R(\mathscr{F}_{[\uparrow(d)]}) & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d) \\
& \mathbf{0} \leq \boldsymbol{u}_t,\ \mathbf{0} \leq \boldsymbol{v}_t,\ \mathbf{0} \leq \boldsymbol{g}_t \leq \overline{\boldsymbol{g}}_t,\ \mathbf{0} \leq \boldsymbol{p}_t \leq \overline{\boldsymbol{p}}_t,\ \mathbf{0} \leq \boldsymbol{z}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& \boldsymbol{s}_t + \rho_t^{\mathrm{u}}\boldsymbol{u}_t - \rho_t^{\mathrm{v}}\boldsymbol{v}_t = \boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_t \leq \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \leq \overline{\boldsymbol{w}}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& \boldsymbol{w}_d = \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) & \forall d \in \mathscr{D},\ \mathbb{P}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_d \leq \boldsymbol{w}_d \leq \overline{\boldsymbol{w}}_d,\ \boldsymbol{w}_0 \leq \boldsymbol{w}_D & \forall d \in \mathscr{D},\ \mathbb{P}\text{-a.s.}
\end{array} \tag{4.1}$$

Note that the end-of-day reservoir bounds in the last line are implied by the end-of-hour reservoir bounds two lines above and are thus redundant. However, they will be instrumental for proving the tightness of a relaxation of problem (4.1) to be derived below. Next, we introduce another auxiliary problem that imposes the daily reservoir balance constraints as inequalities rather than equalities.

$$\begin{array}{lll}
\sup & \sum_{t\in\mathscr{T}} \mathbb{E}\left[\pi_t^{\mathrm{s}} \mathbf{1}^\top \boldsymbol{s}_t + (\pi_t^{\mathrm{u}} + \rho_t^{\mathrm{u}}\psi_t^{\mathrm{u}})\mathbf{1}^\top \boldsymbol{u}_t + (\pi_t^{\mathrm{v}} + \rho_t^{\mathrm{v}}\psi_t^{\mathrm{v}})\mathbf{1}^\top \boldsymbol{v}_t\right] & \\
\text{s.t.} & \boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t \in \mathscr{L}^A(\mathscr{F}_{[\downarrow(t)]}),\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[t]}),\ \boldsymbol{w}_d \in \mathscr{L}^R(\mathscr{F}_{[\uparrow(d)]}) & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d) \\
& \mathbf{0} \leq \boldsymbol{u}_t,\ \mathbf{0} \leq \boldsymbol{v}_t,\ \mathbf{0} \leq \boldsymbol{g}_t \leq \overline{\boldsymbol{g}}_t,\ \mathbf{0} \leq \boldsymbol{p}_t \leq \overline{\boldsymbol{p}}_t,\ \mathbf{0} \leq \boldsymbol{z}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& \boldsymbol{s}_t + \rho_t^{\mathrm{u}}\boldsymbol{u}_t - \rho_t^{\mathrm{v}}\boldsymbol{v}_t = \boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_t \leq \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \leq \overline{\boldsymbol{w}}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& \boldsymbol{w}_d \leq \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) & \forall d \in \mathscr{D},\ \mathbb{P}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_d \leq \boldsymbol{w}_d \leq \overline{\boldsymbol{w}}_d,\ \boldsymbol{w}_0 \leq \boldsymbol{w}_D & \forall d \in \mathscr{D},\ \mathbb{P}\text{-a.s.}
\end{array} \tag{4.2}$$

The next proposition asserts that the optimization problems (4.1) and (4.2) are indeed equivalent. Specifically, we will show that any slacks between the left- and right-hand sides of the daily reservoir balance constraints can systematically be eliminated by spillage. The existence of suitable spillage decisions that make all the reservoir balance constraints binding is guaranteed by Lemma 4.1.

Proposition 4.2. The optimal values of problems (4.1) and (4.2) are equal.

Proof. As problem (4.2) is a relaxation of problem (4.1), it suffices to show that any feasible solution of problem (4.2) corresponds to a feasible solution of problem (4.1) that attains the same objective function value. Thus, consider an arbitrary feasible solution $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t, \boldsymbol{w}_d)\}_{(t,d)\in\mathscr{T}\times\mathscr{D}}$ of (4.2). Since the daily reservoir balance constraints are imposed as inequalities, we may introduce a slack variable $h_{d,r}$ for each day $d \in \mathscr{D}$ and for each reservoir $r \in \mathscr{R} \setminus \{R\}$, which is defined as

$$h_{d,r} = w_{d,r} - w_{d-1,r} - \textstyle\sum_{\tau\in\mathscr{T}(d)} \left[\phi_{\tau,r} + \sum_{a\in\mathscr{A}} m_{r,a}(g_{\tau,a} - p_{\tau,a} + z_{\tau,a})\right].$$

One readily verifies that $h_{d,r} \in \mathscr{L}_{[\uparrow(d)]}$ and that $h_{d,r} \leq 0$ $\mathbb{P}$-almost surely. Lemma 4.1 thus implies that for each day d there exists $\hat{\boldsymbol{z}}_d \in \mathscr{L}^A_{[\uparrow(d)]}$ with

$$\hat{\boldsymbol{z}}_d \geq \mathbf{0} \ \ \mathbb{P}\text{-a.s.} \quad \text{and} \quad \mathbf{M}\hat{\boldsymbol{z}}_d = [\boldsymbol{h}_d^\top, -\mathbf{1}^\top \boldsymbol{h}_d]^\top \ \ \mathbb{P}\text{-a.s.}$$

Next, we show that the solution $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}'_t, \boldsymbol{w}'_d)\}_{(t,d)\in\mathscr{T}\times\mathscr{D}}$ involving the spillage decisions

$$\boldsymbol{z}'_t = \begin{cases} \boldsymbol{z}_t & \text{if } t \notin \uparrow(\mathscr{D}) \\ \boldsymbol{z}_t + \hat{\boldsymbol{z}}_{d(t)} & \text{if } t \in \uparrow(\mathscr{D}) \end{cases}$$

for all $t \in \mathscr{T}$ and the end-of-day reservoir filling levels

$$w'_{d,r} = \begin{cases} w_{d,r} & \text{if } r \neq R \\ \sum_{\tau=1}^{\uparrow(d)} \left[\phi_{\tau,R} + \sum_{a\in\mathscr{A}} m_{R,a}(g_{\tau,a} + z_{\tau,a})\right] - \sum_{d'=1}^{d} \mathbf{1}^\top \boldsymbol{h}_{d'} & \text{if } r = R \end{cases}$$

for all $d \in \mathscr{D}$ and $r \in \mathscr{R}$ is feasible in (4.1). To see this, note that $\boldsymbol{z}'_t$ and $\boldsymbol{w}'_d$ are non-anticipative and $\mathbb{P}$-almost surely non-negative because the dummy reservoir has no children, that is, $m_{R,a} \in \{0, +1\}$ for all $a \in \mathscr{A}$. In addition, for any reservoir $r \in \mathscr{R} \setminus \{R\}$, we have that

$$\begin{aligned} &w'_{d-1,r} + \textstyle\sum_{\tau\in\mathscr{T}(d)} \left[\phi_{\tau,r} + \sum_{a\in\mathscr{A}} m_{r,a}(g_{\tau,a} - p_{\tau,a} + z'_{\tau,a})\right] \\ &\quad = w'_{d-1,r} + \textstyle\sum_{a\in\mathscr{A}} m_{r,a}\hat{z}_{d,a} + \sum_{\tau\in\mathscr{T}(d)} \left[\phi_{\tau,r} + \sum_{a\in\mathscr{A}} m_{r,a}(g_{\tau,a} - p_{\tau,a} + z_{\tau,a})\right] \\ &\quad = w'_{d-1,r} + h_{d,r} + \textstyle\sum_{\tau\in\mathscr{T}(d)} \left[\phi_{\tau,r} + \sum_{a\in\mathscr{A}} m_{r,a}(g_{\tau,a} - p_{\tau,a} + z_{\tau,a})\right] = w'_{d,r} \end{aligned}$$

for all days $d \in \mathscr{D}$, where the three equalities follow from the construction of $\boldsymbol{z}'_t$, the properties

of $\hat{\boldsymbol{z}}_d$ and the definition of $\boldsymbol{h}_d$, respectively. Similarly, for the dummy reservoir R, we have that

$$\begin{aligned}
&w'_{d-1,R} + \textstyle\sum_{\tau\in\mathscr{T}(d)} \left[\phi_{\tau,R} + \sum_{a\in\mathscr{A}} m_{R,a}(g_{\tau,a} - p_{\tau,a} + z'_{\tau,a})\right] \\
&\quad = w'_{d-1,R} + \textstyle\sum_{a\in\mathscr{A}} m_{R,a}\hat{z}_{d,a} + \sum_{\tau\in\mathscr{T}(d)} \left[\phi_{\tau,R} + \sum_{a\in\mathscr{A}} m_{R,a}(g_{\tau,a} - p_{\tau,a} + z_{\tau,a})\right] \\
&\quad = w'_{d-1,R} - \mathbf{1}^\top \boldsymbol{h}_d + \textstyle\sum_{\tau\in\mathscr{T}(d)} \left[\phi_{\tau,R} + \sum_{a\in\mathscr{A}} m_{R,a}(g_{\tau,a} - p_{\tau,a} + z_{\tau,a})\right] \\
&\quad = w'_{d-1,R} - \mathbf{1}^\top \boldsymbol{h}_d + \textstyle\sum_{\tau\in\mathscr{T}(d)} \left[\phi_{\tau,R} + \sum_{a\in\mathscr{A}} m_{R,a}(g_{\tau,a} + z_{\tau,a})\right] = w'_{d,R}
\end{aligned}$$

for all day $d \in \mathscr{D}$, where the third equality holds because no arc incident on reservoir R accommodates a pump, which implies that $m_{R,a}p_{\tau,a} = 0$ for all $a \in \mathscr{A}$. Hence, the new solution satisfies all daily reservoir balance constraints as equalities. Finally, we need to ascertain that the filling levels $\{w'_{d,R}\}_{d\in\mathscr{D}}$ of the dummy reservoir R still respect the hourly reservoir bounds. However, this is trivially the case because $\underline{w}_{t,R} = 0$ and $\overline{w}_{t,R} = +\infty$, regardless of the hour t. □

We are now ready to decompose the individual bidding problem (I), which jointly maximizes over all hourly and daily decisions, into planning and trading subproblems that maximize only over the daily and the hourly decisions, respectively. The proposed decomposition assumes that a fictitious planner and a fictitious trader collaborate so solve problem (I) in the following manner. The planner determines the end-of-day reservoir levels by solving the stochastic program

$$\begin{array}{lll}
\sup & \sum_{d\in\mathscr{D}} \mathbb{E}\big[\Pi^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d, \boldsymbol{\xi}_{[\Downarrow(d)]})\big] & \\
\text{s.t.} & \boldsymbol{w}_d \in \mathscr{L}^R(\mathscr{F}_{[\uparrow(d)]}) & \forall\, d \in \mathscr{D} \\
& \underline{\boldsymbol{w}}_d \le \boldsymbol{w}_d \le \overline{\boldsymbol{w}}_d & \forall\, d \in \mathscr{D},\ \mathbb{P}\text{-a.s.} \\
& \boldsymbol{w}_0 \le \boldsymbol{w}_D & \mathbb{P}\text{-a.s.},
\end{array} \tag{IP}$$

which maximizes the expected daily profits accrued over the entire planning horizon. Here, the function $\Pi^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d, \boldsymbol{\xi}_{[\Downarrow(d)]})$ represents the expected profit earned by the trader on day d via individual bidding, conditional on the information $\boldsymbol{\xi}_{[\Downarrow(d)]}$ available at the beginning of the day and conditional on the initial and terminal reservoir levels $\boldsymbol{w}_{d-1}$ and $\boldsymbol{w}_d$, respectively, imposed by the planner. The function $\Pi^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d, \boldsymbol{\xi}_{[\Downarrow(d)]})$ evaluates the optimal value of the stochastic program

$$\begin{array}{lll}
\sup & \mathbb{E}\big[\sum_{t\in\mathscr{T}(d)} \pi^{\mathrm{s}}_t \mathbf{1}^\top \boldsymbol{s}_t + (\pi^{\mathrm{u}}_t + \rho^{\mathrm{u}}_t\psi^{\mathrm{u}}_t)\mathbf{1}^\top \boldsymbol{u}_t + (\pi^{\mathrm{v}}_t + \rho^{\mathrm{v}}_t\psi^{\mathrm{v}}_t)\mathbf{1}^\top \boldsymbol{v}_t \,\big|\, \boldsymbol{\xi}_{[\Downarrow(d)]}\big] & \\
\text{s.t.} & \boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t \in \mathbb{R}^A,\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[\downarrow(d),t]}) & \forall\, t \in \mathscr{T}(d) \\
& \mathbf{0} \le \boldsymbol{u}_t,\ \mathbf{0} \le \boldsymbol{v}_t,\ \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t & \forall\, t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\
& \boldsymbol{s}_t + \rho^{\mathrm{u}}_t \boldsymbol{u}_t - \rho^{\mathrm{v}}_t \boldsymbol{v}_t = \boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t & \forall\, t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \le \overline{\boldsymbol{w}}_t & \forall\, t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\
& \boldsymbol{w}_d \le \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) & \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.}
\end{array} \tag{IT}$$

solved by the trader. We emphasize that the trader's problem (IT) may be infeasible for some reservoir targets $\boldsymbol{w}_{d-1}$ and $\boldsymbol{w}_d$, in which case $\Pi^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d, \boldsymbol{\xi}_{[\Downarrow(d)]})$ evaluates to $-\infty$ and thus introduces implicit constraints in the planner's problem (IP). We also highlight that problem (IT) constitutes a stochastic program with contextual information (Ban and Rudin 2019, Bertsimas and Kallus 2020), where the contextual covariates $\boldsymbol{\xi}_{[\Downarrow(d)]}$ do not directly impact the problem but impact only the (conditional) distribution of the random variables in the objective and the constraints.

The next proposition asserts that the planner-trader decomposition incurs no loss of optimality.

Proposition 4.3. The optimal values of problems (4.2) and (IP) are equal.

Proof. Observe that if the reservoir filling levels $\{\boldsymbol{w}_d\}_{d\in\mathscr{D}}$ are fixed, then the remaining decisions in problem (4.2) are no longer coupled across days. This allows us to decompose the stochastic program (4.2) into an outer maximization problem over the end-of-day reservoir filling levels $\{\boldsymbol{w}_d\}_{d\in\mathscr{D}}$ and a series of (mutually independent) inner stochastic programs maximizing over the bidding decisions $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t)\}_{t\in\mathscr{T}(d)}$ and the operational decisions $\{(\boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathscr{T}(d)}$, one for each day $d \in \mathscr{D}$. Formally, problem (4.2) is thus equivalent to the outer stochastic program

$$\begin{array}{lll} \sup & \sum_{d\in\mathscr{D}} \hat{\Pi}^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d) & \\ \text{s.t.} & \boldsymbol{w}_d \in \mathscr{L}^R(\mathscr{F}_{[\uparrow(d)]}) & \forall d \in \mathscr{D} \\ & \underline{\boldsymbol{w}}_d \le \boldsymbol{w}_d \le \overline{\boldsymbol{w}}_d & \forall d \in \mathscr{D},\ \mathbb{P}\text{-a.s.} \\ & \boldsymbol{w}_0 \le \boldsymbol{w}_D & \mathbb{P}\text{-a.s.}, \end{array}$$

where $\hat{\Pi}^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d)$ stands for the optimal value of the inner (parametric) stochastic program

$$\begin{array}{lll} \sup & \mathbb{E}\big[\sum_{t\in\mathscr{T}(d)} \pi^{\mathrm{s}}_t \mathbf{1}^\top \boldsymbol{s}_t + (\pi^{\mathrm{u}}_t + \rho^{\mathrm{u}}_t \psi^{\mathrm{u}}_t)\mathbf{1}^\top \boldsymbol{u}_t + (\pi^{\mathrm{v}}_t + \rho^{\mathrm{v}}_t \psi^{\mathrm{v}}_t)\mathbf{1}^\top \boldsymbol{v}_t\big] & \\ \text{s.t.} & \boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t \in \mathscr{L}^A(\mathscr{F}_{[\Downarrow(d)]}),\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[t]}) & \forall t \in \mathscr{T}(d) \\ & \mathbf{0} \le \boldsymbol{u}_t,\ \mathbf{0} \le \boldsymbol{v}_t,\ \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t & \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\ & \boldsymbol{s}_t + \rho^{\mathrm{u}}_t \boldsymbol{u}_t - \rho^{\mathrm{v}}_t \boldsymbol{v}_t = \boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t & \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\ & \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \le \overline{\boldsymbol{w}}_t & \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\ & \boldsymbol{w}_d \le \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) & \mathbb{P}\text{-a.s.}, \end{array} \tag{4.3}$$

which optimizes only over decisions pertaining to day d. Observe that problem (4.3) differs from the trader's problem (IT) in that it maximizes an *un*conditional expectation of the profits earned on day d and in that its decision variables may still adapt to information revealed *prior* to day d.

We will now use Lemma 4.22 from the appendix to show that $\hat{\Pi}^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d)$ coincides with the expectation of $\Pi^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d, \boldsymbol{\xi}_{[\Downarrow(d)]})$. For notational convenience, we abbreviate the random variables revealed before the beginning of day d as $\boldsymbol{\xi}^d_0 = \boldsymbol{\xi}_{[\Downarrow(d)]}$ and those revealed in each hour $h = 1, \ldots, H$ of day d by $\boldsymbol{\xi}^d_{[h]} = \boldsymbol{\xi}_{[\Downarrow(d)+h]}$. Furthermore, we gather the hourly decisions of day d in the vectors

$$\begin{aligned}
\boldsymbol{x}^d_0 &= (\{\boldsymbol{s}_\tau, \boldsymbol{u}_\tau, \boldsymbol{v}_\tau\}^{\uparrow(d)}_{\tau=\downarrow(d)}) \in \mathscr{L}^{3HA}(\mathscr{F}_{[\Downarrow(d)]}), \\
\boldsymbol{x}^d_h &= (\boldsymbol{g}_{\Downarrow(d)+h}, \boldsymbol{p}_{\Downarrow(d)+h}, \boldsymbol{z}_{\Downarrow(d)+h}) \in \mathscr{L}^{3A}(\mathscr{F}_{[\Downarrow(d)+h]}) \qquad \forall h = 1, \ldots, H.
\end{aligned}$$

Using this notation, the profit earned on day d can be expressed more concisely as

$$f^d(\boldsymbol{x}^d_0, \ldots, \boldsymbol{x}^d_H, \boldsymbol{\xi}^d_{[H]}) = \begin{cases} \sum_{t\in\mathscr{T}(d)} \pi^{\mathrm{s}}_t \mathbf{1}^\top \boldsymbol{s}_t + (\pi^{\mathrm{u}}_t + \rho^{\mathrm{u}}_t \psi^{\mathrm{u}}_t)\mathbf{1}^\top \boldsymbol{u}_t + (\pi^{\mathrm{v}}_t + \rho^{\mathrm{v}}_t \psi^{\mathrm{v}}_t)\mathbf{1}^\top \boldsymbol{v}_t & \\ \qquad\qquad \text{if } (\boldsymbol{x}^d_0, \ldots, \boldsymbol{x}^d_H) \text{ satisfies all } \mathbb{P}\text{-almost sure} & \\ \qquad\qquad \text{constraints of problem (4.3) in scenario } \boldsymbol{\xi}^d_{[H]}, & \\ -\infty \qquad\qquad \text{otherwise.} & \end{cases}$$

Problem (4.3) can thus be represented abstractly as

$$\hat{\Pi}^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d) = \begin{cases} \sup & \mathbb{E}\big[f^d(\boldsymbol{x}^d_0, \ldots, \boldsymbol{x}^d_H, \boldsymbol{\xi}^d_{[H]})\big] \\ \text{s.t.} & \boldsymbol{x}^d_0 \in \mathscr{L}^{3HA}(\mathscr{F}_{[\Downarrow(d)]}),\ \boldsymbol{x}^d_h \in \mathscr{L}^{3A}(\mathscr{F}_{[\Downarrow(d)+h]}) \qquad \forall h = 1, \ldots, H. \end{cases}$$

By the extended interchangeability principle established in Lemma 4.22 in the appendix, we thus find

$$\hat{\Pi}^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d) = \mathbb{E}\left[\begin{array}{ll} \sup & \mathbb{E}\big[f^d(\boldsymbol{x}^d_0, \ldots, \boldsymbol{x}^d_H, \boldsymbol{\xi}^d_{[H]}) \,|\, \boldsymbol{\xi}^d_0\big] \\ \text{s.t.} & \boldsymbol{x}^d_0 \in \mathbb{R}^{3HA},\ \boldsymbol{x}^d_h \in \mathscr{L}^{3A}(\mathscr{F}_{[\downarrow(d),\Downarrow(d)+h]}) \qquad \forall h = 1, \ldots, H \end{array}\right].$$

Unravelling the abbreviations shows that the stochastic program inside the expectation coincides with the trader's problem (IT), whose optimal value is given by $\Pi^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d, \boldsymbol{\xi}_{[\Downarrow(d)]})$. □

We will now show that the collective bidding model (C) admits a similar planner-trader decomposition. Using the same slight overload of notation as before, we denote the end-of-day reservoir filling levels as well as their lower and upper bounds as $\{\boldsymbol{w}_d\}_{d\in\mathscr{D}}$, $\{\underline{\boldsymbol{w}}_d\}_{d\in\mathscr{D}}$ and $\{\overline{\boldsymbol{w}}_d\}_{d\in\mathscr{D}}$, respectively. Adapting the water level constraints accordingly, model (C) admits the

equivalent reformulation

$$
\begin{array}{lll}
\sup & \sum_{t\in\mathscr{T}} \mathbb{E}\left[\pi^{\mathrm{s}}_t s_t + (\pi^{\mathrm{u}}_t + \rho^{\mathrm{u}}_t \psi^{\mathrm{u}}_t) u_t + (\pi^{\mathrm{v}}_t + \rho^{\mathrm{v}}_t \psi^{\mathrm{v}}_t) v_t\right] & \\
\text{s.t.} & s_t, u_t, v_t \in \mathscr{L}(\mathscr{F}_{[\downarrow(t)]}),\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[t]}),\ \boldsymbol{w}_d \in \mathscr{L}^R(\mathscr{F}_{[\uparrow(d)]}) & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d) \\
& 0 \le u_t,\ 0 \le v_t,\ \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& s_t + \rho^{\mathrm{u}}_t u_t - \rho^{\mathrm{v}}_t v_t = \boldsymbol{\eta}_t^\top \boldsymbol{g}_t - \boldsymbol{\zeta}_t^\top \boldsymbol{p}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \le \overline{\boldsymbol{w}}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& \boldsymbol{w}_d = \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) & \forall d \in \mathscr{D},\ \mathbb{P}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_d \le \boldsymbol{w}_d \le \overline{\boldsymbol{w}}_d,\ \boldsymbol{w}_0 \le \boldsymbol{w}_D & \forall d \in \mathscr{D},\ \mathbb{P}\text{-a.s.},
\end{array}
\tag{4.4}
$$

where the end-of-day reservoir bounds in the last line are again redundant for now. They will however be useful for strengthening the relaxation (4.5) below, which differs from (4.4) only in that the daily flow-reservation constraints are imposed as inequalities rather than equalities.

$$
\begin{array}{lll}
\sup & \sum_{t\in\mathscr{T}} \mathbb{E}\left[\pi^{\mathrm{s}}_t s_t + (\pi^{\mathrm{u}}_t + \rho^{\mathrm{u}}_t \psi^{\mathrm{u}}_t) u_t + (\pi^{\mathrm{v}}_t + \rho^{\mathrm{v}}_t \psi^{\mathrm{v}}_t) v_t\right] & \\
\text{s.t.} & s_t, u_t, v_t \in \mathscr{L}(\mathscr{F}_{[\downarrow(t)]}),\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[t]}),\ \boldsymbol{w}_d \in \mathscr{L}^R(\mathscr{F}_{[\uparrow(d)]}) & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d) \\
& 0 \le u_t,\ 0 \le v_t,\ \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& s_t + \rho^{\mathrm{u}}_t u_t - \rho^{\mathrm{v}}_t v_t = \boldsymbol{\eta}_t^\top \boldsymbol{g}_t - \boldsymbol{\zeta}_t^\top \boldsymbol{p}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \le \overline{\boldsymbol{w}}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& \boldsymbol{w}_d \le \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) & \forall d \in \mathscr{D},\ \mathbb{P}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_d \le \boldsymbol{w}_d \le \overline{\boldsymbol{w}}_d,\ \boldsymbol{w}_0 \le \boldsymbol{w}_D & \forall d \in \mathscr{D},\ \mathbb{P}\text{-a.s.}
\end{array}
\tag{4.5}
$$

The following proposition asserts that the two optimization problems (4.4) and (4.5) are equivalent.

Proposition 4.4. The optimal values of problems (4.4) and (4.5) are equal.

Proof. Observe that (4.5) is a relaxation of (4.4). Using a very similar construction as in the proof of Proposition 4.2, one can show that for every feasible solution of (4.5) one can construct a feasible solution of (4.4) that attains the same objective function value. Details are omitted for brevity. □

The relaxed collective bidding model (4.5) admits again a planner-trader decomposition. Specifically, the planner determines the end-of-day reservoir levels by solving the stochastic

program

$$\begin{array}{lll} \sup & \sum_{d\in\mathscr{D}} \mathbb{E}\big[\Pi^{\mathrm{C}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d, \boldsymbol{\xi}_{[\Downarrow(d)]})\big] & \\ \text{s.t.} & \boldsymbol{w}_d \in \mathscr{L}^R(\mathscr{F}_{[\uparrow(d)]}) & \forall d \in \mathscr{D} \\ & \underline{\boldsymbol{w}}_d \le \boldsymbol{w}_d \le \overline{\boldsymbol{w}}_d & \forall d \in \mathscr{D},\ \mathbb{P}\text{-a.s.} \\ & \boldsymbol{w}_0 \le \boldsymbol{w}_D & \mathbb{P}\text{-a.s.}, \end{array} \tag{CP}$$

where the function $\Pi^{\mathrm{C}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d, \boldsymbol{\xi}_{[\Downarrow(d)]})$ evaluates the optimal value of the stochastic program

$$\begin{array}{lll} \sup & \sum_{t\in\mathscr{T}(d)} \mathbb{E}\big[\pi^{\mathrm{s}}_t s_t + (\pi^{\mathrm{u}}_t + \rho^{\mathrm{u}}_t \psi^{\mathrm{u}}_t) u_t + (\pi^{\mathrm{v}}_t + \rho^{\mathrm{v}}_t \psi^{\mathrm{v}}_t) v_t \,\big|\, \boldsymbol{\xi}_{[\Downarrow(d)]}\big] & \\ \text{s.t.} & s_t, u_t, v_t \in \mathbb{R},\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[\downarrow(d),t]}) & \forall t \in \mathscr{T}(d) \\ & 0 \le u_t,\ 0 \le v_t,\ \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t & \forall t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\ & s_t + \rho^{\mathrm{u}}_t u_t - \rho^{\mathrm{v}}_t v_t = \boldsymbol{\eta}_t^\top \boldsymbol{g}_t - \boldsymbol{\zeta}_t^\top \boldsymbol{p}_t & \forall t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\ & \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \le \overline{\boldsymbol{w}}_t & \forall t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\ & \boldsymbol{w}_d \le \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) & \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \end{array} \tag{CT}$$

solved by the trader on day d. As in the case of the individual bidding model, the planner-trader decomposition is tight. This claim is formalized in the following proposition.

Proposition 4.5. The optimal values of problems (4.5) and (CP) are equal.

Proof. The proposition follows from an argument similar to the one in the proof of Proposition 4.3, involving Lemma 4.22 from the Appendix. Details are omitted for brevity. □

In conclusion, we emphasize that the proposed planner-trader decompositions both result in exact reformulations of the underlying individual and collective bidding models (I) and (C), respectively, and that they involve no approximations. These decompositions are useful because the trader's subproblems can be reduced to tractable linear programs under a mild information restriction. Sections 4.4 and 4.5 detail this reduction for the individual and collective bidding models, respectively.

4.4 Reduction of the Individual Trading Model

The individual trading problem (IT) constitutes an infinite-dimensional stochastic program because it optimizes over functional flow decisions. Such problems are known to be intractable in general (Dyer and Stougie 2006b, Hanasusanto et al. 2015a). In the remainder of this section

we will show, however, that the stochastic program (IT) is equivalent to a finite-dimensional linear program if the planner chooses the reservoir filling levels one day ahead of time. Note that selecting $\boldsymbol{w}_d$ at the end of day $d-1$ rather than the end of day d amounts to an information restriction whereby the planner sacrifices potentially useful information revealed during day d and thus foregoes some of the achievable expected profit. In other words, requiring that the reservoir filling levels be pre-committed a day in advance leads to a conservative approximation of the individual planning problem (IP). In Section 4.5 below we will impose the same restriction on the *collective* planning problem (CP). For later reference, we formally state this approximation below.

Approximation 4.6 (Information restriction)**.** For every $d \in \mathscr{D}$, the decision variable $\boldsymbol{w}_d$ in the planning problems (IP) and (CP) is restricted to $\mathscr{L}^R(\mathscr{F}_{[\uparrow(d-1)]}) = \mathscr{L}^R(\mathscr{F}_{[\Downarrow(d)]}) \subseteq \mathscr{L}^R(\mathscr{F}_{[\uparrow(d)]})$.

Under Approximation 4.6, the end-of-day reservoir filling level $\boldsymbol{w}_d$ becomes a measurable function of $\boldsymbol{\xi}_{[\Downarrow(d)]}$, which captures the information available to the trader at the beginning of day d. Similarly, the natural inflows $\{\boldsymbol{\phi}_t\}_{t\in\mathscr{T}(d)}$ of day d were assumed to be predictable at the beginning of the day and can therefore also be expressed as measurable functions of $\boldsymbol{\xi}_{[\Downarrow(d)]}$. Finally, the coefficients

$$\mathbb{E}\big[\pi_t^{\mathrm{s}}\mathbf{1}\,\big|\,\boldsymbol{\xi}_{[\Downarrow(d)]}\big], \quad \mathbb{E}\big[(\pi_t^{\mathrm{u}}+\rho_t^{\mathrm{u}}\psi_t^{\mathrm{u}})\mathbf{1}\,\big|\,\boldsymbol{\xi}_{[\Downarrow(d)]}\big] \quad \text{and} \quad \mathbb{E}\big[(\pi_t^{\mathrm{v}}+\rho_t^{\mathrm{v}}\psi_t^{\mathrm{v}})\mathbf{1}\,\big|\,\boldsymbol{\xi}_{[\Downarrow(d)]}\big] \quad \forall\, t\in\mathscr{T}(d)$$

of the (here-and-now) bidding decisions $\{(\boldsymbol{s}_t,\boldsymbol{u}_t,\boldsymbol{v}_t)\}_{t\in\mathscr{T}(d)}$ in the objective function of problem (IT) constitute measurable functions of $\boldsymbol{\xi}_{[\Downarrow(d)]}$ thanks to the properties of conditional expectations. We may thus conclude that, once a particular realization of $\boldsymbol{\xi}_{[\Downarrow(d)]}$ is fixed, the individual bidding problem (IT) reduces to a multistage stochastic program with $H+1$ decision stages, where the reserve activations $\{(\rho_t^{\mathrm{u}},\rho_t^{\mathrm{v}})\}_{t\in\mathscr{T}(d)}$ are the only exogenous uncertain parameters affecting the constraints, and the objective function is independent of the (wait-and-see) flow decisions $\{(\boldsymbol{g}_t,\boldsymbol{p}_t,\boldsymbol{z}_t)\}_{t\in\mathscr{T}(d)}$.

Based on these insights, we should expect that, conditional on $\boldsymbol{\xi}_{[\Downarrow(d)]}$, the wait-and-see decisions $\{(\boldsymbol{g}_t,\boldsymbol{p}_t,\boldsymbol{z}_t)\}_{t\in\mathscr{T}(d)}$ can be restricted to measurable functions of the reserve activations without sacrificing optimality. This additional information restriction gives rise to the optimization problem

$$\begin{array}{lll}
\sup & \mathbb{E}\big[\sum_{t\in\mathscr{T}(d)}\pi_t^{\mathrm{s}}\mathbf{1}^\top\boldsymbol{s}_t+(\pi_t^{\mathrm{u}}+\rho_t^{\mathrm{u}}\psi_t^{\mathrm{u}})\mathbf{1}^\top\boldsymbol{u}_t+(\pi_t^{\mathrm{v}}+\rho_t^{\mathrm{v}}\psi_t^{\mathrm{v}})\mathbf{1}^\top\boldsymbol{v}_t\,\big|\,\boldsymbol{\xi}_{[\Downarrow(d)]}\big] & \\
\text{s.t.} & \boldsymbol{s}_t,\boldsymbol{u}_t,\boldsymbol{v}_t\in\mathbb{R}^A,\ \boldsymbol{g}_t,\boldsymbol{p}_t,\boldsymbol{z}_t\in\mathscr{L}^A(\mathscr{F}^{\rho}_{[\downarrow(d),t]}) & \forall\, t\in\mathscr{T}(d)\\
& \mathbf{0}\le\boldsymbol{u}_t,\ \mathbf{0}\le\boldsymbol{v}_t,\ \mathbf{0}\le\boldsymbol{g}_t\le\overline{\boldsymbol{g}}_t,\ \mathbf{0}\le\boldsymbol{p}_t\le\overline{\boldsymbol{p}}_t,\ \mathbf{0}\le\boldsymbol{z}_t & \forall\, t\in\mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.}\\
& \boldsymbol{s}_t+\rho_t^{\mathrm{u}}\boldsymbol{u}_t-\rho_t^{\mathrm{v}}\boldsymbol{v}_t=\boldsymbol{\eta}_t\circ\boldsymbol{g}_t-\boldsymbol{\zeta}_t\circ\boldsymbol{p}_t & \forall\, t\in\mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.}\\
& \underline{\boldsymbol{w}}_t\le\boldsymbol{w}_{d-1}+\sum_{\tau=\downarrow(d)}^{t}\boldsymbol{\phi}_\tau+\mathbf{M}(\boldsymbol{g}_\tau-\boldsymbol{p}_\tau+\boldsymbol{z}_\tau)\le\overline{\boldsymbol{w}}_t & \forall\, t\in\mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.}\\
& \boldsymbol{w}_d\le\boldsymbol{w}_{d-1}+\sum_{\tau\in\mathscr{T}(d)}\boldsymbol{\phi}_\tau+\mathbf{M}(\boldsymbol{g}_\tau-\boldsymbol{p}_\tau+\boldsymbol{z}_\tau) & \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.},
\end{array} \tag{4.6}$$

where $\mathscr{F}^{\rho}_{[\downarrow(d),t]}=\sigma(\{(\rho_\tau^{\mathrm{u}},\rho_\tau^{\mathrm{v}})\}_{\tau=\downarrow(d)}^{t})$ denotes the σ-algebra generated by the reserve activations within the interval $[\downarrow(d),t]$. The following proposition establishes that problem (4.6) is indeed

equivalent to the individual trading problem (IT), in spite of the additional information restriction.

Proposition 4.7. Under Approximation 4.6, the optimal values of problems (IT) and (4.6) are equal.

Proof. Proof Throughout this proof we will fix a realization of the contextual covariates $\boldsymbol{\xi}_{[\Downarrow(d)]}$, which reflects the trader's information when solving problem (IT) for day d. Approximation 4.6 then implies that $\boldsymbol{w}_{d-1}$ and $\boldsymbol{w}_d$ reduce to deterministic constants. In addition, $\mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}$ becomes an *un*conditional distribution. Below we will denote the expectation with respect to this distribution by $\mathbb{E}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}[\cdot]$.

As problem (4.6) is a restriction of problem (IT), it suffices to show that for every feasible solution $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathscr{T}(d)}$ of (IT) there is a feasible solution $\{(\boldsymbol{s}'_t, \boldsymbol{u}'_t, \boldsymbol{v}'_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\}_{t\in\mathscr{T}(d)}$ of (4.6) that attains the same objective value. Such a solution can readily be constructed as

$$(\boldsymbol{s}'_t, \boldsymbol{u}'_t, \boldsymbol{v}'_t) = (\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t) \quad \text{and} \quad (\boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t) = \mathbb{E}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}[(\boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t) \,|\, \{(\rho^{\mathrm{u}}_\tau, \rho^{\mathrm{v}}_\tau)\}^t_{\tau=\downarrow(d)}] \quad \forall\, t \in \mathscr{T}(d).$$

The equality of objective values is immediate because the objective function only depends on the bidding decisions, which are preserved. The non-anticipativity constraints $\boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t \in \mathscr{L}^A(\mathscr{F}^{\rho}_{[\downarrow(d),t]})$ are also satisfied thanks to the defining properties of conditional expectations. To show that the almost sure constraints hold, we recall first that the reserve activations are serially independent and independent of all other sources of uncertainty. This implies that $\{(\rho^{\mathrm{u}}_\tau, \rho^{\mathrm{v}}_\tau)\}^T_{\tau=t+1}$ is independent of the non-anticipative flow decisions $\boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[\downarrow(d),t]})$ under the distribution $\mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}$, that is

$$(\boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t) = \mathbb{E}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}[(\boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t) \,|\, \{(\rho^{\mathrm{u}}_\tau, \rho^{\mathrm{v}}_\tau)\}^T_{\tau=\downarrow(d)}] \quad \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \tag{4.7}$$

for all $t \in \mathscr{T}(d)$. The feasibility of $\{(\boldsymbol{s}'_t, \boldsymbol{u}'_t, \boldsymbol{v}'_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\}_{t\in\mathscr{T}(d)}$ in (4.6) therefore follows from the feasibility of $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathscr{T}(d)}$ in (IT) and the linearity of the almost sure constraints in (4.6), which implies that they all remain valid under conditional expectations. For example, the new solution $\{(\boldsymbol{s}'_t, \boldsymbol{u}'_t, \boldsymbol{v}'_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\}_{t\in\mathscr{T}(d)}$ obeys the end-of-day reservoir bound in (4.6) because

$$\begin{aligned}
\boldsymbol{w}_d &= \mathbb{E}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\left[\boldsymbol{w}_d \,\middle|\, \{(\rho^{\mathrm{u}}_\tau, \rho^{\mathrm{v}}_\tau)\}^T_{\tau=\downarrow(d)}\right] \\
&\leq \mathbb{E}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\left[\boldsymbol{w}_{d-1} + \textstyle\sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \,\middle|\, \{(\rho^{\mathrm{u}}_\tau, \rho^{\mathrm{v}}_\tau)\}^T_{\tau=\downarrow(d)}\right] \\
&= \boldsymbol{w}_{d-1} + \textstyle\sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}'_\tau - \boldsymbol{p}'_\tau + \boldsymbol{z}'_\tau) \quad \forall\, t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{a.s.},
\end{aligned}$$

where the first equality follows from Approximation 4.6, whereby $\boldsymbol{w}_d$ is deterministic when $\boldsymbol{\xi}_{[\Downarrow(d)]}$ is kept fixed, the inequality holds because $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathscr{T}(d)}$ satisfies the upper reservoir bounds in (IT), and the last equality follows from (4.7). Thus, the claim follows. □

Proposition 4.7 asserts that solving the infinite-dimensional stochastic program (IT) is tantamount to solving the simplified stochastic program (4.6), which is in fact equivalent to a finite-dimensional yet prohibitively large linear program. To see this, recall that the reserve activations $(\rho_t^{\mathrm{u}}, \rho_t^{\mathrm{v}})$ have only three possible realizations $(0,0)$, $(1,0)$ or $(0,1)$ regardless of any past information. Thus, the σ-algebra $\mathscr{F}^{\rho}_{[\downarrow(d),t]}$ generated by all reserve activations revealed over the interval $[\downarrow(d), t]$ accommodates $3^{t-\Downarrow(t)}$ atoms. We conclude that any wait-and-see decision in the space $\mathscr{L}^A(\mathscr{F}^{\rho}_{[\downarrow(d),t]})$ is piecewise constant with up to $3^{t-\Downarrow(t)}$ pieces in $\mathbb{R}^A$ and can therefore be uniquely encoded by $A \cdot 3^{t-\Downarrow(t)}$ parameters. As the linear stochastic program (4.6) encompasses $H+1$ stages overall (one here-and-now stage and H wait-and-see stages), it is equivalent to a linear program with $\mathscr{O}(A \cdot 3^H)$ decision variables and constraints. Thus, its supremum is attained whenever the problem is feasible (which may not be the case for some end-of-day reservoir targets $\boldsymbol{w}_d$) and bounded (which is always the case because the generation and pumping capacities induce bounds on the market decisions).

We now simplify problem (4.6) in two steps and show that it has the same optimal value as a much smaller, efficiently solvable linear program. In the first step we add valid inequalities to problem (4.6), which are needed to ensure that the subsequent simplifications incur no loss of optimality.

Proposition 4.8. Any feasible solution of (4.6) satisfies $\boldsymbol{s}_t - \boldsymbol{v}_t \geq -\boldsymbol{\zeta}_t \circ \overline{\boldsymbol{p}}_t$ for all $t \in \mathscr{T}(d)$.

Proof. Any feasible solution of the stochastic program (4.6) satisfies the energy delivery constraint $\boldsymbol{s}_t + \rho_t^{\mathrm{u}} \boldsymbol{u}_t - \rho_t^{\mathrm{v}} \boldsymbol{v}_t = \boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t$ for all $t \in \mathscr{T}(d)$ $\mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}$-almost surely. Similarly, for any $t \in \mathscr{T}(d)$, we have that $\boldsymbol{g}_t \geq \boldsymbol{0}$ and $\boldsymbol{p}_t \leq \overline{\boldsymbol{p}}_t$ $\mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}$-almost surely. In combination, these inequalities imply

$$\boldsymbol{s}_t + \rho_t^{\mathrm{u}} \boldsymbol{u}_t - \rho_t^{\mathrm{v}} \boldsymbol{v}_t = \boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t \geq -\boldsymbol{\zeta}_t \circ \boldsymbol{p}_t \geq -\boldsymbol{\zeta}_t \circ \overline{\boldsymbol{p}}_t \qquad \forall\, t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.}$$

As the scenario $(\rho_t^{\mathrm{u}}, \rho_t^{\mathrm{v}}) = (0,1)$ has strictly positive probability under $\mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}$, we thus find

$$\boldsymbol{s}_t - \boldsymbol{v}_t \geq -\boldsymbol{\zeta}_t \circ \overline{\boldsymbol{p}}_t \quad \forall\, t \in \mathscr{T}(d),$$

which proves the statement. □

By Proposition 4.8, the deterministic inequalities $\boldsymbol{s}_t - \boldsymbol{v}_t \geq -\boldsymbol{\zeta}_t \circ \overline{\boldsymbol{p}}_t$ are valid for all $t \in \mathscr{T}(d)$ and can therefore be appended to problem (4.6) without affecting its feasible set. Observe that $\boldsymbol{s}_t + \boldsymbol{\zeta}_t \circ \overline{\boldsymbol{p}}_t$ comprises the arc-wise maximum bids in the reserve-down market on which the company can deliver in case of a call-off. To see this, note that $s_{t,a}$ is the amount of energy produced on arc a for the spot market and that $\zeta_{t,a}\overline{p}_{t,a}$ represents the maximum amount of energy that can be absorbed on arc a by pumping. In case of a call-off on the reserve-down market, the energy production on arc a can thus be reduced at most by $s_{t,a} + \zeta_{t,a}\overline{p}_a$.

Appending the valid inequalities to (4.6), we obtain

$$\begin{array}{lll}
\max & \mathbb{E}\big[\sum_{t\in\mathscr{T}(d)} \pi^{\mathrm{s}}_t \mathbf{1}^\top \boldsymbol{s}_t + (\pi^{\mathrm{u}}_t + \rho^{\mathrm{u}}_t \psi^{\mathrm{u}}_t)\mathbf{1}^\top \boldsymbol{u}_t + (\pi^{\mathrm{v}}_t + \rho^{\mathrm{v}}_t \psi^{\mathrm{v}}_t)\mathbf{1}^\top \boldsymbol{v}_t \,\big|\, \boldsymbol{\xi}_{[\Downarrow(d)]}\big] & \\
\text{s.t.} & \boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t \in \mathbb{R}^A,\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}^{\rho}_{[\downarrow(d),t]}) & \forall\, t \in \mathscr{T}(d) \\
& \mathbf{0} \le \boldsymbol{u}_t,\ \mathbf{0} \le \boldsymbol{v}_t,\ \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t & \forall\, t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\
& \boldsymbol{s}_t + \rho^{\mathrm{u}}_t \boldsymbol{u}_t - \rho^{\mathrm{v}}_t \boldsymbol{v}_t = \boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t & \forall\, t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\
& \boldsymbol{s}_t - \boldsymbol{v}_t \ge -\boldsymbol{\zeta}_t \circ \overline{\boldsymbol{p}}_t & \forall\, t \in \mathscr{T}(d) \\
& \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \le \overline{\boldsymbol{w}}_t & \forall\, t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\
& \boldsymbol{w}_d \le \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) & \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.},
\end{array} \tag{4.8}$$

which is equivalent to problem (4.6) by virtue of Proposition 4.8.

Next, we show that problem (4.8) is equivalent to the reduced stochastic program

$$\begin{array}{lll}
\max & \mathbb{E}\big[\sum_{t\in\mathscr{T}(d)} \pi^{\mathrm{s}}_t \mathbf{1}^\top \boldsymbol{s}_t + (\pi^{\mathrm{u}}_t + \rho^{\mathrm{u}}_t \psi^{\mathrm{u}}_t)\mathbf{1}^\top \boldsymbol{u}_t + (\pi^{\mathrm{v}}_t + \rho^{\mathrm{v}}_t \psi^{\mathrm{v}}_t)\mathbf{1}^\top \boldsymbol{v}_t \,\big|\, \boldsymbol{\xi}_{[\Downarrow(d)]}\big] & \\
\text{s.t.} & \boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t \in \mathbb{R}^A,\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathbb{R}^A & \forall\, t \in \mathscr{T}(d) \\
& \mathbf{0} \le \boldsymbol{u}_t,\ \mathbf{0} \le \boldsymbol{v}_t,\ \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t & \forall\, t \in \mathscr{T}(d) \\
& \boldsymbol{s}_t + \boldsymbol{u}_t = \boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t & \forall\, t \in \mathscr{T}(d) \\
& \boldsymbol{s}_t - \boldsymbol{v}_t \ge -\boldsymbol{\zeta}_t \circ \overline{\boldsymbol{p}}_t & \forall\, t \in \mathscr{T}(d) \\
& \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \le \overline{\boldsymbol{w}}_t & \forall\, t \in \mathscr{T}(d) \\
& \boldsymbol{w}_d \le \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau). &
\end{array} \tag{IT$^{\mathrm{r}}$}$$

Problem (IT$^{\mathrm{r}}$) involves only here-and-now decisions, and its constraints are deterministic conditional on $\boldsymbol{\xi}_{[\Downarrow(d)]}$. In fact, it constitutes a linear program with $\mathscr{O}(A \cdot H)$ decisions and constraints.

Proposition 4.9. Under Approximation 4.6, the optimal values of (4.8) and (IT$^{\mathrm{r}}$) are equal.

Proof. We will prove the proposition by showing that every feasible solution of problem (4.8) corresponds to a feasible solution of problem (IT$^{\mathrm{r}}$) with the same objective value and vice versa.

To prove the first direction, fix any feasible solution $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathscr{T}(d)}$ of problem (4.8) and an arbitrary realization of the contextual covariates $\boldsymbol{\xi}_{[\Downarrow(d)]}$. Then, there exists an event $\Omega_0 \in \mathscr{F}$ with $\mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}[\Omega_0] = 1$ such that all constraints of (4.8) are satisfied pointwise for all $\omega \in \Omega_0$. Next, select any sample $\omega_0 \in \Omega_0$ with $\rho^{\mathrm{u}}_t(\omega_0) = 1$ and $\rho^{\mathrm{v}}_t(\omega_0) = 0$ for all $t \in \mathscr{T}(d)$. Note that the set of all samples satisfying these conditions has a positive probability (and is therefore non-empty) by our assumptions about the statistics of the reserve activations. Next, define deterministic flow decisions $\boldsymbol{g}'_t = \boldsymbol{g}_t(\omega_0)$, $\boldsymbol{p}'_t = \boldsymbol{p}_t(\omega_0)$ and $\boldsymbol{z}'_t = \boldsymbol{z}_t(\omega_0)$ for all $t \in \mathscr{T}(d)$. One readily verifies that $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\}_{t\in\mathscr{T}(d)}$ is feasible in (IT$^{\mathrm{r}}$) because the constraints of the stochastic program (4.8) corresponding to scenario ω_0 coincide with the constraints of the deterministic linear program (IT$^{\mathrm{r}}$). Also, the objective value of $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\}_{t\in\mathscr{T}(d)}$

in (IT$^{\mathrm{r}}$) coincides with that of $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathscr{T}(d)}$ in (4.8) because the two solutions involve identical market bids.

To prove the opposite direction, we fix any feasible solution $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathscr{T}(d)}$ of the reduced trader's problem (IT$^{\mathrm{r}}$), and by Lemma 4.23 in the appendix we may assume without loss of generality that this solution satisfies the complementarity constraints $\boldsymbol{g}_t \circ \boldsymbol{p}_t = \boldsymbol{0}$ for all $t \in \mathscr{T}(d)$. We will argue below that one can systematically construct flow decisions $\{(\boldsymbol{g}'_t, \boldsymbol{p}'_t)\}_{t\in\mathscr{T}(d)}$ that satisfy

$$g'_{t,a}, p'_{t,a} \in \mathscr{L}(\mathscr{F}^{\rho}_{[\downarrow(d),t]}) \tag{4.9a}$$

$$0 \le g'_{t,a} \le \overline{g}_{t,a} \qquad \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \tag{4.9b}$$

$$0 \le p'_{t,a} \le \overline{p}_{t,a} \qquad \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \tag{4.9c}$$

$$\eta_{t,a} g'_{t,a} - \zeta_{t,a} p'_{t,a} = s_{t,a} + \rho^{\mathrm{u}}_t u_{t,a} - \rho^{\mathrm{v}}_t v_{t,a} \qquad \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \tag{4.9d}$$

$$g'_{t,a} - p'_{t,a} \le g_{t,a} - p_{t,a} \qquad \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \tag{4.9e}$$

for all $t \in \mathscr{T}(d)$ and $a \in \mathscr{A}$. As $g_{t,a} p_{t,a} = 0$, we can distinguish two cases for each hour-arc pair (t,a), which necessitate two different constructions of the corresponding flows $g'_{t,a}$ and $p'_{t,a}$.

Case 1 ($g_{t,a} = 0$)**:**

We set $g'_{t,a} = 0$ and $p'_{t,a} = -(s_{t,a} + \rho^{\mathrm{u}}_t u_{t,a} - \rho^{\mathrm{v}}_t v_{t,a})/\zeta_{t,a}$. It is easy to verify that $(g'_{t,a}, p'_{t,a})$ satisfies (4.9b) and (4.9d). The non-anticipativity constraints (4.9a) are also met because the reserve activations ρ^{u}_t and ρ^{v}_t are revealed in hour t. In order to establish (4.9c) and (4.9e), we first observe that the given feasible solution of the reduced trader's problem (IT$^{\mathrm{r}}$) satisfies

$$s_{t,a} + u_{t,a} = \eta_{t,a} g_{t,a} - \zeta_{t,a} p_{t,a} \quad \Longrightarrow \quad p_{t,a} = -(s_{t,a} + u_{t,a})/\zeta_{t,a},$$

where the implication holds because $g_{t,a} = 0$. As $\rho^{\mathrm{u}}_t \le 1$ and as ρ^{v}_t, $u_{t,a}$ and $v_{t,a}$ are non-negative, we may thus conclude that $p_{t,a} \le p'_{t,a}$. Hence, the constructed pumping decision $p'_{t,a}$ meets requirement (4.9e). Finally, requirement (4.9c) is satisfied because $0 \le p_{t,a} \le p'_{t,a}$ and because

$$p'_{t,a} \;=\; -(s_{t,a} + \rho^{\mathrm{u}}_t u_{t,a} - \rho^{\mathrm{v}}_t v_{t,a})/\zeta_{t,a} \;\le\; -(s_{t,a} - v_{t,a})/\zeta_{t,a} \le \overline{p}_{t,a},$$

where the second inequality follows from the valid cut derived in Proposition 4.8, which constitutes one of the constraints of the reduced trader's problem (IT$^{\mathrm{r}}$). Thus, $g'_{t,a}$ and $p'_{t,a}$ satisfy (4.9).

Case 2 ($g_{t,a} > 0$)**:**

We set $g'_{t,a} = (s_{t,a} + \rho^{\mathrm{u}}_t u_{t,a} - \rho^{\mathrm{v}}_t v_{t,a})^+/\eta_{t,a}$ and $p'_{t,a} = (s_{t,a} + \rho^{\mathrm{u}}_t u_{t,a} - \rho^{\mathrm{v}}_t v_{t,a})^-/\zeta_{t,a}$. These flow decisions manifestly satisfy the non-anticipativity constraints (4.9a).

If $s_{t,a} + \rho^{\mathrm{u}}_t u_{t,a} - \rho^{\mathrm{v}}_t v_{t,a} \leq 0$, then we have $g'_{t,a} = 0$ and $p'_{t,a} = -(s_{t,a} + \rho^{\mathrm{u}}_t u_{t,a} - \rho^{\mathrm{v}}_t v_{t,a})/\zeta_{t,a}$, and one can proceed as in Case 1 to show that the requirements (4.9b)–(4.9e) are met. From now on we may thus assume that $s_{t,a} + \rho^{\mathrm{u}}_t u_{t,a} - \rho^{\mathrm{v}}_t v_{t,a} > 0$, in which case $g'_{t,a} = (s_{t,a} + \rho^{\mathrm{u}}_t u_{t,a} - \rho^{\mathrm{v}}_t v_{t,a})/\eta_{t,a}$ and $p'_{t,a} = 0$. Note first that the requirements (4.9c) and (4.9d) are trivially met. We further find that

$$g'_{t,a} \;=\; (s_{t,a} + \rho^{\mathrm{u}}_t u_{t,a} - \rho^{\mathrm{v}}_t v_{t,a})/\eta_{t,a} \;\leq\; (s_{t,a} + u_{t,a})/\eta_{t,a} \;=\; g_{t,a} \leq \overline{g}_{t,a}, \tag{4.10}$$

where the second equality follows from the constraint $s_{t,a} + u_{t,a} = \eta_{t,a} g_{t,a} - \zeta_{t,a} p_{t,a}$ of problem (IT$^{\mathrm{r}}$) and the complementarity condition $g_{t,a} p_{t,a} = 0$, which implies that $p_{t,a} = 0$. Hence, requirement (4.9c) is satisfied. Finally, as $p'_{t,a} = p_{t,a} = 0$, the inequality (4.10) also implies that requirement (4.9e) is met. Thus, $g'_{t,a}$ and $p'_{t,a}$ satisfy again all of the requirements listed in (4.9).

Given the flow decisions $\{(\boldsymbol{g}'_t, \boldsymbol{p}'_t)\}_{t\in\mathscr{T}(d)}$ constructed above, we are now ready to introduce compatible spill decisions $\boldsymbol{z}'_t = \boldsymbol{z}_t + (\boldsymbol{g}_t - \boldsymbol{p}_t) - (\boldsymbol{g}'_t - \boldsymbol{p}'_t)$ for all $t \in \mathscr{T}(d)$. In the remainder of the proof we will demonstrate that the constructed solution $\big\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\big\}_{t\in\mathscr{T}(d)}$ is feasible in (4.8).

Note first that we need not be concerned with the constraints that only involve the market decisions $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t)\}_{t\in\mathscr{T}(d)}$, which are trivially satisfied because $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathscr{T}(d)}$ is feasible in problem (IT$^{\mathrm{r}}$). Moreover, all constraints of problem (4.8) that do *not* involve the spill decisions are satisfied because of (4.9). It remains to verify that the spill decisions are non-anticipative as well as non-negative and that the reservoir-balance constraints are satisfied. To this end, we note first that $\boldsymbol{z}'_t$ inherits non-anticipativity from $\boldsymbol{g}'_t$ and $\boldsymbol{p}'_t$. Similarly, $\boldsymbol{z}'_t$ inherits non-negativity from $\boldsymbol{z}_t$ thanks to (4.9e). Finally, we highlight that $(\boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)$ impacts the reservoir balance constraints only through the net water outflows $\boldsymbol{g}'_t - \boldsymbol{p}'_t + \boldsymbol{z}'_t$, which coincide with $\boldsymbol{g}_t - \boldsymbol{p}_t + \boldsymbol{z}_t$ by the construction of $\boldsymbol{z}'_t$. This guarantees via the feasibility of $\big\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\big\}_{t\in\mathscr{T}(d)}$ in (IT$^{\mathrm{r}}$) that the reservoir balance constraints are satisfied. Therefore, $\big\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\big\}_{t\in\mathscr{T}(d)}$ is indeed feasible in (4.8).

The claim now follows because $\big\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\big\}_{t\in\mathscr{T}(d)}$ and $\big\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\big\}_{t\in\mathscr{T}(d)}$ share the same market decisions, which implies that these two solutions attain the same objective values in their respective optimization problems. Thus, we can always find a feasible solution of (4.8) that attains the same objective value as any feasible solution of the reduced trader's problem (IT$^{\mathrm{r}}$). □

In summary, the results of this section show that, as long as the end-of-day reservoir levels are fixed a day in advance, the infinite-dimensional trader's problem (IT) is equivalent to the tractable linear program (IT$^{\mathrm{r}}$), whose size scales linearly with the number A of arcs in the reservoir system and the number H of hours per day. This key insight is formalized in the following theorem.

Theorem 4.10. Under Approximation 4.6, the optimal values of (IT) and (IT$^{\mathrm{r}}$) are equal.

Proof. This is an immediate consequence of Propositions 4.7, 4.8 and 4.9. □

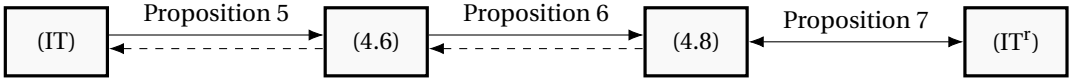


Figure 4.1 – Illustration of the relations between different versions of the trader's problem. Dashed arcs represent trivial relaxations, and solid arcs represent non-trivial implications proved in the referenced propositions.

The relations between the various optimization problems studied in this section are illustrated in Figure 4.1. Every arc encodes a relation $A \leq B$, where A and B represent the optimal values of the problems at the arc's tail and head, respectively. Dashed arcs indicate trivial relaxations, and solid arcs represent non-trivial implications proved in the referenced propositions.

We emphasize that the constant decision rules corresponding to the deterministic optimal flow decisions $\{(\boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathcal{T}(d)}$ of problem (IT$^{\mathrm{r}}$) *fail* to be optimal in (IT). In fact, *any* constant decision rules are infeasible in (IT) unless the trader is inactive on the reserve markets. To see this, assume to the contrary that $\boldsymbol{u}_t \neq \boldsymbol{0}$ or $\boldsymbol{v}_t \neq \boldsymbol{0}$ for some $t \in \mathcal{T}(d)$, in which case the left-hand side of the energy delivery constraint for hour t in problem (IT) is uncertain with different realizations for different reserve activations, while the right-hand side is deterministic. Thus, restricting the wait-and-see decisions in (IT) to constant decision rules would force the reserve market bids $\boldsymbol{u}_t$ and $\boldsymbol{v}_t$ to zero for all $t \in \mathcal{T}(d)$, thus resulting in a useless approximation of the trader's original problem (IT).

Even though naïvely interpreting the optimal solutions of (IT$^{\mathrm{r}}$) as constant decision rules fails to provide feasible, let alone optimal, policies for (IT), it is possible to efficiently convert the optimal solutions of the tractable linear program (IT$^{\mathrm{r}}$) to optimal solutions of the multi-stage stochastic program (IT). For later use, we describe this transformation in the following corollary.

Corollary 4.11. Under Approximation 4.6, if $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathcal{T}(d)}$ is optimal in (IT$^{\mathrm{r}}$) and satisfies $\boldsymbol{g}_t \circ \boldsymbol{p}_t = \boldsymbol{0}$ for all $t \in \mathcal{T}(d)$, then $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\}_{t\in\mathcal{T}(d)}$ is optimal in (IT), where

$$g'_{t,a} = \begin{cases} 0 & \text{if } g_{t,a} = 0, \\ (s_{t,a} + \rho^{\mathrm{u}}_t u_{t,a} - \rho^{\mathrm{v}}_t v_{t,a})^+ / \eta_{t,a} & \text{if } g_{t,a} > 0, \end{cases}$$

$$p'_{t,a} = \begin{cases} -(s_{t,a} + \rho^{\mathrm{u}}_t u_{t,a} - \rho^{\mathrm{v}}_t v_{t,a}) / \zeta_{t,a} & \text{if } g_{t,a} = 0, \\ (s_{t,a} + \rho^{\mathrm{u}}_t u_{t,a} - \rho^{\mathrm{v}}_t v_{t,a})^- / \zeta_{t,a} & \text{if } g_{t,a} > 0, \end{cases}$$

and $\boldsymbol{z}'_t = \boldsymbol{z}_t - (\boldsymbol{g}'_t - \boldsymbol{p}'_t) + (\boldsymbol{g}_t - \boldsymbol{p}_t)$ for all $t \in \mathcal{T}(d)$.

An optimal solution of problem (IT$^{\mathrm{r}}$) satisfying the complementarity conditions $\boldsymbol{g}_t \circ \boldsymbol{p}_t = \mathbf{0}$ is guaranteed to exist by Lemma 4.23 in the appendix. The proof of this lemma also reveals how an arbitrary feasible solution of (IT$^{\mathrm{r}}$) can be converted to another feasible solution with the same objective value that satisfies $\boldsymbol{g}_t \circ \boldsymbol{p}_t = \mathbf{0}$ by using elementary algebraic manipulations.

Proof of Corollary 4.11. From the proof of Proposition 4.9, we know that $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\}_{t\in\mathscr{T}(d)}$ is feasible in problem (4.8), which implies that it is also feasible in the relaxations (4.6) and (IT) of problem (4.8). Evidently, $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\}_{t\in\mathscr{T}(d)}$ attains the same objective value across all these problems because all of them display the same objective function, which only depends on the bidding decisions $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t)\}_{t\in\mathscr{T}(d)}$. Theorem 4.10 thus implies that $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\}_{t\in\mathscr{T}(d)}$ is not only feasible but also optimal in problem (IT). □

In the remainder of this section we leverage Theorem 4.10 to construct a restriction of the individual bidding model (I) that is susceptible to further approximations and numerical solutions. To this end, note first that Approximation 4.6 restricts the planner's original problem (IP) to

$$\begin{array}{lll} \sup & \sum_{d\in\mathscr{D}} \mathbb{E}\big[\Pi^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d, \boldsymbol{\xi}_{[\Downarrow(d)]})\big] & \\ \text{s.t.} & \boldsymbol{w}_d \in \mathscr{L}^R(\mathscr{F}_{[\Downarrow(d)]}) & \forall d\in\mathscr{D} \\ & \underline{\boldsymbol{w}}_d \le \boldsymbol{w}_d \le \overline{\boldsymbol{w}}_d & \forall d\in\mathscr{D},\ \mathbb{P}\text{-a.s.} \\ & \boldsymbol{w}_0 \le \boldsymbol{w}_D & \mathbb{P}\text{-a.s.}, \end{array} \qquad (\mathrm{IP}^{\mathrm{r}})$$

where the end-of-day reservoir levels $\boldsymbol{w}_d$ are chosen one day in advance and thus adapt to information available at time $\Downarrow(d)$. Under this information restriction, Theorem 4.10 allows us to compute the optimal value $\Pi^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d, \boldsymbol{\xi}_{[\Downarrow(d)]})$ of the trader's problem (IT) by solving the linear program (IT$^{\mathrm{r}}$). This insight further enables us to prove that the planner's reduced problem (IP$^{\mathrm{r}}$) is equivalent to

$$\begin{array}{lll} \sup & \sum_{t\in\mathscr{T}} \mathbb{E}\left[\tilde{\pi}^{\mathrm{s}}_t \mathbf{1}^\top \boldsymbol{s}_t + \tilde{\pi}^{\mathrm{u}}_t \mathbf{1}^\top \boldsymbol{u}_t + \tilde{\pi}^{\mathrm{v}}_t \mathbf{1}^\top \boldsymbol{v}_t\right] & \\ \text{s.t.} & \boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t \in \mathscr{L}^A(\mathscr{F}_{[T]}),\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[T]}),\ \boldsymbol{w}_d \in \mathscr{L}^R(\mathscr{F}_{[\Downarrow(d)]}) & \forall d\in\mathscr{D},\ \forall t\in\mathscr{T}(d) \\ & \mathbf{0}\le \boldsymbol{u}_t,\ \mathbf{0}\le \boldsymbol{v}_t,\ \mathbf{0}\le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0}\le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0}\le \boldsymbol{z}_t & \forall d\in\mathscr{D},\ \forall t\in\mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\ & \boldsymbol{s}_t + \boldsymbol{u}_t = \boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t,\ \boldsymbol{s}_t - \boldsymbol{v}_t \ge -\boldsymbol{\zeta}_t \circ \overline{\boldsymbol{p}}_t & \forall d\in\mathscr{D},\ \forall t\in\mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\ & \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \le \overline{\boldsymbol{w}}_t & \forall d\in\mathscr{D},\ \forall t\in\mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\ & \boldsymbol{w}_d \le \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) & \forall d\in\mathscr{D},\ \mathbb{P}\text{-a.s.} \\ & \underline{\boldsymbol{w}}_d \le \boldsymbol{w}_d \le \overline{\boldsymbol{w}}_d,\ \boldsymbol{w}_0 \le \boldsymbol{w}_D & \forall d\in\mathscr{D},\ \mathbb{P}\text{-a.s.}, \end{array} \qquad (\mathrm{I}^{\mathrm{r}})$$

where

$$\tilde{\pi}^{\mathrm{s}}_t = \mathbb{E}\big[\pi^{\mathrm{s}}_t \,\big|\, \boldsymbol{\xi}_{[\Downarrow(t)]}\big], \quad \tilde{\pi}^{\mathrm{u}}_t = \mathbb{E}\big[\pi^{\mathrm{u}}_t + \rho^{\mathrm{u}}_t \psi^{\mathrm{u}}_t \,\big|\, \boldsymbol{\xi}_{[\Downarrow(t)]}\big] \quad \text{and} \quad \tilde{\pi}^{\mathrm{v}}_t = \mathbb{E}\big[\pi^{\mathrm{v}}_t + \rho^{\mathrm{v}}_t \psi^{\mathrm{v}}_t \,\big|\, \boldsymbol{\xi}_{[\Downarrow(t)]}\big] \quad \forall t\in\mathscr{T}.$$

Problem (I$^{\mathrm{r}}$) can be viewed as a reduction of the original individual bidding model (I).

Theorem 4.12. The optimal values of the problems (IP$^{\mathrm{r}}$) and (I$^{\mathrm{r}}$) are equal, and they are smaller than or equal to the optimal value of problem (I).

Proof. From Propositions 4.2 and 4.3 we know that the individual bidding model (I) has the same optimal value as the planner's problem (IP), and it is easy to see that problem (IP$^{\mathrm{r}}$) is obtained by applying Approximation 4.6 to problem (IP). As this approximation consists in restricting the planner's information structure, the optimal value of problem (IP$^{\mathrm{r}}$) is no larger than that of problem (I). It remains to be shown that (IP$^{\mathrm{r}}$) and (I$^{\mathrm{r}}$) share the same optimal value. As Approximation 4.6 is in force, we may conclude via Theorem 4.10 that the optimal value $\Pi^{\mathrm{I}}_d(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d, \boldsymbol{\xi}_{[\Downarrow(d)]})$ of the stochastic program (IT) coincides with the optimal value of the linear program (IT$^{\mathrm{r}}$). Substituting (IT$^{\mathrm{r}}$) into (IP$^{\mathrm{r}}$) and using Theorem 14.60 by Rockafellar and Wets (2010) to move the maximization over the trading and flow decisions out of the expectation and the sum finally yields (I$^{\mathrm{r}}$). □

The individual bidding model (I) and its reduction (I$^{\mathrm{r}}$) differ in the following aspects.

- In problem (I) the market decisions $(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t)$ are adapted to information that is available at the *beginning of day* $d(t)$, while the operational decisions $(\boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)$ are adapted to *real-time information*. In contrast, in (I$^{\mathrm{r}}$) these decisions are taken under *perfect information*.
- Problem (I) has *random recourse* because the reserve market bids in the energy delivery constraints are multiplied by the uncertain reserve activations. In contrast, problem (I$^{\mathrm{r}}$) has *fixed recourse* because all reserve activations were eliminated from the constraints.
- Problem (I$^{\mathrm{r}}$) accommodates the *valid cuts* derived in Proposition 4.8, which are absent in (I).

We highlight again that problem (I$^{\mathrm{r}}$) was obtained from the original individual bidding model (I) by applying a single information restriction, whereby the end-of-day reservoir levels must be chosen one day in advance. No other approximations have been applied.

4.5 Reduction of the Collective Trading Model

Armed with a profound understanding of the *individual* bidding model (I), we are now ready to analyze the more flexible *collective* bidding model (C). Using a similar reasoning as in Section 4.4, we will show that the trader's problem (CT) can again be reformulated as a tractable linear program if the reservoir targets imposed by the planner are chosen a day in advance. We thus subject the planner's problem (4.5) to the information restriction of Approximation 4.6, which reduces the achievable expected revenue but makes the feasible set of the trader's problem (CT) independent of all exogenous uncertainties except for the reserve activations $\{\rho^{\mathrm{u}}_t, \rho^{\mathrm{v}}_t\}_{t\in\mathcal{T}(d)}$. As the objective function of (CT) depends only on the bidding decisions $\{(s_t, u_t, v_t)\}_{t\in\mathcal{T}(d)}$, we may restrict the operational decisions $\{(\boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathcal{T}(d)}$ to depend only on the reserve activations without sacrificing optimality. Doing so results in the

following variant of the trader's problem,

$$\begin{array}{lll}
\sup & \mathbb{E}\big[\sum_{t\in\mathscr{T}(d)} \pi^{\mathrm{s}}_t s_t + (\pi^{\mathrm{u}}_t + \rho^{\mathrm{u}}_t \psi^{\mathrm{u}}_t) u_t + (\pi^{\mathrm{v}}_t + \rho^{\mathrm{v}}_t \psi^{\mathrm{v}}_t) v_t \,\big|\, \boldsymbol{\xi}_{[\Downarrow(d)]}\big] & \\
\text{s.t.} & s_t, u_t, v_t \in \mathbb{R},\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}^{\rho}_{[\downarrow(d),t]}) & \forall t \in \mathscr{T}(d) \\
& 0 \le u_t,\ 0 \le v_t,\ \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t & \forall t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\
& s_t + \rho^{\mathrm{u}}_t u_t - \rho^{\mathrm{v}}_t v_t = \boldsymbol{\eta}_t^\top \boldsymbol{g}_t - \boldsymbol{\zeta}_t^\top \boldsymbol{p}_t & \forall t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \le \overline{\boldsymbol{w}}_t & \forall t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\
& \boldsymbol{w}_d \le \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) & \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.},
\end{array} \tag{4.11}$$

where $\mathscr{F}^{\rho}_{[\downarrow(d),t]}$ is again defined as $\sigma(\{(\rho^{\mathrm{u}}_\tau, \rho^{\mathrm{v}}_\tau)\}_{\tau=\downarrow(d)}^{t})$. Even though (4.11) constitutes a restriction of the trader's problem (CT), their optimal values can be shown to coincide.

Proposition 4.13. Under Approximation 4.6, the optimal values of (CT) and (4.11) are equal.

Proof. As in the proof of Proposition 4.7, one may condition any feasible solution of problem (CT) on the history of reserve activations to construct a feasible solution of problem (4.11) that adopts the same objective value. Details are omitted for brevity. □

Proposition 4.13 reduces the infinite-dimensional stochastic program (CT) to the much simpler stochastic program (4.11). By counting the atoms of the underlying σ-algebras $\mathscr{F}^{\rho}_{[\downarrow(d),t]}$ for all $t \in \mathscr{T}(d)$, one readily verifies that problem (4.11) is equivalent to a prohibitively large linear program with $\mathscr{O}(A \cdot 3^H)$ decision variables and constraints. In analogy to Section 4.4, however, we can apply two additional simplifications, which will reveal that problem (4.11) is indeed also equivalent to an efficiently solvable linear program with $\mathscr{O}(A \cdot H)$ decision variables and constraints. As a first step towards this goal, we derive a family of valid inequalities that may be appended to problem (4.11).

Proposition 4.14. Any feasible solution of (4.11) satisfies $s_t - v_t \ge -\boldsymbol{\zeta}^\top \overline{\boldsymbol{p}}_t$ for all $t \in \mathscr{T}(d)$.

Proof. The proof widely parallels to that of Proposition 4.8 and is thus omitted for brevity. □

The valid inequalities of Proposition 4.14 characterize the maximum bids on the reserve-down market (collectively across all arcs of the reservoir system) that can be honored under all possible realizations of the reserve activations. Appending these valid inequalities to

problem (4.11) yields

$$\begin{array}{lll}
\max & \mathbb{E}\big[\sum_{t\in\mathscr{T}(d)} \pi^{\mathrm{s}}_t s_t + (\pi^{\mathrm{u}}_t + \rho^{\mathrm{u}}_t \psi^{\mathrm{u}}_t) u_t + (\pi^{\mathrm{v}}_t + \rho^{\mathrm{v}}_t \psi^{\mathrm{v}}_t) v_t \,\big|\, \boldsymbol{\xi}_{[\Downarrow(d)]}\big] & \\
\text{s.t.} & s_t, u_t, v_t \in \mathbb{R},\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}^{\rho}_{[\downarrow(d),t]}) & \forall t \in \mathscr{T}(d) \\
& 0 \le u_t,\ 0 \le v_t,\ \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t & \forall t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\
& s_t + \rho^{\mathrm{u}}_t u_t - \rho^{\mathrm{v}}_t v_t = \boldsymbol{\eta}_t^\top \boldsymbol{g}_t - \boldsymbol{\zeta}_t^\top \boldsymbol{p}_t & \forall t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\
& s_t - v_t \ge -\boldsymbol{\zeta}^\top \overline{\boldsymbol{p}}_t & \forall t \in \mathscr{T}(d) \\
& \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \le \overline{\boldsymbol{w}}_t & \forall t \in \mathscr{T}(d),\ \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.} \\
& \boldsymbol{w}_d \le \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) & \mathbb{P}_{|\boldsymbol{\xi}_{[\Downarrow(d)]}}\text{-a.s.}
\end{array} \tag{4.12}$$

which is equivalent to problem (4.11) by Proposition 4.14.

Next, we show that problem (4.12) is equivalent to the reduced stochastic program

$$\begin{array}{lll}
\max & \mathbb{E}\big[\sum_{t\in\mathscr{T}(d)} \pi^{\mathrm{s}}_t s_t + (\pi^{\mathrm{u}}_t + \rho^{\mathrm{u}}_t \psi^{\mathrm{u}}_t) u_t + (\pi^{\mathrm{v}}_t + \rho^{\mathrm{v}}_t \psi^{\mathrm{v}}_t) v_t \,\big|\, \boldsymbol{\xi}_{[\Downarrow(d)]}\big] & \\
\text{s.t.} & s_t, u_t, v_t \in \mathbb{R},\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathbb{R}^A & \forall t \in \mathscr{T}(d) \\
& 0 \le u_t,\ 0 \le v_t,\ \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t & \forall t \in \mathscr{T}(d) \\
& s_t + u_t = \boldsymbol{\eta}_t^\top \boldsymbol{g}_t - \boldsymbol{\zeta}_t^\top \boldsymbol{p}_t & \forall t \in \mathscr{T}(d) \\
& s_t - v_t \ge -\boldsymbol{\zeta}_t^\top \overline{\boldsymbol{p}}_t & \forall t \in \mathscr{T}(d) \\
& \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \le \overline{\boldsymbol{w}}_t & \forall t \in \mathscr{T}(d) \\
& \boldsymbol{w}_d \le \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau), &
\end{array} \tag{CT$^{\mathrm{r}}$}$$

Like problem (IT$^{\mathrm{r}}$), problem (CT$^{\mathrm{r}}$) involves only here-and-now decisions, and its constraints are deterministic conditional on $\boldsymbol{\xi}_{[\Downarrow(d)]}$. In fact, it constitutes a linear program of size $\mathscr{O}(A \cdot H)$.

To demonstrate that problems (CT$^{\mathrm{r}}$) and (4.12) share the same optimal value, we develop a fundamentally new proof strategy that establishes a link to the *individual* bidding models studied in Section 4.4. Specifically, using techniques familiar from Proposition 4.9, we first show that the optimal value of (CT$^{\mathrm{r}}$) is larger or equal to that of (4.12) (see Proposition 4.15 below), and then we develop a new duality argument to prove that the optimal value of (IT$^{\mathrm{r}}$) is larger or equal to that of (CT$^{\mathrm{r}}$) (see Proposition 4.16 below). This new approach is needed because the aggregation of arc-wise market bids into a single collective bid in the energy delivery constraints and in the valid inequalities of Proposition 4.14 make it difficult to construct a feasible solution for problem (4.12) from a feasible solution of problem (CT$^{\mathrm{r}}$) by using similar arguments as in the proof of Proposition 4.9.

Anticipating the results of Propositions 4.15 and 4.16, the relations between the various variants of the trader's problem studied so far are illustrated in Figure 4.2. Every arc encodes a relation $A \le B$, where A and B represent the optimal values of the problems at the arc's tail and head, respectively. Dashed arcs indicate trivial relaxations, and solid arcs represent non-trivial implications proved in the referenced propositions. Interpreting Figure 4.2 as a

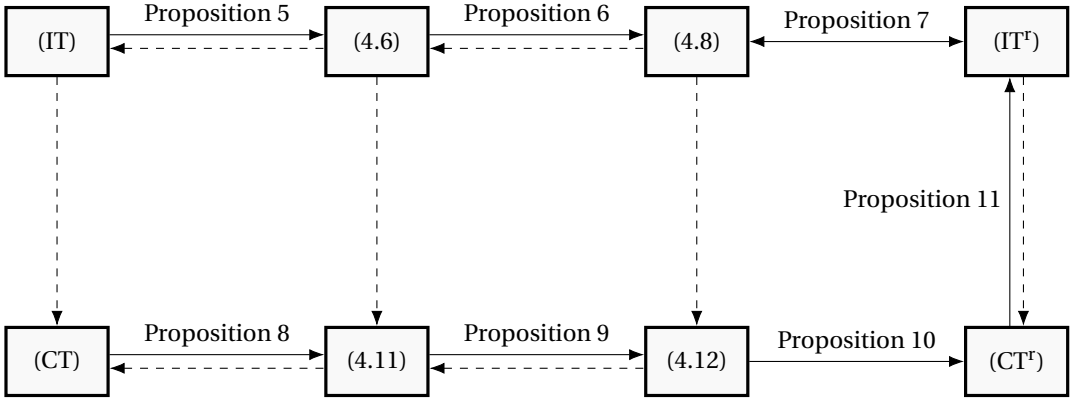


Figure 4.2 – Illustration of the relations between different variants of the trader's problem. Dashed arcs represent trivial relaxations, and solid arcs represent non-trivial implications proved in the referenced propositions.

directed (multi)graph, we note that Propositions 4.15 and 4.16 complete a counter-clockwise loop that visits each node, which implies that the optimal values of *all* optimization problems are in fact equal. This implies in particular that the more flexible collective bidding model does not generate higher revenues than the individual bidding model. We will discuss this insight in more detail after proving Propositions 4.15 and 4.16.

Proposition 4.15. Under Approximation 4.6, the optimal value of problem (CT^{r}) is larger or equal to that of problem (4.12).

Proof. The claim follows from a simple adaptation of the corresponding argument in the proof of Proposition 4.9 and can thus be omitted for brevity. □

Proposition 4.16. Under Approximation 4.6, the optimal value of problem (IT^{r}) is larger or equal to that of problem (CT^{r}).

Proof. The claim follows if we can show that for every feasible solution of problem (CT^{r}) there exists a feasible solution of problem (IT^{r}) that attains the same objective value. To this end, we select an arbitrary feasible solution $\{(s_t, u_t, v_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathcal{T}(d)}$ of problem (CT^{r}) and aim to show that problem (IT^{r}) admits a feasible solution $\{(\boldsymbol{s}'_t, \boldsymbol{u}'_t, \boldsymbol{v}'_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathcal{T}(d)}$ with the same flow decisions that satisfies $\mathbf{1}^\top \boldsymbol{s}'_t = s_t$, $\mathbf{1}^\top \boldsymbol{u}'_t = u_t$ and $\mathbf{1}^\top \boldsymbol{v}'_t = v_t$ for all $t \in \mathcal{T}(d)$. These identities ensure that the two solutions adopt the same objective values in their respective optimization problems.

As the flow decisions $\{(\boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathcal{T}(d)}$ are preserved, their upper and lower bounds as well as the reservoir level constraints are trivially satisfied. It thus suffices to show that there exist individual market bids $\{(\boldsymbol{s}'_t, \boldsymbol{u}'_t, \boldsymbol{v}'_t)\}_{t\in\mathcal{T}(d)}$ that are consistent with the prescribed collective

market bids $\{(s_t, u_t, v_t\}_{t\in\mathscr{T}(d)}$ and that satisfy all remaining constraints of problem (IT$^{\mathrm{r}}$), that is, the non-negativity constraints, the energy delivery constraints, and the valid inequalities from Proposition 4.14. Formally, such individual market bids exist if and only if the optimal value of the feasibility problem

$$\begin{array}{ll} \min & 0 \\ \text{s.t.} & \boldsymbol{s}'_t \in \mathbb{R}^A,\ \boldsymbol{u}'_t \in \mathbb{R}^A_+,\ \boldsymbol{v}'_t \in \mathbb{R}^A_+ \\ & \mathbf{1}^\top \boldsymbol{s}'_t = s_t,\ \mathbf{1}^\top \boldsymbol{u}'_t = u_t,\ \mathbf{1}^\top \boldsymbol{v}'_t = v_t \\ & \boldsymbol{s}'_t + \boldsymbol{u}'_t = \boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t \\ & \boldsymbol{s}'_t - \boldsymbol{v}'_t \geq -\boldsymbol{\zeta}_t \circ \overline{\boldsymbol{p}}_t \end{array} \tag{4.13}$$

vanishes for each hour $t \in \mathscr{T}(d)$, provided that $\{(s_t, u_t, v_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathscr{T}(d)}$ is feasible in (CT$^{\mathrm{r}}$). Assigning dual variables $\alpha, \beta, \gamma \in \mathbb{R}$ to the consistency constraints for the market bids, $\boldsymbol{\lambda} \in \mathbb{R}^A$ to the energy delivery constraints and $\boldsymbol{\mu} \in \mathbb{R}^A_+$ to the valid inequalities from Proposition 4.14, the linear program dual to the above feasibility problem can be represented as

$$\begin{array}{ll} \max & \alpha s_t + \beta u_t + \gamma v_t + \boldsymbol{\lambda}^\top(\boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t) - \boldsymbol{\mu}^\top(\boldsymbol{\zeta}_t \circ \overline{\boldsymbol{p}}_t) \\ \text{s.t.} & \alpha, \beta, \gamma \in \mathbb{R},\ \boldsymbol{\lambda} \in \mathbb{R}^A,\ \boldsymbol{\mu} \in \mathbb{R}^A_+ \\ & \alpha\mathbf{1} + \boldsymbol{\lambda} + \boldsymbol{\mu} = \mathbf{0} \\ & \beta\mathbf{1} + \boldsymbol{\lambda} \leq \mathbf{0} \\ & \gamma\mathbf{1} - \boldsymbol{\mu} \leq \mathbf{0}. \end{array} \tag{4.14}$$

Strong duality holds because the feasible set of the dual problem contains the origin. In the remainder we will argue that the objective value of any feasible solution $(\alpha, \beta, \gamma, \boldsymbol{\lambda}, \boldsymbol{\mu})$ of the dual linear program (4.14) is bounded above by 0. To this end, we define $\underline{\mu} = \min_{a\in\mathscr{A}} \mu_a \geq 0$ and note that combining the first two constraints in (4.14) yields $(\beta - \alpha)\mathbf{1} - \boldsymbol{\mu} \leq \mathbf{0}$, which in turn implies that $\beta - \alpha - \underline{\mu} \leq 0$. We thus find

$$\begin{aligned} & \alpha s_t + \beta u_t + \gamma v_t + \boldsymbol{\lambda}^\top(\boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t) - \boldsymbol{\mu}^\top(\boldsymbol{\zeta}_t \circ \overline{\boldsymbol{p}}_t) \\ = \;& \alpha s_t + \beta u_t + \gamma v_t + (\boldsymbol{\lambda} + \boldsymbol{\mu})^\top(\boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t) - \boldsymbol{\mu}^\top(\boldsymbol{\eta}_t \circ \boldsymbol{g}_t + \boldsymbol{\zeta}_t \circ (\overline{\boldsymbol{p}}_t - \boldsymbol{p}_t)) \\ \leq \;& \alpha s_t + \beta u_t + \gamma v_t - \alpha(\boldsymbol{\eta}_t^\top \boldsymbol{g}_t - \boldsymbol{\zeta}_t^\top \boldsymbol{p}_t) - \underline{\mu}(\boldsymbol{\eta}_t^\top \boldsymbol{g}_t + \boldsymbol{\zeta}_t^\top (\overline{\boldsymbol{p}}_t - \boldsymbol{p}_t)) \\ = \;& \alpha s_t + \beta u_t + \gamma v_t - \alpha(s_t + u_t) - \underline{\mu}(s_t + u_t + \boldsymbol{\zeta}_t^\top \overline{\boldsymbol{p}}_t) \\ = \;& (\beta - \alpha - \underline{\mu}) u_t + \gamma v_t - \underline{\mu}(s_t + \boldsymbol{\zeta}_t^\top \overline{\boldsymbol{p}}_t) \\ \leq \;& (\beta - \alpha - \underline{\mu}) u_t + \underline{\mu}(v_t - s_t - \boldsymbol{\zeta}_t^\top \overline{\boldsymbol{p}}_t) \;\leq\; 0, \end{aligned}$$

where the first inequality exploits the relations $\boldsymbol{\lambda} + \boldsymbol{\mu} = -\alpha\mathbf{1}$ (by the feasibility of α, $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ in (4.14)) and $\underline{\mu}\mathbf{1} \leq \boldsymbol{\mu}$ (by the construction of $\underline{\mu}$) together with the non-negativity of $\boldsymbol{g}_t$ and $\overline{\boldsymbol{p}}_t - \boldsymbol{p}_t$ (by the feasibility of $\boldsymbol{g}_t$ and $\boldsymbol{p}_t$ in (CT$^{\mathrm{r}}$)). The second equality follows from the energy delivery constraints in (CT$^{\mathrm{r}}$), and the second inequality holds because $\gamma \leq \underline{\mu}$ (by the feasibility of γ in (4.14) and the construction of $\underline{\mu}$) and $v_t \geq 0$ (by the feasibility of v_t in (CT$^{\mathrm{r}}$)). The last

inequality, finally, follows from our earlier observations that $\beta - \alpha - \underline{\mu} \leq 0$ and $\underline{\mu} \geq 0$ combined with the relations $u_t \geq 0$ and $v_t - s_t - \boldsymbol{\zeta}_t^\top \overline{\boldsymbol{p}}_t \leq 0$ (by the feasibility of $\overline{s}_t$, u_t and $\overline{v}_t$ in (CT$^\mathrm{r}$)).

In conclusion, we have demonstrated that the optimal value of the dual feasibility problem (4.14)—and thus also that of its primal counterpart—must vanish. As our arguments hold for any feasible solution $\{(s_t, u_t, v_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathscr{T}(d)}$ of (CT$^\mathrm{r}$) and for any $t \in \mathscr{T}(d)$, the claim follows. □

The results of Sections 4.4 and 4.5 as visualized in Figure 4.2 culminate in the following main theorem.

Theorem 4.17. Under Approximation 4.6, the optimal values of the trader's collective bidding models (CT) and (CT$^\mathrm{r}$) and the trader's individual bidding models (IT) and (IT$^\mathrm{r}$) are all equal.

Proof. By construction, the problems (CT) and (CT$^\mathrm{r}$) are relaxations of the problems (IT) and (IT$^\mathrm{r}$), respectively. This is because the energy delivery constraints in the collective bidding models can be obtained by aggregating those in the individual bidding models over all arcs $a \in \mathscr{A}$. In view of this, Proposition 4.16 implies that the optimal values of (IT$^\mathrm{r}$) and (CT$^\mathrm{r}$) are in fact equal. From Theorem 4.10 we further know that problem (IT) also shares the same optimal value. Finally, as problem (CT) is a relaxation of (IT) and as Propositions 4.13, 4.14 and 4.15 imply that the optimal value of (CT) is smaller or equal to that of (CT$^\mathrm{r}$), we conclude that (IT) displays also the same optimal value as (CT) and (CT$^\mathrm{r}$) (and (IT$^\mathrm{r}$)). This observation completes the proof. □

In the remainder we use Theorem 4.17 to simplify the collective bidding model (C). To this end, note first that Approximation 4.6 restricts the planner's problem (CP) to

$$\begin{array}{lll} \sup & \sum_{d\in\mathscr{D}} \mathbb{E}\big[\Pi_d^{\mathrm{C}}(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d, \boldsymbol{\xi}_{[\Downarrow(d)]})\big] & \\ \text{s.t.} & \boldsymbol{w}_d \in \mathscr{L}^R(\mathscr{F}_{[\Downarrow(d)]}) & \forall\, d \in \mathscr{D} \\ & \underline{\boldsymbol{w}}_d \leq \boldsymbol{w}_d \leq \overline{\boldsymbol{w}}_d & \forall\, d \in \mathscr{D},\ \mathbb{P}\text{-a.s.} \\ & \boldsymbol{w}_0 \leq \boldsymbol{w}_D & \mathbb{P}\text{-a.s.} \end{array} \tag{CP$^\mathrm{r}$}$$

Theorem 4.17 thus allows us to compute the optimal value $\Pi_d^{\mathrm{C}}(\boldsymbol{w}_{d-1}, \boldsymbol{w}_d, \boldsymbol{\xi}_{[\Downarrow(d)]})$ of the trader's problem (CT) by solving the linear program (CT$^\mathrm{r}$). One can also show that (CP$^\mathrm{r}$) is equivalent

to

$$
\begin{array}{lll}
\sup & \sum_{t\in\mathscr{T}} \mathbb{E}\left[\tilde{\pi}^{\mathrm{s}}_t s_t + \tilde{\pi}^{\mathrm{u}}_t \mathbf{1}^\top \boldsymbol{u}_t + \tilde{\pi}^{\mathrm{v}}_t \mathbf{1}^\top \boldsymbol{v}_t\right] & \\
\text{s.t.} & s_t, u_t, v_t \in \mathscr{L}(\mathscr{F}_{[T]}),\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[T]}),\ \boldsymbol{w}_d \in \mathscr{L}^R(\mathscr{F}_{[\Downarrow(d)]}) & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d) \\
& 0 \le u_t,\ 0 \le v_t,\ \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& s_t + u_t = \boldsymbol{\eta}_t^\top \boldsymbol{g}_t - \boldsymbol{\zeta}_t^\top \boldsymbol{p}_t,\ s_t - v_t \ge -\boldsymbol{\zeta}_t^\top \overline{\boldsymbol{p}}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_{d-1} + \sum_{\tau=\downarrow(d)}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \le \overline{\boldsymbol{w}}_t & \forall d \in \mathscr{D},\ \forall t \in \mathscr{T}(d),\ \mathbb{P}\text{-a.s.} \\
& \boldsymbol{w}_d \le \boldsymbol{w}_{d-1} + \sum_{\tau\in\mathscr{T}(d)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) & \forall d \in \mathscr{D},\ \mathbb{P}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_d \le \boldsymbol{w}_d \le \overline{\boldsymbol{w}}_d,\ \boldsymbol{w}_0 \le \boldsymbol{w}_D & \forall d \in \mathscr{D},\ \mathbb{P}\text{-a.s.},
\end{array}
\tag{C$^{\mathrm{r}}$}
$$

where $\tilde{\pi}^{\mathrm{s}}_t$, $\tilde{\pi}^{\mathrm{u}}_t$ and $\tilde{\pi}^{\mathrm{v}}_t$ are defined as in Section 4.4. Problem (C$^{\mathrm{r}}$) can be viewed as a reduction of (C).

Theorem 4.18. The optimal values of the problems (CP$^{\mathrm{r}}$) and (C$^{\mathrm{r}}$) are equal, and they are smaller than or equal to the optimal value of problem (C).

Proof. The proof is similar to that of Theorem 4.12 and thus omitted. □

The collective bidding model (C) differs from (C$^{\mathrm{r}}$) in the same way as the individual bidding model (I) differs from (I$^{\mathrm{r}}$). Specifically, in contrast to (C), the reduced collective bidding model (C$^{\mathrm{r}}$) has fixed recourse, and all market bids and flow decisions in (C$^{\mathrm{r}}$) are taken under perfect information. Unlike problem (C), its reduction (C$^{\mathrm{r}}$) further accommodates the valid inequalities derived in Proposition 4.14. Note that (C$^{\mathrm{r}}$) was simply obtained from (C) by applying Approximation 4.6.

4.6 Numerical Solution of the Reduced Collective Bidding Model

Approximation 4.6 allowed us to reduce the individual bidding model (I) to (I$^{\mathrm{r}}$) and the collective bidding model (C) to (C$^{\mathrm{r}}$). The reduced models (I$^{\mathrm{r}}$) and (C$^{\mathrm{r}}$) still constitute infinite-dimensional linear programs over spaces of measurable functions, and thus there is little hope to solve them *exactly*. Key advantages of the reduced models (I$^{\mathrm{r}}$) and (C$^{\mathrm{r}}$) over the respective original bidding models are that they have *fixed recourse* and that all market bids and operational decisions are taken under *perfect information*, while only the reservoir targets obey complicating non-anticipativity constraints. Below we will show that (I$^{\mathrm{r}}$) and (C$^{\mathrm{r}}$) can be addressed with standard techniques.

From Theorem 4.17 we know that, under Approximation 4.6, the individual and collective bidding models generate the same expected revenue, which implies that the hydropower producer should have no strict preference for either of these models. In the remainder of the paper, we will thus focus on the collective bidding model, which has the advantage of

reflecting standard market practice. To make problem (C^r) amenable to numerical solution, we further reduce it to a two-stage stochastic program by restricting the reservoir filling levels to parsimonious *affine decision rules* that depend on the observable random parameters only through a few judiciously chosen features.

Approximation 4.19 (Affine Decision Rule Restriction)**.** For all $d \in \mathscr{D}$, the end-of-day reservoir filling levels are representable as

$$\boldsymbol{w}_d = \boldsymbol{\lambda}_d + \boldsymbol{\Phi}_d \left(\sum_{\tau \in \mathscr{T}(d)} \boldsymbol{\phi}_\tau \right) + \Psi_d \left(\sum_{\tau=1}^{\uparrow(d-1)} \boldsymbol{\phi}_\tau \right) + \boldsymbol{\mu}_d \left(\tfrac{1}{H} \sum_{\tau \in \mathscr{T}(d)} \pi_\tau^{\mathrm{s}} \right)$$

for some fixed vectors $\boldsymbol{\lambda}_d, \boldsymbol{\mu}_d \in \mathbb{R}^R$ and matrices $\boldsymbol{\Phi}_d, \boldsymbol{\Psi}_d \in \mathbb{R}^{R \times R}$.

Approximation 4.19 restricts the reservoir filling levels at the end of day d to affine functions of the following features: (i) the cumulative natural inflows into the reservoirs across day d, (ii) the cumulative natural inflows into the reservoirs across the planning horizon until the end of day $d-1$ and (iii) the average spot price on day d. The proposed affine decision rules are parsimonious as they compress the history of all observations into a few relevant features, but they are flexible enough to allow the planner to set different reservoir targets for wet and dry days, for wet and dry seasons as well as for high- and low-price days. Thanks to our standing assumptions, all of the proposed features are observable at the beginning of day d, and therefore the non-anticipativity conditions $\boldsymbol{w}_d \in \mathscr{L}^R(\mathscr{F}_{[\Downarrow(d)]})$, $d \in \mathscr{D}$, imposed by Approximation 4.6 are automatically satisfied.

Note that Approximation 4.19 restricts the functional form of the reservoir targets, and thus it results in a conservative lower bound on the optimal value of problem (C^r), which itself underestimates the optimal value of the original collective bidding model (C) by virtue of Theorem 4.18. We further emphasize that Approximation 4.19 reduces problem (C^r) to a a two-stage stochastic program with here-and-now decisions $\boldsymbol{\lambda}_d, \boldsymbol{\mu}_d \in \mathbb{R}^R$ and $\boldsymbol{\Phi}_d, \boldsymbol{\Psi}_d \in \mathbb{R}^{R \times R}$, $d \in \mathscr{D}$, which are chosen without any information about $\boldsymbol{\xi}_{[T]}$, and with wait-and-see decisions $s_t, u_t, v_t \in \mathscr{L}(\mathscr{F}_{[T]})$ and $\boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[T]})$, $t \in \mathscr{T}$, which are chosen under perfect information about $\boldsymbol{\xi}_{[T]}$. The emerging two-stage stochastic program can then be solved with the popular sample average approximation (Shapiro et al. 2009).

Approximation 4.20 (Sample Average Approximation)**.** The original probability measure $\mathbb{P}$ is replaced with a discrete empirical measure $\widehat{\mathbb{P}} = \frac{1}{N} \sum_{n=1}^{N} \delta_{\omega^{(n)}}$, where $\delta_{\omega^{(n)}}$ stands for the Dirac point mass at $\omega^{(n)}$, and where $\omega^{(n)} \in \Omega$, $n = 1, \ldots, N$, constitute independent samples from $\mathbb{P}$.

We emphasize that, even though we use an empirical probability measure $\widehat{\mathbb{P}}$ when solving the decision rule approximation of problem (C^r), the conditional expectations $\tilde{\pi}_t^{\mathrm{s}}$, $\tilde{\pi}_t^{\mathrm{u}}$ and $\tilde{\pi}_t^{\mathrm{v}}$ in the objective function of (C^r) are pre-computed under the original measure $\mathbb{P}$.

Affine decision rule approximations are now routinely used in linear adjustable robust optimization (Ben-Tal et al. 2004b) and linear multistage stochastic programming (Chen et al. 2008). If the uncertain problem parameters are supported on a conic representable set, then

the best affine decision rules can be computed *exactly* by solving a tractable convex program. For general uncertainty sets, the best affine decision rules can still be computed *approximately* via sampling techniques. This idea has been investigated by Vayanos et al. (2012) and Bodur and Luedtke (2018) in the context of robust optimization and stochastic programming, respectively. If *all* wait-and-see decisions are restricted to affine decision rules, then the original multistage problem reduces to a one-stage problem. Better approximations can be obtained by restricting only a subset of decision variables (the *state variables*) to affine decision rules while retaining the flexibility of the remaining ones (the *recourse variables*). If the state variables are chosen so that the optimization problem decouples temporally, then Bodur and Luedtke (2018) argue that the original multistage problem reduces to a two-stage problem where the decision rule coefficients of the state variables and the values of the recourse variables are determined in the first and the second stage, respectively. In our context, the end-of-day water levels can be chosen as the state variables and the bidding and flow decisions as the recourse variables, respectively. Under the Approximations 4.6 and 4.19, Theorem 4.18 again ensures that the resulting two-stage approximation enjoys a fixed recourse.

While the collective bidding model (C) determines a policy for the entire planning horizon $\mathscr{D}$, in practice one solves the problem repeatedly in a rolling horizon fashion and each time only implements the decisions of the first day for the following reasons: (i) The true state of the world $\boldsymbol{\omega}$ may differ from all discretization points $\boldsymbol{\omega}^{(n)}$, $n = 1,\ldots,N$, of the empirical measure $\widehat{\mathbb{P}}$, and thus the sample average approximation does not provide any recourse decisions corresponding to $\boldsymbol{\omega}$. (ii) Problem (C) only models a finite time window of the perpetual operation of the reservoir system, which adversely affects the decisions towards the end of the planning horizon $\mathscr{D}$. (iii) Resolving problem (C) on a daily basis ensures that the end-of-day reservoir targets adapt to all available information, and it reduces the conservatism of the affine decision rule approximation.

In the following we describe two complementary approaches for approximately solving the collective bidding model (C) and for constructing near-optimal market and operational decisions for the first day of the planning horizon. Both approaches first solve the linear program obtained by applying Approximations 4.6, 4.19 and 4.20 to problem (C) and record the optimal reservoir targets $\boldsymbol{w}_1$ for the end of day 1. Note that $\boldsymbol{w}_1$ is deterministic by virtue of Approximation 4.6. Next, both approaches solve the linear program (CT$^{\mathrm{r}}$) for $d = 1$ with inputs $\boldsymbol{w}_0$ and $\boldsymbol{w}_1$, which is equivalent to the trader's problem (CT) on day 1, and record the optimal collective market bids $\{(s_t, u_t, v_t)\}_{t\in\mathscr{T}(1)}$.

From now on the two approaches proceed differently. The first—naïve—approach solves the linear program (4.13) to compute for each hour $t \in \mathscr{T}(1)$ a set of individual market bids $(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t)$ that are compatible with the given collective market bids (s_t, u_t, v_t). Recall that problem (4.13) is guaranteed to be feasible by Proposition 4.16. The resulting solution $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathscr{T}(1)}$ is optimal for problem (IT$^{\mathrm{r}}$) and can thus be converted to an optimal solution $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\}_{t\in\mathscr{T}(1)}$ for problem (IT) by applying the procedure described in Corollary 4.11. By construction, the resulting solution $\{(s_t, u_t, v_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\}_{t\in\mathscr{T}(1)}$ is there-

fore optimal for problem (CT).

The naïve approach suffers from the following shortcoming. Once the reservoir targets $\boldsymbol{w}_1$ and the market bids $\{(s_t, u_t, v_t)\}_{t\in\mathscr{T}(1)}$ are fixed, the trader's problem (CT) reduces to a feasibility problem for the flow decisions. While we have proved that at least one set of feasible flow decisions exists (which is explicitly constructed by the naïve approach described above), common sense suggests that there are usually many flow decisions that meet the reservoir targets and are compatible with the market bids. Problem (CT) is indifferent between these feasible solutions and sets no incentive to exceed the reservoir targets, which have been chosen to be satisfiable even under the most adverse call-off scenarios on the reserve markets. Under other call-off scenarios, however, it is often possible to exceed the reservoir targets. This would be beneficial in view of Lemma 4.1, which implies that higher reservoir levels must lead to non-inferior future profits because excess water can always be spilled through to the dummy reservoir R. By constructing arbitrary feasible flow decisions, the naïve approach for solving (CT) may unnecessarily sacrifice future profits.

The second—more forward-looking—approach for solving the trader's problem (CT) aims at exceeding the reservoir targets. Specifically, among all feasible flow decisions, this approach seeks those that maximize the total value of all reservoirs, which is obtained by multiplying the vector of reservoir filling levels with a prescribed vector $\boldsymbol{\nu} \in \mathbb{R}^R$ [\$/m^3] of *water values*. While the proposed approach will be independent of the particular choice of $\boldsymbol{\nu}$, it is natural to set $\boldsymbol{\nu}$ to a supergradient of the planner's maximum expected profit with respect to the initial reservoir filling levels $\boldsymbol{w}_0$, which can be estimated by solving the dual of problem (C) under the Approximations 4.6, 4.19 and 4.20. Lemma 4.1 then implies that $\boldsymbol{\nu} \geq \mathbf{0}$. In addition, we have $\nu_R = 0$ because water in the dummy reservoir is lost for future energy production. Given the water values, we then construct the flow decisions $(\boldsymbol{g}_\theta, \boldsymbol{p}_\theta, \boldsymbol{z}_\theta)$ sequentially for each $\theta \in \mathscr{T}(1)$ by solving the multistage stochastic program

$$
\begin{array}{lll}
\sup & \boldsymbol{\nu}^\top \big(\boldsymbol{w}_0 + \sum_{\tau=1}^{\theta} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \big) & \\
\text{s.t.} & \boldsymbol{g}_\theta, \boldsymbol{p}_\theta, \boldsymbol{z}_\theta \in \mathbb{R}^A,\ \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathscr{L}^A(\mathscr{F}_{[\theta+1,t]}) & \forall\, t \in [\theta+1, H] \\
& \mathbf{0} \leq \boldsymbol{g}_t \leq \overline{\boldsymbol{g}}_t,\ \mathbf{0} \leq \boldsymbol{p}_t \leq \overline{\boldsymbol{p}}_t,\ \mathbf{0} \leq \boldsymbol{z}_t & \forall\, t \in [\theta, H],\ \mathbb{P}_{|\boldsymbol{\xi}_{[\theta]}}\text{-a.s.} \\
& s_t + \rho_t^{\mathrm{u}} u_t - \rho_t^{\mathrm{v}} v_t = \boldsymbol{\eta}_t^\top \boldsymbol{g}_t - \boldsymbol{\zeta}_t^\top \boldsymbol{p}_t & \forall\, t \in [\theta, H],\ \mathbb{P}_{|\boldsymbol{\xi}_{[\theta]}}\text{-a.s.} \\
& \underline{\boldsymbol{w}}_t \leq \boldsymbol{w}_0 + \sum_{\tau=1}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \leq \overline{\boldsymbol{w}}_t & \forall\, t \in [\theta, H],\ \mathbb{P}_{|\boldsymbol{\xi}_{[\theta]}}\text{-a.s.} \\
& \boldsymbol{w}_1 \leq \boldsymbol{w}_0 + \sum_{\tau\in\mathscr{T}(1)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) & \mathbb{P}_{|\boldsymbol{\xi}_{[\theta]}}\text{-a.s.}
\end{array}
\tag{4.15}
$$

with inputs $\boldsymbol{w}_0$, $\boldsymbol{w}_1$, $\{(s_t, u_t, v_t)\}_{t\in\mathscr{T}(1)}$ and $\{(\boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t=1}^{\theta-1}$. Problem (4.15) is designed to avoid unnecessary spillage in hour θ. Specifically, it seeks here-and-now decisions $(\boldsymbol{g}_\theta, \boldsymbol{p}_\theta, \boldsymbol{z}_\theta)$ that maximize the total value of the reservoirs at the end of hour θ, while ensuring that all end-of-day reservoir targets can be met and all market commitments can be honored by some future wait-and-see decisions $\{(\boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t=\theta+1}^{H}$. We emphasize that the flow decisions constructed with the naïve approach described above are always feasible in (4.15) and that the forward-looking approach described here is guaranteed to yield a feasible solution for the trader's

problem (CT).

As problem (4.15) constitutes a multistage stochastic program with random recourse, it appears to be computationally hard. Nevertheless, (4.15) is equivalent to the tractable linear program

$$\begin{array}{lll} \max & \boldsymbol{\nu}^\top \big(\boldsymbol{w}_0 + \sum_{\tau=1}^{\theta} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \big) & \\ \text{s.t.} & \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t \in \mathbb{R}^A & \forall\, t \in [\theta, H] \\ & \mathbf{0} \le \boldsymbol{g}_t \le \overline{\boldsymbol{g}}_t,\ \mathbf{0} \le \boldsymbol{p}_t \le \overline{\boldsymbol{p}}_t,\ \mathbf{0} \le \boldsymbol{z}_t & \forall\, t \in [\theta, H] \\ & s_\theta + \rho_\theta^{\mathrm{u}} u_\theta - \rho_\theta^{\mathrm{v}} v_\theta = \boldsymbol{\eta}_\theta^\top \boldsymbol{g}_\theta - \boldsymbol{\zeta}_\theta^\top \boldsymbol{p}_\theta & \\ & s_t + u_t = \boldsymbol{\eta}_t^\top \boldsymbol{g}_t - \boldsymbol{\zeta}_t^\top \boldsymbol{p}_t & \forall\, t \in [\theta + 1, H] \\ & \underline{\boldsymbol{w}}_t \le \boldsymbol{w}_0 + \sum_{\tau=1}^{t} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau) \le \overline{\boldsymbol{w}}_t & \forall\, t \in [\theta, H] \\ & \boldsymbol{w}_1 \le \boldsymbol{w}_0 + \sum_{\tau \in \mathcal{T}(1)} \boldsymbol{\phi}_\tau + \mathbf{M}(\boldsymbol{g}_\tau - \boldsymbol{p}_\tau + \boldsymbol{z}_\tau). & \end{array} \tag{4.16}$$

Proposition 4.21. Under Approximation 4.6, the optimal values of (4.15) and (4.16) are equal.

Proof. The proof reuses arguments that were developed to show the equivalence of (IT) and (IT$^{\mathrm{r}}$) as well as that of (CT) and (CT$^{\mathrm{r}}$). As no new ideas are needed, details are omitted for brevity. □

The proposed approach for constructing flow decisions is motivated by the solution degeneracy of problem (CT). Specifically, our approach seeks flow decisions that are not only compatible with the given reservoir targets and market commitments but also maximize the flexibility for future operations by saving as much water as possible. More generally, solution degeneracy poses a formidable challenge in robust optimization, where it has been recognized that some optimal solutions of a robust optimization problem perform significantly better than others in non-worst-case scenarios. General techniques for constructing Pareto-optimal solutions of robust optimization problems are described by Iancu and Trichakis (2014). Our approach is inspired by these techniques.

4.7 Case Study: Gasteiner Tal Cascade

We apply our planner-trader decomposition to a hydropower cascade located in the Gasteiner Tal, Austria. We describe the problem instance in Section 4.7.1, and we present our results in Section 4.7.2.

4.7.1 Problem Parameters

The Gasteiner Tal cascade comprises three reservoirs: the Bockhartsee annual reservoir with a capacity of 18,500,000 m^3, the Nassfeld daily reservoir with a capacity of 230,000 m^3 and the Remsach compensation reservoir with a capacity of 40,000 m^3. The Bockhartsee reser-

voir is connected to the Nassfeld reservoir by the Nassfeld pumped-storage plant that has a generating capacity of 40,600 m^3/h and a generation efficiency of 6.68×10^{-4} MWh/m^3, as well as a pumping capacity of 28,500 m^3/h and an inverse pumping efficiency of 9.35×10^{-4} MWh/m^3, respectively. The Nassfeld reservoir is connected to the Remsach reservoir by the Böckstein plant with a generating capacity of 41,400 m^3/h and an efficiency of 10.3×10^{-4} MWh/m^3, and the Remsach reservoir is connected to the Gasteiner Ache river by the Remsach plant with a generating capacity of 50,400 m^3/h and an efficiency of 5.41×10^{-4} MWh/m^3. Together, the cascade produces 264,000 MWh of electricity per year and covers the demand of approximately 75,000 households.[III]

We model the hourly natural inflows into the three reservoirs as

$$\phi_{t,r} = \left(\alpha_r \mu_{d(t),r} + \beta_{d(t),r}\right)^+ / H \quad \text{with} \quad \alpha_r \sim \mathcal{N}(1, \sigma_r^2) \text{ and } \beta_{d,r} \sim \mathcal{N}(0, \sigma_{d,r}^2),$$

where the average daily inflow $\mu_{d,r}$, $d \in \mathscr{D}$, into reservoir $r \in \mathscr{R}$ is derived from meteorological data. The affine transformation of the average daily inflows is determined by a random long-term factor α_r that is constant throughout the planning horizon and captures the hydrological year type (dry vs. wet year), as well as random additive noises $\beta_{d,r}$ that capture daily fluctuations. The variances σ_r^2 of α_r and $\sigma_{d,r}^2$ of $\beta_{d,r}$ are derived from the variations observed in the historical meteorological data. Note that this model implicitly assumes that every reservoir receives a constant rate of inflow throughout each day.

We further assume that the hourly spot prices fluctuate around a given deterministic price forward curve. Following Haarbrücker and Kuhn (2009), we set

$$\pi_t^{\mathrm{s}} = \pi_{0,t}^{\mathrm{s}} \cdot \frac{\exp(x_t)}{\mathbb{E}\left[\exp(x_t)\right]},$$

where $\{\pi_{0,t}^{\mathrm{s}}\}_{t \in \mathscr{T}}$ denotes the price forward curve observed at time 0,[IV] and where the risk factor x_t follows a mean-reverting process with a stochastic long-term mean y_t. Specifically, x_t and y_t are governed by the stochastic difference equations

$$x_{t+1} = x_t + \alpha(y_t - x_t) + \sigma_x \omega_t^x \quad \text{and} \quad y_{t+1} = y_t + \sigma_y \omega_t^y$$

initialized by $x_0 = y_0 = 0$; see Pilipovic (2007). Here, ω_t^x and ω_t^y follow i.i.d. standard Gaussian distributions. Note that this model satisfies $\mathbb{E}\left[\pi_t^{\mathrm{s}}\right] = \pi_{0,t}^{\mathrm{s}}$ for all $t \in \mathscr{T}$.

We assume that the hourly capacity prices satisfy $\pi_t^{\mathrm{u}} = \pi_t^{\mathrm{v}} = 0$ throughout the planning horizon. In fact, the capacity prices have eroded in the recent past, and they tend to be significantly smaller than the spot and the activation prices. Our assumption of vanishing capacity prices implies that we are underestimating the revenues to be earned on the reserve market. In line

[III] Further details of the cascade can be found online at `https://www.salzburg-ag.at/content/dam/web18/dokumente/unternehmen/erzeugung/Kraftwerke-GasteinerTal.pdf`.

[IV] The price forward curve was constructed from historical EEX spot prices, available at `https://www.eex.com`.

with historical averages, we draw the reserve activations $(\rho_t^{\mathrm{u}}, \rho_t^{\mathrm{v}})$ independently according to

$$(\rho_t^{\mathrm{u}}, \rho_t^{\mathrm{v}}) = \begin{cases} (0,0) & \text{with probability } 0.98, \\ (1,0) & \text{with probability } 0.01, \\ (0,1) & \text{with probability } 0.01. \end{cases}$$

This construction satisfies our assumptions that there are no simultaneous call-offs on the reserve-up and the reserve-down market, that all three call-off scenarios have a strictly positive probability for every $t \in \mathcal{T}$ and that the call-offs are serially independent as well as independent of all other sources of uncertainty. Finally, we assume that the activation prices ψ_t^{u} and ψ_t^{v} are independent of any information $\boldsymbol{\xi}_{[\Downarrow(t)]}$ available at the beginning of day $d(t)$, which implies that they enter problem $(\mathrm{C}^{\mathrm{r}})$ only through their unconditional expectations. We identify these expectations with their respective empirical averages, which we compute separately for peak and off-peak hours.

4.7.2 Results

We now apply the decomposition methods developed in this paper to the Gasteiner Tal case study from Section 4.7.1. In our experiments, we set $D = 364$, and we employ $N = 50$ scenarios for the sample average approximation. All results are averaged over 25 statistically independent runs.

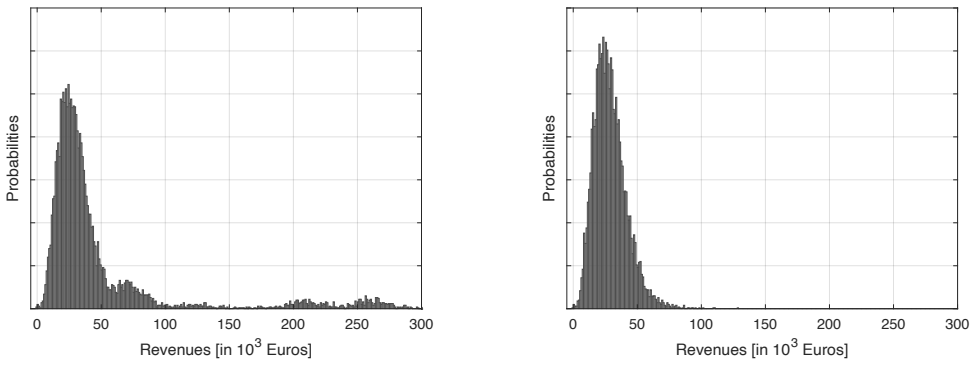


Figure 4.3 – Histogram of daily out-of-sample revenues for the spot-reserve (left) and the spot-only model (right). For given revenues on the x-axis, the y-axis reports the probability that any given day achieves these revenues.

We first evaluate the benefits of simultaneously participating on the spot and reserve markets. To this end, we compare the out-of-sample revenues generated by the collective bidding model (hereafter 'spot-reserve model') with those of a variant of the collective bidding model that only operates on the spot market, thus enforcing $u_t = v_t = 0$ $\mathbb{P}$-a.s. for all $t \in \mathcal{T}$ (hereafter 'spot-only model'). Figure 4.3 compares the daily revenues of both models. It turns out that the daily revenues of the spot-reserve model exceed those of the spot-only model only on 171.21 days on average (*i.e.*, with a 47.04% chance). The expected margin of the spot-reserve model

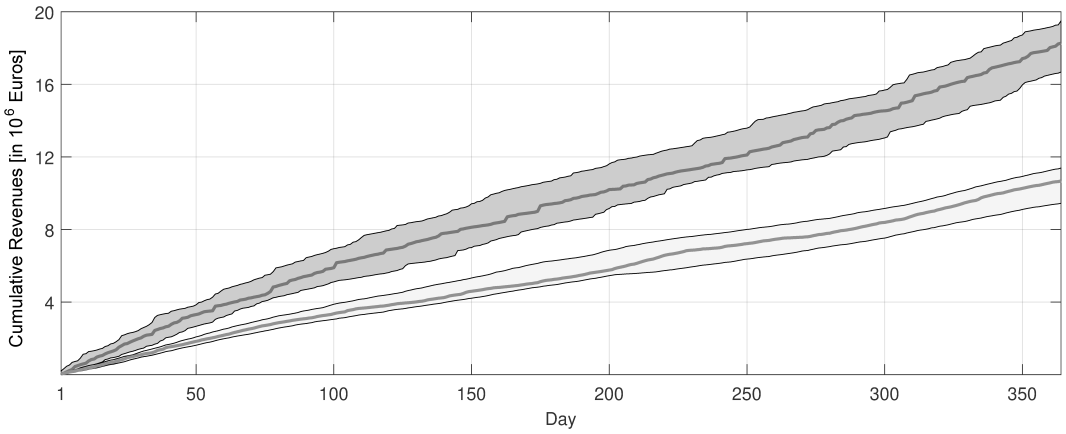


Figure 4.4 – Cumulative out-of-sample revenues of the spot-reserve (dark region) and the spot-only model (light region). The shaded regions show the intervals between the 10% and 90% percentiles, whereas the bold lines represent the medians.

over the spot-only model, conditioned on those days where the spot-reserve outperforms the spot-only model amounts to 46,184.03 €/d. However, the conditional expected margin of the spot-only model over the spot-reserve model is only 581.52 €/d. Figure 4.3 shows that this difference is caused by large tail margins that can be generated by participating in the reserve market. Even though these tail margins are realized with a small probability, they are crucial for the long-term profitability of the cascade. Indeed, Figure 4.4 compares the cumulative out-of-sample revenues generated by both models. The simultaneous bidding on the spot and reserve markets results in average daily revenues of 50,127.79 €/d, whereas the exclusive bidding on the spot market results in daily revenues of 28,761.69 €/d. In other words, trading on the reserve market enables the hydropower producer to increase the daily revenues by more than 74% on average.

A more detailed analysis of the revenues revealed that the spot-reserve model sells more energy on the spot market than the spot-only model and that it simultaneously offers all of this energy (as well as some excess energy) on the reserve-down market. This is caused by the attractive activation prices on the reserve-down market, and it explains why the spot-reserve model generates slightly lower revenues than the spot-only model on a typical day: By operating more aggressively on the spot market, the model accepts moderately lower daily revenues on 52.96% of the days in exchange for significantly larger daily revenues on 47.04% of the days. The spot-reserve model does not participate in the reserve-up market since the activation prices on that market are less attractive.

We next investigate the predictive accuracy of the spot-reserve model. To this end, Figure 4.5 visualizes a realization of the out-of-sample daily revenues across the time horizon, together with the in-sample predictions of the daily revenues on day 0 as well as day 180. We observe that the out-of-sample daily revenues tend to stay inside the predicted intervals, but that

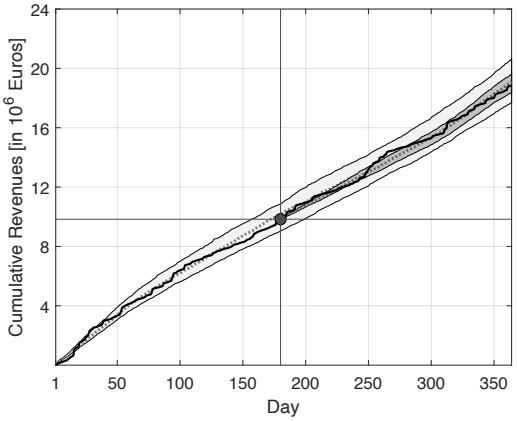


Figure 4.5 – Predicted vs. realized daily revenues in the spot-reserve model. The bold line shows a random realization of the realized out-of-sample revenues, whereas the light-shaded (dark-shaded) region represents the interval between the 10% and 90% percentiles of the future in-sample revenues predicted on day 0 (day 180).

the intervals widen as the model looks further into the future. We emphasize that the daily revenue predictions are updated every day, rather than only twice during the horizon as shown in the figure.

Figure 4.6 visualizes the daily out-of-sample filling levels of the three reservoirs. We observe that the decision rules correctly capture the different time scales of the reservoir dynamics: The filling level of the Bockhartsee annual reservoir changes slowly throughout the year, and its dynamics are mostly informed by the annual snow melt that happens around day 300-350, whereas the Nassfeld and the Remsach reservoir operate at a much faster pace. Note that in our experiments, all lower reservoir bounds are set to zero. In practice, one typically imposes strictly positive lower reservoir bounds to protect the aquatic life as well as to avoid a negative impact on tourism.

Decision rule architecture	**Avg. daily revenues**	**Percentage loss**
Perfect foresight solution	55,754.47 €	**0.00%**
$\boldsymbol{w}_d = \boldsymbol{\lambda}_d + \boldsymbol{\Phi}_d\left(\sum \boldsymbol{\phi}_\tau\right) + \Psi_d\left(\sum \boldsymbol{\phi}_\tau\right) + \boldsymbol{\mu}_d\left(\frac{1}{H}\sum \pi^{\mathrm{s}}_\tau\right)$	53,050.22 €	**-4.85%**
$\boldsymbol{w}_d = \boldsymbol{\lambda}_d + \boldsymbol{\Phi}_d\left(\sum \boldsymbol{\phi}_\tau\right) + \boldsymbol{\mu}_d\left(\frac{1}{H}\sum \pi^{\mathrm{s}}_\tau\right)$	52,915.51 €	**-5.09%**
$\boldsymbol{w}_d = \boldsymbol{\lambda}_d + \Psi_d\left(\sum \boldsymbol{\phi}_\tau\right) + \boldsymbol{\mu}_d\left(\frac{1}{H}\sum \pi^{\mathrm{s}}_\tau\right)$	52,654.82 €	**-5.56%**
$\boldsymbol{w}_d = \boldsymbol{\lambda}_d + \boldsymbol{\Phi}_d\left(\sum \boldsymbol{\phi}_\tau\right) + \Psi_d\left(\sum \boldsymbol{\phi}_\tau\right)$	52,194.82 €	**-6.38%**
$\boldsymbol{w}_d = \boldsymbol{\lambda}_d$	51,454.53 €	**-7.71%**

Table 4.1 – Average daily revenues as well as percentage losses (relative to the perfect foresight solution) of various decision rule architectures.

We close with an investigation of the suboptimality of our affine decision rule approximation.

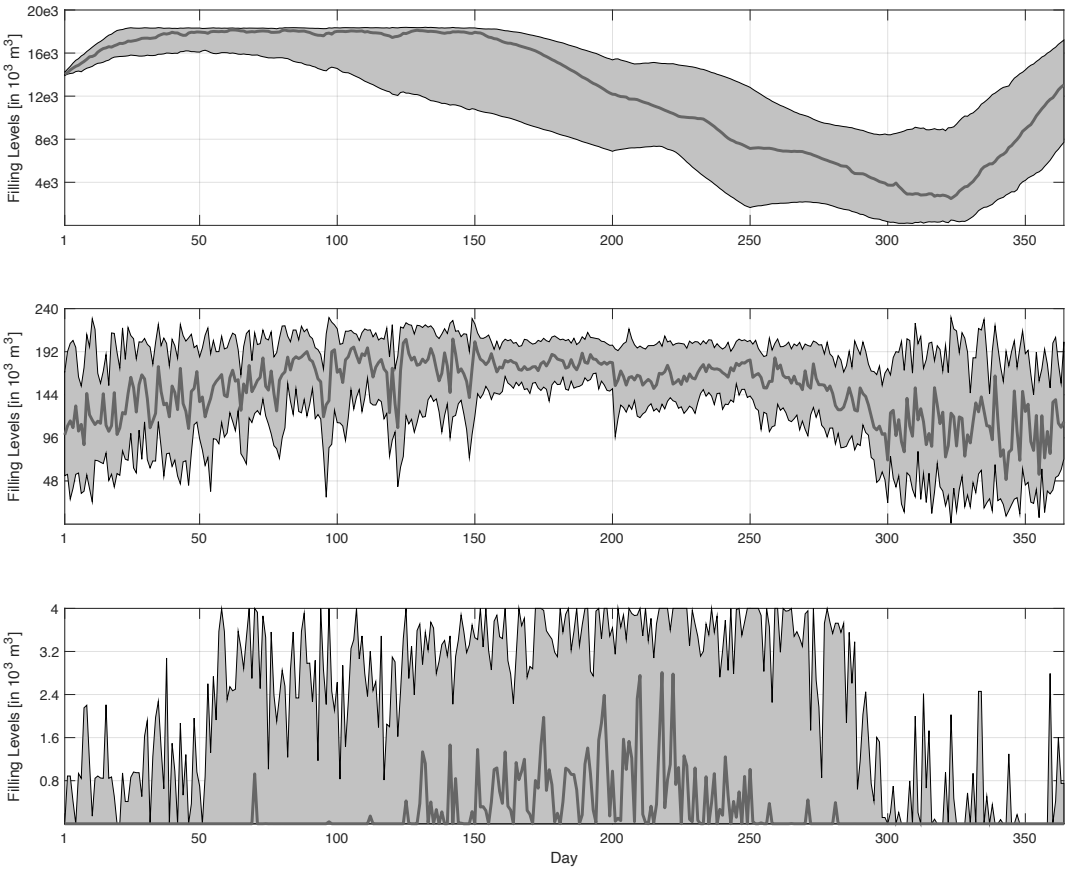


Figure 4.6 – Daily filling levels of the Bockhartsee, Nassfeld and Remsach reservoirs (top to bottom). The shaded regions represent the intervals between the 10% and 90% percentiles; the bold lines denote the median values.

To this end, Table 4.1 compares the average daily revenues generated by our decision rules,

$$\boldsymbol{w}_d = \boldsymbol{\lambda}_d + \boldsymbol{\Phi}_d \left(\sum_{\tau \in \mathscr{T}(d)} \boldsymbol{\phi}_\tau \right) + \Psi_d \left(\sum_{\tau=1}^{\uparrow(d-1)} \boldsymbol{\phi}_\tau \right) + \boldsymbol{\mu}_d \left(\frac{1}{H} \sum_{\tau \in \mathscr{T}(d)} \pi^{\mathrm{s}}_\tau \right)$$

(*cf.* Approximation 4.19), with the average daily revenues generated by more restricted decision rule architectures that set one or all of the feature weights $\boldsymbol{\Phi}_d$, Ψ_d and $\boldsymbol{\mu}_d$ to zero. We also report the average daily revenues generated by the (unachievable) perfect foresight solution that chooses all water levels after observing the entire realized sample path $\boldsymbol{\xi}_{[T]}$. We observe that each feature of the decision rule contributes to the overall revenues and that our decision rule approximation closes 37.11% of the gap between the static solution $\boldsymbol{w}_d = \boldsymbol{\lambda}_d$ and the perfect foresight solution.

Appendix

Consider a generic multistage stochastic program of the form

$$\begin{array}{ll} \sup & \mathbb{E}\big[f(\boldsymbol{x}_0,\ldots,\boldsymbol{x}_T,\boldsymbol{\xi}_{[T]})\big] \\ s.t. & \boldsymbol{x}_0 \in \mathscr{L}^{n_0}(\mathscr{F}_0),\ \boldsymbol{x}_t \in \mathscr{L}^{n_t}(\mathscr{F}_{[t]}) \quad \forall\, t = 1,\ldots,T, \end{array} \tag{4.17}$$

where $\mathbb{P}$ denotes the joint distribution of the random vectors $\boldsymbol{\xi}_t \in \mathbb{R}^{k_t}$, $t = 0,\ldots,T$, and f constitutes a normal integrand in the sense of (Rockafellar and Wets 2010, Definition 14.27), that is, its epigraph $\mathrm{epi}(f(\cdot,\boldsymbol{\xi}_{[T]}))$ is a closed-valued measurable multifunction of $\boldsymbol{\xi}_{[T]}$. We assume that f may adopt the value $+\infty$ and thereby encode $\mathbb{P}$-almost sure constraints on the decisions variables. We also retain the notational conventions for combined random vectors and σ-algebras introduced in Section 4.2. The proofs of Propositions 4.3 and 4.5 rely on the following lemma, which follows from Theorem 14.60 by Rockafellar and Wets (2010) and the properties of conditional expectations.

Lemma 4.22. The optimal value of the multistage stochastic program (4.17) equals

$$\mathbb{E}\left[\begin{array}{ll} \sup & \mathbb{E}\big[f(\boldsymbol{x}_0,\ldots,\boldsymbol{x}_T,\boldsymbol{\xi}_{[T]})\,\big|\,\boldsymbol{\xi}_0\big] \\ s.t. & \boldsymbol{x}_0 \in \mathbb{R}^{n_0},\ \boldsymbol{x}_t \in \mathscr{L}^{n_t}(\mathscr{F}_{[1,t]}) \quad \forall\, t = 1,\ldots,T \end{array}\right], \tag{4.18}$$

where $\mathbb{P}_{|\boldsymbol{\xi}_0}$ denotes the distribution of $\boldsymbol{\xi}_{[T]}$ conditional on $\boldsymbol{\xi}_0$.

Proof. As joint optimization over the stage-wise decision variables is equivalent to sequential optimization, we can rewrite the multistage stochastic program (4.17) as

$$\sup_{\boldsymbol{x}_0 \in \mathscr{L}^{n_0}(\mathscr{F}_0)} \ \sup_{\boldsymbol{x}_1 \in \mathscr{L}^{n_1}(\mathscr{F}_{[1]})} \cdots \sup_{\boldsymbol{x}_{T-1} \in \mathscr{L}^{n_{T-1}}(\mathscr{F}_{[T-1]})} \ \sup_{\boldsymbol{x}_T \in \mathscr{L}^{n_T}(\mathscr{F}_{[T]})} \mathbb{E}\big[f(\boldsymbol{x}_0,\ldots,\boldsymbol{x}_T,\boldsymbol{\xi}_{[T]})\big].$$

By (Rockafellar and Wets 2010, Theorem 14.60), the above problem is equivalent to

$$\sup_{\boldsymbol{x}_0 \in \mathscr{L}^{n_0}(\mathscr{F}_0)} \ \sup_{\boldsymbol{x}_1 \in \mathscr{L}^{n_1}(\mathscr{F}_{[1]})} \cdots \sup_{\boldsymbol{x}_{T-1} \in \mathscr{L}^{n_{T-1}}(\mathscr{F}_{[T-1]})} \mathbb{E}\left[\sup_{\boldsymbol{x}_T \in \mathbb{R}^{n_T}} f(\boldsymbol{x}_0,\ldots,\boldsymbol{x}_T,\boldsymbol{\xi}_{[T]})\right].$$

Recalling the tower property of conditional expectations, we can rewrite the resulting problem as

$$\sup_{\boldsymbol{x}_0 \in \mathscr{L}^{n_0}(\mathscr{F}_0)} \ \sup_{\boldsymbol{x}_1 \in \mathscr{L}^{n_1}(\mathscr{F}_{[1]})} \cdots \sup_{\boldsymbol{x}_{T-1} \in \mathscr{L}^{n_{T-1}}(\mathscr{F}_{[T-1]})} \mathbb{E}\left[\mathbb{E}\left[\sup_{\boldsymbol{x}_T \in \mathbb{R}^{n_T}} f(\boldsymbol{x}_0,\ldots,\boldsymbol{x}_T,\boldsymbol{\xi}_{[T]})\,\middle|\,\boldsymbol{\xi}_{[T-1]}\right]\right]$$

and invoke the interchangeability theorem of Rockafellar and Wets (2010) once again to prove its equivalence to

$$\sup_{\boldsymbol{x}_0 \in \mathscr{L}^{n_0}(\mathscr{F}_0)} \ \sup_{\boldsymbol{x}_1 \in \mathscr{L}^{n_1}(\mathscr{F}_{[1]})} \cdots \mathbb{E}\left[\sup_{\boldsymbol{x}_{T-1} \in \mathbb{R}^{n_{T-1}}} \mathbb{E}\left[\sup_{\boldsymbol{x}_T \in \mathbb{R}^{n_T}} f(\boldsymbol{x}_0,\ldots,\boldsymbol{x}_T,\boldsymbol{\xi}_{[T]})\,\middle|\,\boldsymbol{\xi}_{[T-1]}\right]\right]$$

Repeating this argument, it is possible to move all supremum operators inside the appropriate conditional expectation layers such as to obtain the equivalent problem

$$\mathbb{E}\left[\sup_{\boldsymbol{x}_0\in\mathbb{R}^{n_0}}\mathbb{E}\left[\sup_{\boldsymbol{x}_1\in\mathbb{R}^{n_1}}\mathbb{E}\left[\sup_{\boldsymbol{x}_2\in\mathbb{R}^{n_2}}\cdots\mathbb{E}\left[\sup_{\boldsymbol{x}_T\in\mathbb{R}^{n_T}} f(\boldsymbol{x}_0,\ldots,\boldsymbol{x}_T,\boldsymbol{\xi}_{[T-1]})\,\middle|\,\boldsymbol{\xi}_{[T-1]}\right]\cdots\,\middle|\,\boldsymbol{\xi}_{[1]}\right]\,\middle|\,\boldsymbol{\xi}_0\right]\right].$$

Applying the interchangeability theorem of Rockafellar and Wets (2010) in the reverse direction, we can then move the second supremum out of its conditional expectation layer to obtain

$$\mathbb{E}\left[\sup_{\boldsymbol{x}_0\in\mathbb{R}^{n_0}}\sup_{\boldsymbol{x}_1\in\mathscr{L}^{n_1}(\mathscr{F}_1)}\mathbb{E}\left[\sup_{\boldsymbol{x}_2\in\mathbb{R}^{n_2}}\cdots\mathbb{E}\left[\sup_{\boldsymbol{x}_T\in\mathbb{R}^{n_T}} f(\boldsymbol{x}_0,\ldots,\boldsymbol{x}_T,\boldsymbol{\xi}_{[T-1]})\,\middle|\,\boldsymbol{\xi}_{[T-1]}\right]\cdots\,\middle|\,\boldsymbol{\xi}_0\right]\right].$$

Iterating this argument from the outside to the inside, it is indeed possible to move all supremum operators into the outermost conditional expectation layer and thus obtain

$$\mathbb{E}\left[\sup_{\boldsymbol{x}_0\in\mathbb{R}^{n_0}}\sup_{\boldsymbol{x}_1\in\mathscr{L}^{n_1}(\mathscr{F}_1)}\sup_{\boldsymbol{x}_2\in\mathscr{L}^{n_2}(\mathscr{F}_{[1,2]})}\cdots\sup_{\boldsymbol{x}_T\in\mathscr{L}^{n_T}(\mathscr{F}_{[1,T]})}\mathbb{E}\Big[f(\boldsymbol{x}_0,\ldots,\boldsymbol{x}_T,\boldsymbol{\xi}_0)\,\Big|\,\boldsymbol{\xi}_0\Big]\right].$$

Re-combining the stage-wise supremum operators finally yields problem (4.18). □

The proof of Proposition 4.9 relies on a technical yet intuitive lemma, which asserts that there is no benefit in simultaneous generation and pumping.

Lemma 4.23. The optimal value of the reduced planner's problem (IT$^{\mathrm{r}}$) does not decrease if we append the complementarity constraints $\boldsymbol{g}_t \circ \boldsymbol{p}_t = \boldsymbol{0}$ for all $t \in \mathscr{T}(d)$.

Proof. Consider any feasible solution $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}_t, \boldsymbol{p}_t, \boldsymbol{z}_t)\}_{t\in\mathscr{T}(d)}$ of problem (IT$^{\mathrm{r}}$), and construct a new solution $\{(\boldsymbol{s}_t, \boldsymbol{u}_t, \boldsymbol{v}_t, \boldsymbol{g}'_t, \boldsymbol{p}'_t, \boldsymbol{z}'_t)\}_{t\in\mathscr{T}(d)}$ with adjusted flow decisions

$$\boldsymbol{g}'_t = \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{\Delta}_t, \quad \boldsymbol{p}'_t = \boldsymbol{p}_t - \boldsymbol{\eta}_t \circ \boldsymbol{\Delta}_t, \quad \boldsymbol{z}'_t = \boldsymbol{z} + (\boldsymbol{\zeta}_t - \boldsymbol{\eta}_t) \circ \boldsymbol{\Delta}_t,$$

where $\boldsymbol{\Delta}_t \in \mathbb{R}^A$ is defined through $\Delta_{t,a} = \min\{g_{t,a}/\zeta_{t,a}, p_{t,a}/\eta_{t,a}\}$. This new solution preserves the market decisions and consequently the objective value of the original solution, and it is readily seen to satisfy the complementarity constraints $\boldsymbol{g}'_t \circ \boldsymbol{p}'_t = \boldsymbol{0}$ for all $t \in \mathscr{T}(d)$. The claim thus follows if we can show that the new solution is feasible in (IT$^{\mathrm{r}}$). To this end, note that $\boldsymbol{\Delta}_t \geq \boldsymbol{0}$, which implies that $\boldsymbol{g}'_t \leq \boldsymbol{g}_t \leq \overline{\boldsymbol{g}}$, $\boldsymbol{p}'_t \leq \boldsymbol{p}_t \leq \overline{\boldsymbol{p}}$, and $\boldsymbol{z}'_t \geq \boldsymbol{0}$, where the last inequality exploits our standing assumption that $\boldsymbol{\zeta}_t > \boldsymbol{\eta}_t$. Similarly, the inequalities $\boldsymbol{\zeta}_t \circ \boldsymbol{\Delta}_t \leq \boldsymbol{g}_t$ and $\boldsymbol{\eta}_t \circ \boldsymbol{\Delta}_t \leq \boldsymbol{p}_t$ ensure that $\boldsymbol{g}'_t \geq \boldsymbol{0}$ and $\boldsymbol{p}'_t \geq \boldsymbol{0}$. Finally, observe that $\boldsymbol{g}'_t - \boldsymbol{p}'_t + \boldsymbol{z}'_t = \boldsymbol{g}_t - \boldsymbol{p}_t + \boldsymbol{z}_t$ (*i.e.*, the net reservoir outflows remain unchanged) and $\boldsymbol{\eta}_t \circ \boldsymbol{g}'_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}'_t = \boldsymbol{\eta}_t \circ \boldsymbol{g}_t - \boldsymbol{\zeta}_t \circ \boldsymbol{p}_t$ (*i.e.*, the arc-wise net energy production quantities remain unchanged), which implies that the new solution satisfies the reservoir balance constraints and the energy delivery constraints, respectively. In summary, we have shown that the new solution is indeed feasible in (IT$^{\mathrm{r}}$), and thus the claim follows. □

5 Conclusion

Since its inception in the 1950s, stochastic programming has advanced to one of the most widely adopted methodologies for addressing uncertainty in optimization problems. Despite its widespread use, stochastic programs are generically intractable, even if only two time stages are considered (Dyer and Stougie 2006a, Hanasusanto et al. 2016). While the computational burden of two-stage problems can be alleviated by approximating the probability distribution governing the random parameters through a discrete distribution supported on finitely many atoms, multi-stage stochastic programs are severely computationally intractable even if only approximate solutions are sought. Similar difficulties arise in the more recent distributionally robust optimization paradigm when the ambiguity sets are defined relative to nominal distributions; see, *e.g.*, Mohajerin Esfahani and Kuhn (2017).

Chapter 2 of this thesis revisited the scenario reduction problem, whose objective is to approximate a given probability distribution with one that has fewer atoms and that thus alleviates the computational burden of solving the associated stochastic program. By measuring the quality of the resulting approximation through the Wasserstein distance, we extended the existing body of literature on scenario reduction in several dimensions. In particular, we proved that we can restrict the search for an approximating probability distribution without much loss to those distributions whose atoms are contained in the support of the original distribution. We also showed that the scenario reduction problem, despite being computationally intractable, can be approximated to a constant factor by a variant of a well-known local search algorithm.

Chapter 3 studied the K-means clustering problem. This problem is a special case of the scenario reduction problem of Chapter 2 when the probability distribution is discrete and places equal weight on all atoms and when the approximation quality is measured by the type-2 Wasserstein distance induced by the Euclidean norm. We discussed several theoretical and practical drawbacks of the plain vanilla K-means clustering problem. These disadvantages motivated us to introduce cardinality constraints which limit the number of data points that each cluster can contain. We solved the emerging cardinality-constrained clustering problem by semidefinite and linear programming relaxations. We proved that our relaxations are tight for the balanced clustering problem as long as a separation condition holds, and we

showed experimentally that the relaxations remain almost tight even in unbalanced clustering problems where the separation condition is violated.

In Chapter 4, we determined an approximately optimal reservoir management and multi-market bidding strategy for a hydropower producer who simultaneously operates on both the spot and the reserve markets. While this problem can be formulated as a multi-stage stochastic program, its solution is very challenging due to the large number of time stages as well as the presence of a random recourse caused by the stochastic activation of the reserve capacities. By restricting the information base of certain decisions, we successively reduced the problem to a two-stage problem with fixed recourse that is amenable to an efficient solution. In other words, by disregarding some of the information available to the hydropower producer, we obtained a potentially less profitable policy that can however be computed efficiently. We carried out numerical experiments to showcase the benefits of our solution approach and the profitability of operating on the reserve market.

5.1 Future Research Avenues

While preparing the material of Chapters 2 to 4, we identified several promising areas for future research. It would be instructive, for example, to study how the scenario reduction approach of Chapter 2 could incorporate the risk preferences of the decision maker; see, *e.g.*, Arpon et al. (2018). Likewise, we have so far been unable to characterize the gap between the semidefinite and the linear programming relaxation for the cardinality-constrained clustering problem (*cf.* Chapter 3). While it is clear that the semidefinite relaxation is at least as tight as the linear one, it would be insightful to develop a priori bounds on their difference. Finally, our hydropower scheduling problem in Chapter 4 employs several simplifying assumptions that should be lifted in future work. In particular, we made the crucial assumption that the transmission system operator is allowed to call upon the producer's reserve capacities throughout the planning horizon (albeit with a small probability). In reality, the transmission system operator may not be able to exploit reserve capacities consecutively for a prolonged period of time. Likewise, we assumed that the hydropower producer is a price taker that faces exogenous capacity and activation fees, whereas in practice the hydropower producer submits an offer curve to the transmission system operator.

Bibliography

C. Aggarwal. *Outlier Analysis*. Springer, 2013.

F. Alizadeh and D. Goldfarb. Second-order cone programming. *Mathematical Programming*, 95(1): 3–51, 2003.

D. Aloise, A. Deshpande, P. Hansen, and P. Popat. NP-hardness of Euclidean sum-of-squares clustering. *Machine Learning*, 75(2):245–248, 2009.

B. Ames. Guaranteed clustering and biclustering via semidefinite programming. *Mathematical Programming*, 147(1–2):429–465, 2014.

K. Anstreicher. Semidefinite programming versus the reformulation-linearization technique for non-convex quadratically constrained quadratic programming. *Journal of Global Optimization*, 43: 471–484, 2009.

S. Arpon, T. H. de Mello, and B. Pagnoncelli. Scenario reduction for risk-averse stochastic programs. *Available on Optimization Online*, 2018.

D. Arthur and S. Vassilvitskii. k-means++: The advantages of careful seeding. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1027–1035, 2007.

V. Arya, N. Garg, R. Khandekar, A. Meyerson, K. Munagala, and V. Pandit. Local search heuristics for k-median and facility location problems. *SIAM Journal on Computing*, 33(3):544–562, 2004.

P. Awasthi, A. Bandeira, M. Charikar, R. Krishnaswamy, S. Villar, and R. Ward. Relax, no need to round: Integrality of clustering formulations. In *Conference on Innovations in Theoretical Computer Science 6*, pages 191–200, 2015.

M. Balcan, S. Ehrlich, and Y. Liang. Distributed k-means and k-median clustering on general topologies. In *Advances in Neural Information Processing Systems 26*, pages 1995–2003. 2013.

G.-Y. Ban and C. Rudin. The big data newsvendor: Practical insights from machine learning. *Operations Research*, 67(1):90–108, 2019.

A. Banerjee and J. Ghosh. Scalable clustering algorithms with balancing constraints. *Data Mining and Knowledge Discovery*, 13(3):365–395, 2006.

M. Beck and M. Scherer. Basic principles of ancillary service products. Swissgrid Ltd., 2015.

A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004a.

A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjusting robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004b.

K. Bennett, P. Bradley, and A. Demiriz. Constrained K-means clustering. Technical Report, Microsoft Research, 2000.

A. Bertoni, M. Goldwurm, J. Lin, and F. Saccà. Size constrained distance clustering: separation properties and some complexity results. *Fundamenta Informaticae*, 115(1):125–139, 2012.

D. Bertsimas and N. Kallus. From predictive to prescriptive analytics. *Management Science*, 66(3): 1025–1044, 2020.

M. Bodur and J. R. Luedtke. Two-stage linear decision rules for multi-stage stochastic programming. *Mathematical Programming*, 2018. (in press).

S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. ISBN 0521833787.

R. Burkard. The quadratic assignment problem. In P. Pardalos, D. Du, and R. Graham, editors, *Handbook of Combinatorial Optimization*, pages 2741–2814. Springer, 2013.

R. Burkard, M. Dell'Amico, and S. Martello. *Assignment Problems*. SIAM, 2009.

P. Carpentier, J.-P. Chancelier, and M. de Lara. Algorithms for two-time scales stochastic optimization with applications to long term management of energy storagey. *HAL archives-ouvertes.fr*, 2019.

M. Charikar and S. Li. A dependent LP-rounding approach for the k-median problem. In *Proceedings of the 39th International Colloquium Conference on Automata, Languages, and Programming*, pages 194–205, 2012.

S. Chawla and A. Gionis. k-means--: A unified approach to clustering and outlier detection. In *SIAM International Conference on Data Mining 13*, pages 189–197, 2013.

X. Chen, M. Sim, P. Sun, and J. Zhang. A linear decision-based approximation approach to stochastic programming. *Operations Research*, 56(2):344–357, 2008.

Y. Chen, Y. Zhang, and X. Ji. Size regularized cut for data clustering. In *Advances in Neural Information Processing Systems 18*, pages 211–218. 2006.

A. Conejo, M. Carrión, and J. Morales. *Decision Making Under Uncertainty in Electricity Markets*. Springer, 2010a.

A. J. Conejo, M. Carrión, and J. M. Morales. *Decision Making Under Uncertainty in Electricity Markets*. Springer, 2010b.

W. Cook and A. Rohe. Computing minimum-weight perfect matchings. *INFORMS Journal on Computing*, 11(2):138–148, 1999.

L. Costa, D. Aloise, and N. Mladenović. Less is more: basic variable neighborhood search heuristic for balanced minimum sum-of-squares clustering. *Information Sciences*, 415–416:247–253, 2017.

S. Dasgupta. CSE 291: Topics in unsupervised learning, 2008. URL http://cseweb.ucsd.edu/~dasgupta/291-unsup/.

Z. Drezner and H. Hamacher. *Facility Location: Applications and Theory*. Springer, 2004.

J. Dupačová. Stability and sensitivity-analysis for stochastic programming. *Annals of Operations Research*, 27(1):115–142, 1990.

J. Dupačová, N. Gröwe-Kuska, and W. Römisch. Scenario reduction in stochastic programming: an approach using probability metrics. *Mathematical Programming*, 95(3):493–511, 2003.

M. Dyer and L. Stougie. Computational complexity of stochastic programming problems. *Mathematical Programming*, 106(3):423–432, 2006a.

M. Dyer and L. Stougie. Computational complexity of stochastic programming problems. *Mathematical Programming*, 106(3):423–432, 2006b.

E. Elhamifar, G. Sapiro, and R. Vidal. Finding exemplars from pairwise dissimilarities via simultaneous sparse recovery. In *Advances in Neural Information Processing Systems 25*, pages 19–27. 2012.

R. Gao and A. Kleywegt. Distributionally robust stochastic optimization with Wasserstein distance, 2016. *arXiv* #1604.02199.

J. Gareth, D. Witten, T. Hastie, and R. Tibshirani. *An Introduction to Statistical Learning with Applications in R*. Springer, 2017.

K. Gatermann and P. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1–3):95–128, 2004.

C. Gauvin, E. Delage, and M. Gendreau. Decision rule approximations for the risk averse reservoir management problem. *European Journal of Operational Research*, 261(1):317–336, 2017.

C. Gauvin, E. Delage, and M. Gendreau. A stochastic program with time series and affine decision rules for the reservoir management problem. *European Journal of Operational Research*, 267(2): 716–732, 2018.

G. Golub and C. Loan. *Matrix Computations*. John Hopkins University Press, 1996.

S. Graf and H. Luschgy. *Foundations of Quantization for Probability Distributions*. Springer, 2000.

R. M. Gray. *Toeplitz and Circulant Matrices: A Review*. Now Publishers, 2006.

G. Haarbrücker and D. Kuhn. Valuation of electricity swing options by multistage stochastic programming. *Automatica*, 45(4):889 – 899, 2009.

G. Hanasusanto and D. Kuhn. Conic programming reformulations of two-stage distributionally robust linear programs over Wasserstein balls, 2016. *arXiv* #1609.07505.

G. Hanasusanto, D. Kuhn, and W. Wiesemann. A comment on "computational complexity of stochastic programming problems". *Mathematical Programming*, page (Forthcoming), 2015a.

G. Hanasusanto, D. Kuhn, and W. Wiesemann. k-adaptability in two-stage robust binary programming. *Operations Research*, 63(4):877–891, 2015b.

G. Hanasusanto, D. Kuhn, and W. Wiesemann. A comment on "computational complexity of stochastic programming problems". *Mathematical Programming*, 159(1):557–569, 2016.

S. Hasegawa, H. Imai, M. Inaba, N. Katoh, and J. Nakano. Efficient algorithms for variance-based k-clustering. In *Pacific Conference on Computer Graphics and Applications 1*, pages 75–89, 1993.

H. Heitsch and W. Römisch. Scenario reduction algorithms in stochastic programming. *Computational Optimization and Applications*, 24(2):187–206, 2003.

H. Heitsch and W. Römisch. A note on scenario reduction for two-stage stochastic programs. *Operations Research Letters*, 35(6):731 – 738, 2007.

L. Hirth and I. Ziegenhagen. Balancing power and variable renewables: Three links. *Renewable and Sustainable Energy Reviews*, 50:1035–1051, 2015. ISSN 1364-0321.

R. Hochreiter and G. Pflug. Financial scenario generation for stochastic multi-stage decision processes as facility location problems. *Annals of Operations Research*, 152(1):257–272, 2007.

J. Hopcroft and R. Kannan. Computer science theory for the information age, 2012. URL https://www.cs.cmu.edu/~venkatg/teaching/CStheory-infoage/.

D. A. Iancu and N. Trichakis. Pareto efficiency in robust optimization. *Management Science*, 60(1): 130–147, 2014.

T. Iguchi, D. Mixon, J. Peterson, and S. Villar. Probably certifiably correct k-means clustering. *Mathematical Programming*, 165(2):605–642, 2017.

M. Inaba, N. Katoh, and I. Hiroshi. Applications of weighted Voronoi diagrams and randomization to variance-based k-clustering. In *Symposium on Computational Geometry 10*, pages 332–339. 1994.

A. Jain. Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 31(8):651–666, 2010a.

A. Jain. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8):651–666, 2010b.

A. Jain, M. Murty, and P. Flynn. Data clustering: a review. *ACM Computing Surveys*, 31(3):264–323, 1999.

T. Kanungo, D. Mount, N. Netanyahu, C. Piatko, R. Silverman, and A. Wu. A local search approximation algorithm for k-means clustering. *Computational Geometry*, 28(2):89–112, 2004.

O. Kariv and S. Hakimi. An algorithmic approach to network location problems. ii: The p-medians. *SIAM Journal on Applied Mathematics*, 37(3):539–560, 1979.

D. Kuhn, W. Wiesemann, and A. Georghiou. Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming*, 130(1):177–209, 2011.

S. Li and O. Svensson. Approximating k-median via pseudo-approximation. *SIAM Journal on Computing*, 45(2):530–547, 2016.

S. Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982.

N. Löhndorf. An empirical analysis of scenario generation methods for stochastic optimization. *European Journal of Operational Research*, 255(1):121–132, 2016.

N. Löhndorf, D. Wozabal, and S. Minner. Optimizing trading decisions for hydro storage systems using approximate dual dynamic programming. *Operations Research*, 61(4):810–823, 2013.

M. Mahajan, P. Nimbhorkar, and K. Varadarajan. The planar k-means problem is NP-hard. In *Proceedings of the 3rd International Workshop on Algorithms and Computation*, pages 274–285, 2009.

M. Malinen and P. Fränti. Balanced k-means for clustering. In *Structural, Syntactic, and Statistical Pattern Recognition*, pages 32–41, 2014.

J. Mayer. Electricity production and spot-prices in Germany 2014. Fraunhofer Institute for Solar Energy Systems, 2014.

P. Mohajerin Esfahani and D. Kuhn. Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations, 2017. URL https://doi.org/10.1007/s10107-017-1172-1. Mathematical Programming.

C. Morris and M. Pehnt. Energy transition: The German energiewende. Heinrich Böll Foundation, 2015.

J. Mulvey and M. Beck. Solving capacitated clustering problems. *European Journal of Operational Research*, 18(3):339–348, 1984.

A. Nellore and R. Ward. Recovery guarantees for exemplar-based clustering. *Information and Computation*, 245(C):165–180, 2015. ISSN 0890-5401.

M. Parvania and M. Fotuhi-Firuzabad. Demand response scheduling by stochastic SCUC. *IEEE Transactions on Smart Grid*, 1(1):89–98, 2010.

J. Peng and Y. Wei. Approximating K-means-type clustering via semidefinite programming. *SIAM Journal on Optimization*, 18(1):186–205, 2007.

G. Pflug. Scenario tree generation for multiperiod financial optimization by optimal discretization. *Mathematical Programming*, 89(2):251–271, 2001.

G. Pflug and A. Pichler. Approximations for probability distributions and stochastic optimization problems. In M. Bertocchi, G. Consigli, and M. A. H. Dempster, editors, *Stochastic Optimization Methods in Finance and Energy: New Financial Products and Energy Market Strategies*, pages 343–387. Springer, 2011.

D. Pilipovic. *Energy Risk: Valuing and Managing Energy Derivatives*. McGraw-Hill, 2007.

G. Pritchard, A. Philpott, and P. Neame. Hydroelectric reservoir optimization in a pool market. *Mathematical Programming*, 103(3):445–461, 2005.

A. Pyatkin, D. Aloise, and N. Mladenović. NP-hardness of balanced minimum sum-of-squares clustering. *Pattern Recognition Letters*, 97:44–45, 2017.

S. Rachev and W. Römisch. Quantitative stability in stochastic programming: The method of probability metrics. *Mathematics of Operations Research*, 27(4):792–818, 2002.

R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer, 2010. ISBN 9783642083044.

W. Römisch and R. Schultz. Stability analysis for stochastic programs. *Annals of Operations Research*, 30(1):241–266, 1991.

W. Römisch and R. Wets. Stability of ε-approximate solutions to convex stochastic programs. *SIAM Journal on Optimization*, 18(3):961–979, 2007.

R. Rubinstein and D. Kroese. *Simulation and the Monte Carlo Method*. Wiley, 2007.

P. Ruiz, C. Philbrick, E. Zak, K. Cheung, and P. Sauer. Uncertainty management in the unit commitment problem. *IEEE Transactions on Power Systems*, 24(2):642–651, 2009.

A. Shapiro and A. Nemirovski. On complexity of stochastic programming problems. In V. Jeyakumar and A. Rubinov, editors, *Continuous Optimization: Current Trends and Modern Applications*, pages 111–146. Springer, 2005.

A. Shapiro, D. Dentcheva, and A. Ruszczynski. *Lectures on Stochastic Programming: Modeling and Theory*. SIAM, 2009.

M. Steele. *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*. Cambridge University Press, 2004.

R. Stockbridge and G. Bayraksan. A probability metrics approach for reducing the bias of optimality gap estimators in two-stage stochastic linear programming. *Mathematical Programming*, 142(1): 107–131, 2013.

P. Vayanos, D. Kuhn, and B. Rustem. A constraint sampling approach for multi-stage robust optimization. *Automatica*, 48(3):459–471, 2012.

R. Vinayak and B. Hassibi. Similarity clustering in the presence of outliers: Exact recovery via convex program. In *IEEE International Symposium on Information Theory*, pages 91–95, 2016.

H. Wirth. Recent facts about photovoltaics in Germany. Fraunhofer Institute for Solar Energy Systems, 2016.

D. Wozabal and G. Rameseder. Optimal bidding of a virtual power plant on the spanish day-ahead and intraday market for electricity. *European Journal of Operational Research*, 280(2):639–655, 2020.

H. Zha, X. He, C. Ding, H. Simon, and M. Gu. Spectral relaxation for K-means clustering. In *Advances in Neural Information Processing Systems 14*, pages 1057–1064, 2002.

H. Zhang, C. Royo, and L. Ma. Solving the quadratic assignment problem by means of general purpose mixed integer linear programming solvers. *Annals of Operations Research*, 207:261–278, 2013.

C. Zhao and Y. Guan. Data-driven risk-averse stochastic optimization with Wasserstein metric. *Available on Optimization Online*, 2015.

Curriculum Vitae Kilian Schindler

1. Personal Information

Date of Birth: October 29, 1989 - in Zürich, Switzerland
Citizenships: Switzerland, and Italy (from mother), and Chile (from father)
Address: Chemin du Coteau 23, 1009 Pully, Switzerland
Contact: kilian.schindler@epfl.ch | +41 76 438 78 78

2. Summary

Kilian studied Operations Research at UC Berkeley and Mechanical Engineering at ETH Zürich. Currently, he is a PhD candidate at EPFL (École Polytechnique Fédérale de Lausanne) in Switzerland, where he works together with Prof. Daniel Kuhn on dynamic data-driven optimization under uncertainty. Specifically, Kilian develops quantitative methods to handle the increasing dimensionality of modern optimization problems and to ensure their robustness against uncertainty. Impact areas of his work so far include Operations Research, Machine Learning, and the Renewable Energy Sector.

Kilian grew up in Zürich, Switzerland, and is the eldest son of his Italian mother and Chilean father. He is fluent in five languages, passionate about teaching, and in his spare time he likes to row, hike, and travel.

3. Education

Ph.D. in Operations Research, EPFL (École Polytechnique Fédérale de Lausanne) Sep/2016 - present

Supervisor:	Prof. Daniel Kuhn - Chair of Risk Analytics and Optimization
Research:	Dynamic Data-Driven Optimization under Uncertainty (Scalability and Robustness)
Papers:	K. Schindler, N. Rujeerapaiboon, D. Kuhn, W. Wiesemann. A Day-Ahead Decision Rule Method for Multi-Market Multi-Reservoir Management. *Working Paper.*
	N. Rujeerapaiboon, K. Schindler, D. Kuhn, W. Wiesemann. Size Matters: Cardinality Constrained Clustering and Outlier Detection via Conic Optimization. *SIAM Journal on Optimization. April 2019.*
	N. Rujeerapaiboon, K. Schindler, D. Kuhn, W. Wiesemann. Scenario Reduction Revisited: Fundamental Limits and Guarantees. *Mathematical Programming. April 2018.*
Conferences:	Intern. Conf. on Continuous Optimization | TU Berlin, Germany | August 2019 Intern. Conf. on Stochastic Programming | NTNU Trondheim, Norway | July 2019 Intern. Symp. on Mathematical Programming | University of Bordeaux, France | July 2018 Computational Management Science | NTNU Trondheim, Norway | May 2018 Modern Convex Optimization and Applications | The Fields Institute, Canada | July 2017 Computational Management Science | University of Bergamo, Italy | May 2017
Teaching:	Best Teaching Assistant Award 2019 | awarded annually by the graduating student cohort Convex Optimization | Prof. Daniel Kuhn | Fall 2018; Fall 2019 Data Science for Managers | Prof. Kenneth Younge | Fall 2018; Spring, Summer, Fall 2019 Applied Probability and Stochastic Processes | Dr. Napat Rujeerapaiboon | Fall 2017 Data Science for Business | Prof. Kenneth Younge | Spring 2017
Coursework:	Advanced Topics in Machine Learning | Mathematics of Data | Data Science for Business | Optimization Methods and Models | Optimization and Simulation | Advanced Probability and Applications | Supply Chain Management | Microeconomics | Econometrics

M.Eng. in Industrial Engineering and Operations Research, UC Berkeley Aug/2015 - May/2016

Overall GPA:	3.88 (max: 4.0) - among top 3% of the overall promotion
Capstone:	Fault-Tolerant Localization in Autonomous Driving using GPS, Camera, and LIDAR
Coursework:	Optimization Analytics | Mathematical Programming | Applied Stochastic Processes | Computational Optimization | Learning and Optimization | Engineering Leadership

M.Sc. with Distinction in Mechanical Engineering, ETH Zürich Feb/2012 - Dec/2013

Overall GPA:	5.81 (max: 6.0) - among top 3% of the overall promotion

Thesis: Chattering Dynamics and related Nonlinear Motion in the Bouncing Ball System

Paper: K. Schindler, R. Leine. Paradoxical Chaos-Like Chattering in the Bouncing Ball System. *ASME International Design Engineering Technical Conferences. August 2018.*

Coursework: Theory of Robotics and Mechatronics | Dynamic Programming and Optimal Control | Model Predictive Control | Recursive Estimation | Linear Systems Theory | Nonlinear Dynamics | Dynamics of Multi-Body / Structure-Variant Systems | Orbital Dynamics

B.Sc. in Mechanical Engineering, ETH Zürich Sep/2008 - Sep/2011

Overall GPA: 5.20 (max: 6.0) - among top 10% of the overall promotion

Thesis: Programming a 7-axis robot arm to find the shortest path in a maze using haptic feedback

Coursework: Calculus | Linear Algebra | Probability and Statistics | Physics | Mechanics | Electronics | Thermodynamics | Fluid Dynamics | System Dynamics and Control | Manufacturing

Matura (High-School Graduation), Freies Gymnasium Zürich 2002 - 2008

Overall GPA: 5.89 (max: 6.0) - best of the overall promotion

4. Experience

Member of the Autonomous Driving Capstone Team at UC Berkeley - CA / USA Aug/2015 - May/2016

The advancement of Autonomous Driving Technology involves finding solutions to many challenges and integrating them into a safe overall system. Our team worked towards a solution for localization, that is, determining in real time the exact location of the vehicle. By fusing GPS, camera, and LIDAR sensors, we developed an algorithm to handle the potential loss of the GPS signal and implemented it as a stand-alone software module, which was deployed on a prototype autonomous car and used for ongoing research.

Full-time position as Vehicle Dynamics Engineer at Stadler Rail - Switzerland May/2014 - Jul/2015

Modeling the train as multi-body mechanical system, simulating its dynamic behavior using appropriate software and conducting the required prototype tests to ensure safety against derailment, optimal riding comfort, running stability, and other specific properties. The obtained insights determine various design parameters of the train suspension and inform the legal authorization procedure for operating the train.

A total of 14 appointments as Teaching Assistant across 7 different courses at ETH Zürich 2009 - 2013

Appointments were with three different institutes, namely: Institute for Dynamic Systems and Control, Institute of Mechanical Systems, and Institute of Machine Tools and Manufacturing.

The main responsibilities included: preparation and delivery of weekly classes to groups of 30+ students, supervising and correcting mid-term exams, and acting as first person of contact for students. Over the years I had the pleasure to teach more than 500 hours to a total of about 600 students at ETH Zürich.

Industrial Internships

R&D internship at Trumpf Maschinen AG in Baar/Zug Sep/2011 - Jan/2012

Production floor internship at Kaba AG in Wetzikon/Zürich Winter 2009

5. Other

Languages: Fluent in English, German, Italian, Spanish, and French

Software: Matlab, Python, C/C++ | Gurobi, Cplex, Mosek | MS Office, LaTeX

Interests: Technology, International Affairs, Teaching, Sports (Rowing/Hiking), and Travelling

- Member of the "Zunft zum Kämbel" (one of Zürich's traditional "Guilds") - since 2010
- Board Member of the Alumni Association of the Freies Gymnasium Zürich (2011 - 2016)
- Vice-President of the Student Association of the Freies Gymnasium Zürich (2006 - 2008)
- Selected and annually confirmed as Member of the Swiss Study Foundation (2009 - 2014)
- Member of the local Polling Bureau at the commune of Kilchberg/Zürich (2010 - 2018)

I have traveled extensively across Europe and have been to North- and South-America. I have visited wide parts of Australia, some countries in South-East Asia, as well as the Gulf Emirates.