Exact and Heuristic Solution Techniques for Mixed-Integer Quantile Minimization Problems

Diego Cattaruzza, Martine Labbé, Matteo Petris,
Marius Roland, Martin Schmidt

Abstract. We consider mixed-integer linear quantile minimization problems that yield large-scale problems that are very hard to solve for real-world instances. We motivate the study of this problem class by two important real-world problems: a maintenance planning problem for electricity networks and a quantile-based variant of the classic portfolio optimization problem. For these problems, we develop valid inequalities and present an overlapping alternating direction method. Moreover, we discuss an adaptive scenario clustering method for which we prove that it terminates after a finite number of iterations with a global optimal solution. We study the computational impact of all presented techniques and finally show that their combination leads to an overall method that can solve the maintenance planning problem on large-scale real-world instances provided by the ROADEF/EURO challenge 2020[1] and that they also lead to significant improvements when solving a quantile-version of the classic portfolio optimization problem.

1. Introduction

Many real-world planning and investment problems face a significant amount of uncertainty since they inevitably need to incorporate aspects that lie in the future and are thus unknown at the time of decision making. Consequently, usual objective functions in this context combine the minimization of expected costs (or the maximization of expected profits) with some kind of risk minimization. In this paper, we consider general planning and investment problems in which we minimize a convex combination of the expected costs and the cost's quantile. In other words, the objective function is a combination of the expected value and the Value at Risk (VaR) of some function that is linear in the problem's variables. It is well known that VaR is nothing but the τ-quantile. It is a measure of risk used in various domains. In portfolio optimization, [8, 9] consider the problem of maximizing the VaR subject to a lower bound on the expected return while [1, 2, 4, 17] maximize the expected return given a lower bound on the VaR.

Limiting the VaR of a random variable is a particular type of a chance constraint since it is equivalent to setting up a lower bound on the probability that the random variable takes a value larger than the said limit. Chance-constrained formulations have been proposed for various applications such as the design of reliable networks [28], the packing of objects with random weights [29], or the allocation of scarce vaccines to prevent the occurrence of disease epidemics [30]. Moreover, the VaR is also used in real-world regulatory frameworks such as Basel or Solvency.

Date: July 4, 2023.
2020 *Mathematics Subject Classification.* 90B25, 90C11, 90C90, 90C15.
Key words and phrases. Quantile Minimization, Value-at-Risk (VaR), Mixed-Integer Optimization, Valid Inequalities, Adaptive Clustering.

[1]See https://www.roadef.org/challenge/2020/en/index.php.

Besides considering the quantile minimization, which already poses a computational challenge on its own, we study both continuous and mixed-integer linear settings, which are required by many real-world problems to properly model planning or investment decisions. Thus, in total, we consider the challenging class of mixed-integer linear quantile minimization problems. To this end, stochasticity is modeled via finite scenario sets, which leads to large-scale mixed-integer linear problems that can hardly be solved with state-of-the-art solvers.

Throughout the paper we use two examples for the general class of problems under consideration: A maintenance planning problem in electricity networks as it was posed in the ROADEF/EURO challenge 2020 and a variant of the classic portfolio optimization problem. The grid operation based outage maintenance planning problem (MPP) consists in determining the start time of maintenance interventions in a high-voltage transmission network over a given time horizon. Each of the interventions lasts a certain number of time units that depends on the start time of the intervention. All interventions must be planned and finished before the end of the time horizon. Further, some interventions cannot take place at the same time. Finally, each intervention consumes resources and the total amount of resources used at each time step is bounded from below and above. The objective is to minimize the risk of the maintenance plan. More precisely, a set of scenarios is given and for each such scenario, we know, at each time period, the cost of each intervention. The goal is to minimize a combination of the expectation and the quantile of the cost.

The second problem is a variant of the well-known portfolio optimization problem. The goal in portfolio optimization is twofold: maximize the return and minimize the risk for which different measures have been proposed; see, e.g., [4, 9, 17, 19]. Among them, the VaR or τ-quantile has attracted particular attention, namely because it is used to measure market risk by regulators; see, e.g., [3] and the references therein.

As we already mentioned above, the studied models lead to large-scale mixed-integer linear problems (MILPs). For these problems, we develop tailored solution techniques. In Section 2, we introduce the general problem class and the two specific examples. In Section 3, we propose problem-tailored valid inequalities. They are derived from duality theory applied to a properly chosen linear optimization problem that models the quantile. In Section 4, we present an overlapping alternating direction method that serves as a primal heuristic for quickly computing feasible points of good quality. In Section 5, we then present an adaptive scenario clustering method for which we prove that it computes an approximate global optimal solution after finitely many iterations. We illustrate the computational impact of all presented techniques in our numerical study in Section 6 before we close this paper with some concluding remarks and a brief discussion of potential future research in Section 7.

2. Problem Statement

In this section, we first state the general problem class that we consider in the following. Afterward, we present two specific examples for the general modeling framework to underline the importance of the studied class of problems.

We consider a discrete set of indices $t \in \mathcal{T} = \{1, \dots, T\}$. With this at hand, the general problem is given by

$$\min_{x} \quad \alpha \sum_{t \in \mathcal{T}} \mathbb{E}[c_t^\top x] + (1-\alpha) \sum_{t \in \mathcal{T}} f\left(\mathbb{Q}[c_t^\top x]\right) \tag{1a}$$

$$\text{s.t.} \quad x \in X \subseteq \mathbb{R}^n. \tag{1b}$$

The feasible set X is a non-empty and closed set that may also include integrality restrictions for all or some of the variables. For each index t, we are given a finite

set $\mathcal{S}_t$ of scenarios and for each scenario $s \in \mathcal{S}_t$, c_t^s is the respective cost vector and $p_t^s \in [0,1]$ is the associated probability with $\sum_{s\in\mathcal{S}_t} p_t^s = 1$. The expected value is then defined as

$$\mathbb{E}[c_t^\top x] = \sum_{s\in\mathcal{S}_t} p_t^s (c_t^s)^\top x$$

and the τ-quantile is given by

$$\mathbb{Q}[c_t^\top x] = \min\left\{ q \in \mathbb{R} \colon \sum_{s\in\mathcal{N}(q,c_t^\top x)} p_t^s \geq \tau \right\}, \quad \mathcal{N}(q, c_t^\top x) = \left\{ s \in \mathcal{S}_t \colon (c_t^s)^\top x \leq q \right\}.$$

Further, $\alpha \in [0,1]$ is a scaling factor that either puts more emphasis on the expected value terms $\mathbb{E}[c_t^\top x]$ or on the τ-quantile terms $f(\mathbb{Q}[c_t^\top x])$, where f is an arbitrary function depending on the τ-quantile $\mathbb{Q}[\cdot]$.

Usually, the τ-quantile $\mathbb{Q}[\cdot]$ cannot be stated in closed form. However, it can be expressed by the solution of the following quantile optimization problem in an extended variable space,

$$\begin{aligned}
\mathbb{Q}[c_t^\top x] = \min_{q_t, y_t^s} \quad & q_t && \text{(2a)}\\
\text{s.t.} \quad & q_t \geq (c_t^s)^\top x + M_t^s (y_t^s - 1), \quad s \in \mathcal{S}_t, && \text{(2b)}\\
& \sum_{s\in\mathcal{S}_t} y_t^s p_t^s \geq \tau, && \text{(2c)}\\
& y_t^s \in \{0,1\}, \quad s \in \mathcal{S}_t, && \text{(2d)}
\end{aligned}$$

where M_t^s are sufficiently large numbers. We will discuss specific choices of these parameters later when we consider concrete examples. Using such a technique leads to the reformulation

$$\begin{aligned}
\min_{x,z} \quad & \alpha \sum_{t\in\mathcal{T}} \mathbb{E}[c_t^\top x] + (1-\alpha) \sum_{t\in\mathcal{T}} g(z_t) && \text{(3a)}\\
\text{s.t.} \quad & x \in X, \quad z_t \in Z_t(x),\ t \in \mathcal{T}, && \text{(3b)}
\end{aligned}$$

where the (possibly mixed-integer) constraint sets $Z_t(x)$, which are required to model the quantile, depend on the original variables x. Moreover, we have $z = (z_t)_{t\in\mathcal{T}}$ and g is an arbitrary function depending on the newly introduced variable vector z.

To highlight the generality of this class of optimization models, we now consider two examples in the following subsections.

2.1. **The Maintenance Planning Problem.** Let $\mathcal{I}$ denote the set of interventions to be scheduled, let $\mathcal{T} = \{1, \dots, T\}$ be the set of time indices representing the time horizon, i.e., the set of time steps at which interventions can take place, and let $\mathcal{R}$ be the set of resources used for the interventions. Further, for each time $t \in \mathcal{T}$, $\mathcal{S}_t$ represents the set of scenarios at this time, which all have the same probability, i.e., $p_t^s = p_t = 1/|\mathcal{S}_t|$ holds for all $t \in \mathcal{T}$. The duration of intervention $i \in \mathcal{I}$, if it starts at time $t \in \mathcal{T}$, is given by Δ_t^i. The amount of resource r used at time t by intervention i starting at time t' is given by $r_{t',t}^i$. The total amount of resource r used by all interventions in process at time t must be at least l_t^r and cannot be larger than u_t^r. Moreover, $\sigma_{t',t}^{i,s}$ denotes the cost in scenario s at time t of intervention i if it starts at time t'. We are also given a set $\mathcal{D}$ of triplets (i,j,t) such that intervention i and j cannot be both in process at time t.

Intervention preemption is not allowed and each intervention must be terminated at time T. We thus denote by $\mathcal{T}(i) = \{t \in \mathcal{T} \colon t + \Delta_t^i \leq T\}$ the set of feasible starting times of intervention i. Further, we denote by $\mathcal{T}(i,t) = \{t' \in \mathcal{T} \colon t' \leq t,\ t' + \Delta_{t'}^i \geq t\}$ the set of starting times of intervention i for which the intervention is in process at time t.

To model the maintenance planning problem (MPP), we use a set of binary variables x: for $i \in \mathcal{I}$ and $t \in \mathcal{T}(i)$, we have $x_t^i = 1$ if intervention i starts at time t. Further, for $t \in \mathcal{T}$, q_t and ε_t are continuous variables. First, q_t represents the τ-quantile. Second, ε_t models the maximum of zero and the difference between the τ-quantile and the average of the cost at time t for the different scenarios in $\mathcal{S}_t$. We now present the MILP:

$$\min_{x,\varepsilon,q} \quad \alpha \frac{1}{T} \sum_{t\in\mathcal{T}} \frac{1}{|\mathcal{S}_t|} \sum_{s\in\mathcal{S}_t} \sum_{i\in\mathcal{I}} \sum_{t'\in\mathcal{T}(i,t)} \sigma_{t',t}^{i,s} x_{t'}^i + (1-\alpha)\frac{1}{T}\sum_{t\in\mathcal{T}} \varepsilon_t \tag{4a}$$

$$\text{s.t.} \quad \sum_{t\in\mathcal{T}(i)} x_t^i = 1, \quad i \in \mathcal{I}, \tag{4b}$$

$$l_t^r \le \sum_{i\in\mathcal{I}} \sum_{t'\in\mathcal{T}(i,t)} r_{t',t}^i x_{t'}^i \le u_t^r, \quad r \in \mathcal{R},\, t \in \mathcal{T}, \tag{4c}$$

$$\sum_{t'\in\mathcal{T}(i,t)} x_{t'}^i + \sum_{t'\in\mathcal{T}(j,t)} x_{t'}^j \le 1, \quad (i,j,t) \in \mathcal{D}, \tag{4d}$$

$$\varepsilon_t \ge q_t - \frac{1}{|\mathcal{S}_t|} \sum_{s\in\mathcal{S}_t} \sum_{i\in\mathcal{I}} \sum_{t'\in\mathcal{T}(i,t)} \sigma_{t',t}^{i,s} x_{t'}^i, \quad t \in \mathcal{T}, \tag{4e}$$

$$\varepsilon_t \ge 0, \quad t \in \mathcal{T}, \tag{4f}$$

$$x_t^i \in \{0,1\}, \quad i \in \mathcal{I},\, t \in \mathcal{T}, \tag{4g}$$

$$q_t \;=\; Z_t(x), \quad t \in \mathcal{T}, \tag{4h}$$

with $x = (x_t)_{t\in\mathcal{T}}$ and $x_t = (x_t^i)_{i\in\mathcal{I}}$. Analogous vector notation is used to define ε and q.

This problem is the one of the ROADEF/EURO challenge 2020. The first term of the objective function represents the average cost and the second term represents the average excess, i.e., the average of the maximum of zero and the difference between the cost's τ-quantile and the cost's average. They are weighted with coefficients α and $(1-\alpha)$, $\alpha \in [0,1]$, respectively. Constraints (4b) specify that each intervention must start in exactly one time period. Constraints (4c) indicate that the amount of each resource must be within its lower and upper bounds at each time period. Constraints (4d) forbid pairs of interventions to be in process at the same time when they are in conflict. Constraints (4e) together with the objective function define the excess, at each time t, as the difference between the τ-quantile and the average of the costs. Note that we used that all scenarios have the same probability in this setting while defining the quantile. Finally, (4f)–(4g) specify the type of the different variables.

The set $Z_t(x)$ appearing in Constraints (4h) is the set of optimal solutions of the following problem that states that the quantile denoted by q_t must be the smallest value larger than or equal to at least $p_t = \lceil \tau|\mathcal{S}_t| \rceil$ cost values $\sum_{i\in\mathcal{I}} \sum_{t'\in\mathcal{T}(i,t)} \sigma_{t',t}^{i,s} x_{t'}^i$. To this end, it uses the variables y_t^s that take binary values. Thus, we have

$$Z_t(x) = \min_{q_t, y_t^s} \quad q_t \tag{5a}$$

$$\text{s.t.} \quad q_t \ge \sum_{i\in\mathcal{I}} \sum_{t'\in\mathcal{T}(i,t)} \sigma_{t',t}^{i,s} x_{t'}^i + M_t^s (y_t^s - 1), \quad s \in \mathcal{S}_t, \tag{5b}$$

$$\sum_{s\in\mathcal{S}_t} y_t^s p_t \ge \tau, \tag{5c}$$

$$y_t^s \in \{0,1\}, \quad s \in \mathcal{S}_t. \tag{5d}$$

Given that the objective function (4a) to be minimized is non-decreasing in the quantiles q_t, we can replace Constraints (4h) by Constraints (5b)–(5d). Note further

that Problem (5) also ensures that every quantile is non-negative if all costs are non-negative.

Constraints (5b) involve the big-M constants M_t^s that must be an upper bound on $\sum_{i\in\mathcal{I}}\sum_{t'\in\mathcal{T}(i,t)}\sigma_{t',t}^{i,s}x_{t'}^i$. Given that variables x_t^i satisfy (4b), we can choose

$$M_t^s = \sum_{i\in\mathcal{I}} \max_{t'\in\mathcal{T}(i,t)} \sigma_{t',t}^{i,s}. \tag{6}$$

A stronger big-M can be computed via

$$M_t^s = \max \quad \sum_{i\in\mathcal{I}}\sum_{t'\in\mathcal{T}(i,t)} \sigma_{t',t}^{i,s}x_{t'}^i \tag{7a}$$

$$\text{s.t.} \quad (4b), (4c), (4d) \quad \text{and} \quad x_t^i \in \{0,1\},\ i\in\mathcal{I},\ t'\in\mathcal{T}(i,t). \tag{7b}$$

A compromise consists in solving the LP relaxation of (7). Later in our numerical experiments, we will use (6).

Finally, to show that Problem (4) is a special instance of the general problem (1), we first note that for every scenario $s\in\mathcal{S}_t$, the random variable in the maintenance planning problem is given by

$$(c_t^s)^\top x = \sum_{i\in\mathcal{I}}\sum_{t'\in\mathcal{T}(i,t)} \sigma_{t',t}^{i,s}x_{t'}^i, \quad s\in\mathcal{S}_t,$$

with associated probability $p_t^s = 1/|\mathcal{S}_t|$. Then, we replace the excess ε_t in the second term of the objective function with

$$f(\mathbb{Q}[c_t^\top x]) = \max\left\{0,\ \mathbb{Q}[c_t^\top x] - \frac{1}{|\mathcal{S}_t|}\sum_{s\in\mathcal{S}_t}\sum_{i\in\mathcal{I}}\sum_{t'\in\mathcal{T}(i,t)}\sigma_{t',t}^{i,s}x_{t'}^i\right\}.$$

Thus, the variables ε_t, $t\in\mathcal{T}$, and Constraints (4e), (4f), and (4h) are not required anymore.

By putting this all together and by re-scaling the objective function with T, for X defined by (4b)–(4d) and (4g), we obtain

$$\min \quad \alpha\sum_{t\in\mathcal{T}}\mathbb{E}[c_t^\top x] + (1-\alpha)\sum_{t\in\mathcal{T}} f(\mathbb{Q}[c_t^\top x]) \quad \text{s.t.} \quad x=(x_t)_{t\in\mathcal{T}}\in X,$$

which is exactly of the general form (1).

2.2. **Portfolio Optimization.** In the classic portfolio optimization problem [20] we are given a budget B that we need to invest in a set of n equities so that the expected return is maximized while the corresponding risk, calculated by the standard deviation, is limited to be at most of a given value.

In recent years the risk measure that has mostly been used in the financial community is the Value-at-Risk (VaR) [4]: to minimize the portfolio's risk, one wishes that, for a given value of the parameter τ, that the corresponding quantile (or VaR) is large.

According to [2], the VaR at the $100\tau\,\%$ confidence level of a risky portfolio is the rate of return v^q such that $F(-v^q) = 1-\tau$ and $F(\cdot)$ is the cumulative distribution function of the portfolio's rate of return at the end of the period.

Let r_i be the return of equity i after, e.g., one year, which is a random variable. Moreover, let $x\in\mathbb{R}^n$ be the vector describing the investment. This means that we invest x_iB in equity i and it holds

$$\sum_{i=1}^n x_i = 1, \quad x\geq 0.$$

The return of the entire portfolio is then given by $r^\top x$.

The resulting general model for portfolio optimization reads

$$\max_x \quad \alpha\mathbb{E}[r^\top x] - (1-\alpha)\mathbb{Q}[-r^\top x] \tag{8a}$$

$$\text{s.t.} \quad x \in X = \left\{ x \in \mathbb{R}^n \colon \sum_{i=1}^n x_i = 1,\ \mathbb{E}[r^\top x] \geq \rho,\ x \geq 0 \right\} \tag{8b}$$

with $\alpha \in [0,1]$, ρ is the minimum expected return of the portfolio and $\mathbb{Q}[-r^\top x]$ is the $(1-\tau)$-quantile of $-r^\top x$. This model is a particular case of Model (1) with $\mathcal{T}$ being a singleton and $c = -r$. Indeed,

$$\max_x \quad \alpha\mathbb{E}[r^\top x] - (1-\alpha)\mathbb{Q}[-r^\top x]$$

is obviously equivalent to

$$-\min_x \quad \alpha\mathbb{E}[-r^\top x] + (1-\alpha)\mathbb{Q}[-r^\top x].$$

For $\alpha = 0$ this problem amounts to minimize the VaR with a minimum expected return. This is the model proposed in [9] and we can deduct from [4] that it is strongly NP-hard. However, (8) also encompasses the possibility of a linear combination of both objectives or to relax the constraint of a minimal expected return by choosing a value for ρ that is sufficiently small.

A common approach to determine an optimal portfolio consists in using historical or simulated data. In this context, return vectors r^s with their associated probability p^s are given for a set $\mathcal{S}$ of scenarios so that

$$\mathbb{E}[r^\top x] = \sum_{s\in\mathcal{S}} p^s (r^s)^\top x$$

and

$$\mathbb{Q}[-r^\top x] = \min\left\{ -q \colon \sum_{s\in\mathcal{N}(-q,-r^\top x)} p^s \geq 1-\tau \right\},$$
$$\mathcal{N}(-q,-r^\top x) = \left\{ s \in \mathcal{S} \colon -(r^s)^\top x \leq -q \right\}$$

holds. Thus, we obtain

$$\mathbb{Q}[-r^\top x] = -\max\left\{ q \colon \sum_{s\in\mathcal{N}(-q,-r^\top x)} p^s \geq 1-\tau \right\},$$
$$\mathcal{N}(-q,-r^\top x) = \left\{ s \in \mathcal{S} \colon (r^s)^\top x \geq q \right\}.$$

Finally, the portfolio optimization problem (POP) can then be formulated as the MILP

$$\max_{x,q,y} \quad \alpha \sum_{s\in\mathcal{S}} p^s (r^s)^\top x + (1-\alpha) q \tag{9a}$$

$$\text{s.t.} \quad x \in X, \tag{9b}$$

$$q \leq (r^s)^\top x + M^s y^s, \quad s \in \mathcal{S}, \tag{9c}$$

$$\sum_{s\in\mathcal{S}} p^s y^s \leq \tau, \tag{9d}$$

$$y^s \in \{0,1\}, \quad s \in \mathcal{S}, \tag{9e}$$

where M^s is a sufficiently large constant that can be set equal to

$$\min\left\{ q \colon \sum_{s'\in\mathcal{M}(q)} p^{s'} \geq \tau \right\} - \min_i r_i^s \quad \text{with} \quad \mathcal{M}(q) = \left\{ s' \in \mathcal{S} \colon \max_i r_i^{s'} \leq q \right\}.$$

Again, this is a special case of Model (3) in which $z = (q, y)$ as well as $g(z) = q$ holds and $Z(x)$ is defined by Constraints (9c)–(9e).

3. Valid Inequalities

For the ease of notation, we omit the index t in this section when there is no possible ambiguity. Lower bounds on the variable q representing the τ-quantile can be obtained in two different ways. The first one uses the strong-duality property of linear optimization while the second one is based on a combinatorial argument.

In what follows, we set $p(\bar{\mathcal{S}}) = \sum_{s\in\bar{\mathcal{S}}} p^s$ and $c_i(\bar{\mathcal{S}}) = \sum_{s\in\bar{\mathcal{S}}} p^s c_i^s$ for $\bar{\mathcal{S}} \subseteq \mathcal{S}$.

Proposition 1. *The following inequality is valid for the quantile problem (2) for all subsets $\bar{\mathcal{S}} \subseteq \mathcal{S}$ with $p(\bar{\mathcal{S}}) < \tau$:*

$$\big(\tau - p(\bar{\mathcal{S}})\big)\, q \geq \sum_{i=1}^{n} \big(b_i - c_i(\bar{\mathcal{S}})\big)\, x_i \tag{10}$$

with

$$b_i = \min_{w} \left\{ \sum_{s\in\mathcal{S}} c_i^s w^s \colon \sum_{s\in\mathcal{S}} w^s = \tau,\ 0 \leq w^s \leq p^s,\ s \in \mathcal{S} \right\}.$$

Furthermore, it can be separated in polynomial time.

Proof. Let us consider $d^s := (c^s)^\top x$ and p^s, $s \in \mathcal{S}$, as the realizations of a discrete random variable and its corresponding probability. It is well known, see, e.g., [19], that the τ-quantile q is the optimal solution of the linear optimization problem

$$\max_{u,q} \quad \tau q - \sum_{s\in\mathcal{S}} p^s u^s \tag{11a}$$

$$\text{s.t.} \quad q - u^s \leq d^s, \quad u^s \geq 0, \quad s \in \mathcal{S}, \tag{11b}$$

$$q \in \mathbb{R}. \tag{11c}$$

The dual of this problem reads

$$\min_{w} \quad \sum_{s\in\mathcal{S}} d^s w^s \tag{12a}$$

$$\text{s.t.} \quad \sum_{s\in\mathcal{S}} w^s = \tau, \tag{12b}$$

$$0 \leq w^s \leq p^s, \quad s \in \mathcal{S}. \tag{12c}$$

Using strong duality of linear optimization, the quantile q must satisfy

$$\tau q \geq \sum_{s\in\mathcal{S}} d^s w^s + \sum_{s\in\mathcal{S}} p^s u^s.$$

Since $u^s \geq \max\{q - d^s, 0\}$, the following inequality is valid for all $\bar{\mathcal{S}} \subseteq \mathcal{S}$:

$$\tau q \geq \sum_{s\in\mathcal{S}} d^s w^s + \sum_{s\in\bar{\mathcal{S}}} p^s (q - d^s). \tag{13}$$

In our general framework, the realization d^s is a linear function $\sum_{i=1}^n c_i^s x_i$ so that the resulting inequality (13) is nonlinear. However, the linear inequality (10) can be obtained by noticing that

$$\sum_{s\in\mathcal{S}} d^s w^s = \sum_{s\in\mathcal{S}} \sum_{i=1}^{n} c_i^s x_i w^s \geq \sum_{i=1}^{n} b_i x_i, \quad b_i = \min_{w} \left\{ \sum_{s\in\mathcal{S}} c_i^s w^s \colon (12b), (12c) \right\} \tag{14}$$

and by rearranging terms.

The separation problem for Inequality (10) is easy. Given a solution $\bar{x}, \bar{q}$, it suffices to choose $\bar{\mathcal{S}} = \{s \in \mathcal{S} \colon \bar{q} > \sum_{i=1}^n c_i^s \bar{x}_i\}$ and check whether the resulting inequality (10) is violated. □

Note that these inequalities can be seen as a special case of the valid inequalities discussed in [15] for bilevel optimization.

The second approach to derive a lower bound on the τ-quantile uses a covering argument and can be seen as a generalization of the idea proposed by [24] for the case where at most k linear inequalities among n given ones are allowed to be violated. In our context, if for a subset $\bar{\mathcal{S}}$ of scenarios the probability satisfies $p(\bar{\mathcal{S}}) < \tau$, then $q \geq (c^s)^\top x$ holds for some scenarios $s \in \mathcal{S} \setminus \bar{\mathcal{S}}$.

Proposition 2. *The following inequality is valid for the quantile problem (2) for all subsets $\bar{\mathcal{S}} \subseteq \mathcal{S}$ with $p(\bar{\mathcal{S}}) < \tau$:*

$$\left(\tau - p(\bar{\mathcal{S}})\right) q \geq \sum_{i=1}^{n} b_i(\bar{\mathcal{S}}) x_i \tag{15}$$

where

$$b_i(\bar{\mathcal{S}}) = \min_{w} \left\{ \sum_{s \in \mathcal{S} \setminus \bar{\mathcal{S}}} c_i^s w^s : \sum_{s \in \mathcal{S} \setminus \bar{\mathcal{S}}} w^s = \tau - p(\bar{\mathcal{S}}), 0 \leq w^s \leq p^s, s \in \mathcal{S} \setminus \bar{\mathcal{S}} \right\}.$$

Proof. Since $p(\bar{\mathcal{S}}) < \tau$ holds, there exists a subset $\mathcal{S}^c \subset \mathcal{S} \setminus \bar{\mathcal{S}}$ such that $p(\mathcal{S}^c) > \tau - p(\bar{\mathcal{S}})$ and $q \geq (c^s)^\top x$ holds for all $s \in \mathcal{S}^c$. Taking a weighted sum of these inequalities with coefficients v^s for $s \in \mathcal{S}^c$ with $\sum_{s \in \mathcal{S}^c} v^s = \tau - p(\bar{\mathcal{S}})$ and $0 \leq v^s \leq p^s$ yields

$$\left(\tau - p(\bar{\mathcal{S}})\right) q \geq \sum_{i=1}^{n} \sum_{s \in \mathcal{S}^c} v^s c_i^s x_i. \tag{16}$$

Inequality (15) is obtained by additionally using that $\sum_{s \in \mathcal{S}^c} v^s c_i^s \geq b_i(\bar{\mathcal{S}})$ holds. □

Let us remark that the definitions of b_i and $b_i(\bar{\mathcal{S}})$ imply that $b_i = b_i(\emptyset)$. To avoid any ambiguities in the sequel we will rather use the notation $b_i(\emptyset)$.

The following example shows that a priori there is no dominance relation between the inequalities in (15).

Example 1. *Consider four scenarios $s \in \{1, 2, 3, 4\}$ with equal probability $1/4$ and four variables x_i with $i \in \{1, 2, 3, 4\}$ and let $c_i^s = 0$, if $i = s$ and $c_i^s = 1$ otherwise. If $\tau = 3/4$, then the inequalities in (15) are the following (after rescaling):*

- $q \geq 2/3 \sum_{i=1}^{4} x_i$, *for* $\bar{\mathcal{S}} = \emptyset$
- $q \geq \sum_{i \in \bar{\mathcal{S}}} x_i + 1/2 \sum_{i \notin \bar{\mathcal{S}}} x_i$, *if* $|\bar{\mathcal{S}}| = 1$
- $q \geq \sum_{i \in \bar{\mathcal{S}}} x_i$, *if* $|\bar{\mathcal{S}}| = 2$.

None of them is dominated by a nonnegative linear combination of the others.

As shown in the following proposition, the separation of the inequalities in (15) is difficult at least for a fixed value of $p(\bar{\mathcal{S}})$.

Proposition 3. *For a fixed value of $p(\bar{\mathcal{S}})$, the separation problem for the inequalities in (15) is NP-hard even in the special case where $p^s = 1/|\mathcal{S}|$ for all $s \in \mathcal{S}$, $\tau = k/|\mathcal{S}|$, and $c_i^s \in \{0, 1\}$.*

Proof. Under the above conditions, the decision version (D-SEP) of the separation problem of (15) for a point (x^*, q^*) consists in determining whether there exists a subset $\bar{\mathcal{S}}$ such that $|\bar{\mathcal{S}}| = B$ and

$$|\mathcal{S}| \sum_{i=1}^{n} b_i(\bar{\mathcal{S}}) x_i^* > q^*(k - B) \tag{17}$$

holds.

This problem clearly belongs to NP. Further, we show that CLIQUE reduces to it; see, e.g., Problem GT19 in [10]. To this end, for an instance of CLIQUE given by a graph $G = (V, E)$ and an integer B, we define an instance of (D-SEP) as follows. We set $\mathcal{S} = V$, $\mathcal{I} = E$, $k = B + 1$, $x_i^* = \chi$ for all i, $q^* = \chi(B(B-1)/2 - 1)$, and $c_i^s = 0$ if edge i is incident to vertex s and $c_i^s = 1$ otherwise. Then, $|\mathcal{S}|b_i(\bar{\mathcal{S}}) = \min\left\{c_i^s \colon s \in \mathcal{S} \setminus \bar{\mathcal{S}}\right\} = 1$ if both end vertices of i belong to $\bar{\mathcal{S}}$ and 0 otherwise. Hence, (17) reads

$$\chi|\mathcal{S}|\sum_{i=1}^{n} b_i(\bar{\mathcal{S}}) = \chi|E(\bar{\mathcal{S}})| > q^*(k-B) = \chi(B(B-1)/2-1)$$

and is satisfied if and only if $\bar{\mathcal{S}}$ is a clique of size B. □

The following proposition shows that the inequalities in (15) are stronger than the inequalities in (10).

Proposition 4. *For $\bar{\mathcal{S}} \in \mathcal{S}$, Inequality (15) dominates Inequality (10).*

Proof. The left-hand sides of both inequalities are equal. Further, we have

$$b_i(\emptyset) = \min_w \left\{ \sum_{s\in\mathcal{S}} c_i^s w^s \colon \sum_{s\in S} w^s = \tau, 0 \le w^s \le p^s, s \in S \right\} \le c_i(\bar{\mathcal{S}}) + b_i(\bar{\mathcal{S}})$$

Hence, the ith coefficient of (15) is larger than or equal to the corresponding one of (10). □

The above results suggest to use the separation procedure for Inequality (10) but to add the corresponding stronger inequality (15).

3.1. **Application to the Maintenance Planning Problem.** First recall that, in the MPP, we have a set of scenarios $\mathcal{S}_t$ for each time step t and $p^s = 1/|\mathcal{S}_t|$ holds for all $s \in \mathcal{S}_t$.

The following proposition shows how to adapt the valid inequalities (10) and (15) to MPP.

Proposition 5. *The following two inequalities are valid for the MPP (4) for all subsets $\bar{\mathcal{S}} \subseteq \mathcal{S}_t$ with $|\bar{\mathcal{S}}| < \lceil\tau|\mathcal{S}_t|\rceil$:*

$$\left(\lceil\tau|\mathcal{S}_t|\rceil - |\bar{\mathcal{S}}|\right) q_t \ge \sum_{i\in I} \sum_{t'\in\mathcal{T}(i,t)} \left(b_{t,t'}^i(\emptyset) - \sum_{s\in\bar{\mathcal{S}}} \sigma_{t',t}^{i,s} x_{t'}^i \right) \tag{18}$$

and

$$\left(\lceil\tau|\mathcal{S}_t|\rceil - |\bar{\mathcal{S}}|\right) q_t \ge \sum_{i\in I} \sum_{t'\in\mathcal{T}(i,t)} b_{t,t'}^i(\bar{\mathcal{S}}) x_{t'}^i, \tag{19}$$

where

$$b_{t,t'}^i(\bar{\mathcal{S}}) = \min_w \left\{ \sum_{s\in\mathcal{S}_t\setminus\bar{\mathcal{S}}} \sigma_{t',t}^{i,s} w^s \colon \sum_{s\in\mathcal{S}_t\setminus\bar{\mathcal{S}}} w^s = \lceil\tau|\mathcal{S}_t|\rceil - |\bar{\mathcal{S}}|, 0 \le w^s \le 1, s \in \mathcal{S}_t \setminus \bar{\mathcal{S}} \right\}$$

holds. In addition, Inequality (18) can be separated in polynomial time but is dominated by Inequality (19), whose separation is NP-hard.

Proof. The objective function of Problem (11) can be rewritten as

$$\max_{u,q} \quad \lceil\tau|\mathcal{S}|\rceil q - \sum_{s\in\mathcal{S}} u^s, \tag{20}$$

which leads to the dual formulation

$$\min_{w} \quad \sum_{s\in\mathcal{S}} d^s w^s \tag{21a}$$

$$\text{s.t.} \quad \sum_{s\in\mathcal{S}} w^s = \lceil \tau|\mathcal{S}| \rceil, \tag{21b}$$

$$0 \le w^s \le 1, \quad s \in \mathcal{S}. \tag{21c}$$

Note that in the MPP, we need to select $\lceil \tau|\mathcal{S}| \rceil$ scenarios with a cost lower than or equal to the quantile, which allows us to use the constant in the right-hand side of (21b) and in the objective function (20). Next, by strong duality, the quantile must then satisfy

$$\lceil \tau|\mathcal{S}| \rceil q \ge \sum_{s\in\mathcal{S}} d^s w^s + \sum_{s\in\bar{\mathcal{S}}} (q - d^s). \tag{22}$$

After re-introducing the t-index and bounding the nonlinear terms as in (14), Inequality (22) applied to the MPP reads

$$\lceil \tau|\mathcal{S}_t| \rceil q_t \ge \sum_{i\in I} \sum_{t'\in\mathcal{T}(i,t)} b^i_{t,t'}(\emptyset) x^i_{t'} + \sum_{s\in\bar{\mathcal{S}}} \left(q_t - \sum_{i\in\mathcal{I}} \sum_{t'\in\mathcal{T}(i,t)} \sigma^{i,s}_{t',t} x^i_{t'} \right), \quad t \in \mathcal{T}, \tag{23}$$

with

$$b^i_{t,t'}(\emptyset) = \min_{w} \left\{ \sum_{s\in\mathcal{S}_t} \sigma^{i,s}_{t',t} w^s : \sum_{s\in\mathcal{S}_t} w^s = \lceil \tau|\mathcal{S}_t| \rceil,\ 0 \le w^s \le 1, s \in \mathcal{S}_t \right\}.$$

The coefficient $b^i_{t,t'}(\emptyset)$ is the sum of the $\lceil \tau|\mathcal{S}_t| \rceil$ smallest cost values $\sigma^{i,s}_{t',t}$ of intervention $i \in \mathcal{I}$ at time $t \in \mathcal{T}$ that has started at $t' \in \mathcal{T}(i,t)$.

By grouping the terms in q in the left-hand side, Inequality (23) becomes (18) and its separation can be done in polynomial time since it suffices to include in $\bar{\mathcal{S}}$ each scenario $s \in \mathcal{S}_t$ if the value of the second term of the right-hand side of (23) is positive for the current solution.

The proofs of the validity of Inequality (18), of the NP-hardness of its separation, and the fact that it dominates Inequality (18) are similar to those of Propositions 2–4 while taking again into account that the scenarios of a set $\mathcal{S}_t$ have equal probability. □

3.2. **Application to the Portfolio Optimization Problem.** Note that in the case of the portfolio optimization problem, given that $-q$ is the $(1-\tau)$-quantile of the linear functions $-(r^s)^\top x$, we can directly apply Propositions 2–4 with the modified data $\tilde{c}^s = -r^s$, $\tilde{\tau} = 1 - \tau$, and variables $\tilde{q} = -q$. The following proposition summarizes these results.

Proposition 6. *The following two inequalities are valid for the portfolio optimization problem (9) for all subsets $\bar{\mathcal{S}} \subseteq \mathcal{S}$ with $p(\bar{\mathcal{S}}) < 1 - \tau$:*

$$(1 - \tau - p(\bar{\mathcal{S}}))q \le \sum_{i\in I} (b_i(\emptyset) - \sum_{s\in\bar{\mathcal{S}}} p^s r^s_i) x_i, \tag{24}$$

and

$$(1 - \tau - p(\bar{\mathcal{S}}))q \le \sum_{i\in I} b_i(\bar{\mathcal{S}}) x_i, \tag{25}$$

where

$$b_i(\bar{\mathcal{S}}) = \max_{w} \left\{ \sum_{s\in\mathcal{S}\setminus\bar{\mathcal{S}}} r^s_i w^s : \sum_{s\in\mathcal{S}\setminus\bar{\mathcal{S}}} w^s = 1 - \tau - p(\bar{\mathcal{S}}), 0 \le w^s \le p^s, s \in \mathcal{S} \setminus \bar{\mathcal{S}} \right\}$$

holds. In addition, Inequality (24) *can be separated in polynomial time but is dominated by Inequality* (25)*, whose separation is NP-hard.*

4. An Overlapping Alternating Direction Method

In this section, we describe an overlapping alternating direction method to quickly compute feasible points of Problem (3). Alternating Direction Methods (ADMs) have been initially proposed in [7, 14] as extensions of Lagrangian methods. They are iterative procedures typically used to tackle problems defined by means of two vectors of decision variables, which are subject to some coupling constraints. Instead of solving the monolithic original problem, at each iteration of an ADM, one sequentially solves two smaller subproblems each of which determines a new value for one of the variable vectors, having fixed the value of the other one. In recent years, ADMs have been exploited to solve large-scale optimization problems in the field of gas transport [12, 13], machine learning [5, 18], bilevel problems [16], or supply chain problems [25]. In our work, we devise an overlapping ADM (OADM), which can be seen as a variant of an ADM to solve problems for which the vector of variables is partitioned into three subvectors. As in usual ADMs, two subproblems related to two variable subvectors are identified and solved sequentially. However, the remaining variable subvector is to be determined in both subproblems, since it is part of both. The idea behind OADMs can be traced back to overlapping Schwarz methods (see, e.g., [6, 21]) used in the field of partial differential equations to solve boundary value problems defined on a domain that is a union of some intersecting subdomains. Recently, it also has been applied very successfully to graph-structured problems; see, e.g., [26].

In the following, we present an OADM for Problem (3), which is motivated by the special structure of the problem itself and by our formulation for the quantile. We observe that Problem (3) makes use of the pair (x, z) of variable vectors, where z is introduced only to reformulate the term of the objective function (3a) that is related to the quantile. Specifically, $z \in Z_t(x)$, $t \in \mathcal{T}$, encodes the formulation of the quantile. In both of our applications, sets of constraints $Z_t(x)$, $t \in \mathcal{T}$, are defined over two vectors of variables; see Problem (5) and (9). Hence, in this section, we generalize the setting of Problem (3) by considering it defined over the 3-tuple (x, z_1, z_2) of variable vectors, where z_1 and z_2 are two subvectors of z, i.e., $z = (z_1, z_2)$. In what follows, we may still write z in lieu of (z_1, z_2) if the explicit decomposition is not required. In our OADM, we identify a subproblem related to x and one related to z_1. The variable vector z_2 is overlapping, i.e., it is part of the optimization for both subproblems. Finally, we highlight that our OADM enjoys two special features. The value of z depends only on the one of x (see, e.g., Constraints (3b)) and z has no influence on the feasible set of the problem, i.e., at each iteration, solving the two subproblems provides feasible points for Problem (3).

We apply the OADM outlined in Algorithm 1 to determine a feasible point of Problem (3), which improves on an initial one (x^0, z_1^0, z_2^0) in terms of the objective function value. The initial point (x^0, z_1^0, z_2^0) is defined by

$$x^0 \in \arg\min_x \left\{ \alpha \sum_{t \in \mathcal{T}} \mathbb{E}[c_t^\top x] \colon x \in X \right\},$$
$$z^0 = (z_1^0, z_2^0) \in \left\{ z = (z_t)_{t \in \mathcal{T}} \colon z_t = (z_{1,t}, z_{2,t}) \in Z_t(x^0) \right\}.$$

Specifically, x^0 is chosen among the solutions of the problem obtained by Problem (3) after removing the constraints and the objective function term involving the variable vector z. Clearly, (x^0, z^0) is a feasible point of Problem (3).

Now, we describe the iterative procedure outlined in Algorithm 1. In what follows, we write

$$v(x,z) := \alpha \sum_{t\in\mathcal{T}} \mathbb{E}[c_t^\top x] + (1-\alpha)\sum_{t\in\mathcal{T}} g(z_t)$$

to lighten the notation. In Line 1, we set the iteration counter j to zero. In iteration $j+1$, the algorithm first solves Problem (3) in the direction of (x, z_2) while having fixed the value of z_1 to z_1^j to determine a new value (x^{j+1}, z_2^{j+1}) for (x, z_2); see Line 3. Then, Problem (3) is solved in the direction of (z_1, z_2) while having fixed the value of x to x^{j+1} to determine a new value (z_1^{j+1}, z_2^{j+1}) for (z_1, z_2). Finally, the algorithm stops (Line 7) once a given stopping criterion is met such as that a time limit is reached or that the improvement of the value of the feasible points is less than a given threshold.

Algorithm 1: An Overlapping Alternating Direction Method

Input : An initial feasible point $(x^0, z^0 = (z_1^0, z_2^0))$ of Problem (3).
Output: A feasible point $(x^j, z^j = (z_1^j, z_2^j))$ of Problem (3).

1 Set $j \leftarrow 0$.
2 **while** *stopping criterion is not satisfied* **do**
3 Compute $(x^{j+1}, z_2^{j+1}) \in \arg\min_{x,z_2}\{v(x, z_1^j, z_{2,t}) \colon x \in X, (z_{1,t}^j, z_2) \in Z_t(x), t \in \mathcal{T}\}$.
4 Compute $z^{j+1} = (z_1^{j+1}, z_2^{j+1}) \in \arg\min_{z=(z_1,z_2)}\{v(x^{j+1}, z) \colon z_t \in Z_t(x^{j+1}), t \in \mathcal{T}\}$.
5 Increment $j \leftarrow j+1$.
6 **end**
7 **return** (x^j, z^j)

4.1. **Application to the Maintenance Planning Problem.** In this section, we discuss how Algorithm 1 is applied to the MPP. Specifically, we apply our OADM to the variant of Problem (4) that makes use of Constraints (5). This problem is defined on four vectors of variables: $x = (x_t)_{t\in\mathcal{T}}$ with $x_t = (x_t^i)_{i\in\mathcal{I}}$, $y = (y_t)_{t\in\mathcal{T}}$ with $y_t = (y_t^s)_{s\in\mathcal{S}_t}$, $q = (q_t)_{t\in\mathcal{T}}$, and $\varepsilon = (\varepsilon_t)_{t\in\mathcal{T}}$. In the OADM for the MPP, variable vectors x and y play the role of x and z_1 and variable vectors q and ε play the role of the overlapping variables z_2 in Algorithm 1. Hence, Problem (4) is solved in the direction of x and y and, in both directions, q and ε are also part of the optimization. Here, we consider the variable vector ε as part of z. This leads to a re-definition of the sets $Z_t(x)$, $t \in \mathcal{T}$, where Constraints (4e) and (4f) are included.

An initial point $(x^0, y^0, q^0, \varepsilon^0)$ of Problem (4) is retrieved as for the general case. First, we select a planning x^0 for the interventions among the feasible solutions of the problem obtained by Problem (4) without taking into account the quantile related variables, constraints, and the objective function term, i.e.,

$$x^0 \in \arg\min_x \left\{ \alpha \frac{1}{T} \sum_{t\in\mathcal{T}} \frac{1}{|\mathcal{S}_t|} \sum_{s\in\mathcal{S}_t} \sum_{i\in\mathcal{I}} \sum_{t'\in\mathcal{T}(i,t)} \sigma_{t',t}^{i,s} x_{t'}^i \colon (4b), (4c), (4d), (4g) \right\}.$$

Then, we determine values y^0, q^0, and ε^0 for variable vectors y, q, and ε as when solving the problem in the direction of y. We do so as follows.

In iteration $j+1$, the subproblem in the x-direction corresponds to finding a maintenance planning for the interventions of $\mathcal{I}$ in the time horizon $\mathcal{T}$ while having fixed the scenarios used to calculate the quantile at each time instant. Hence, the values x^{j+1} for x are selected among the feasible points of Problem (4) with the

values of variables in vector y fixed to y^j. We observe that this subproblem is NP-hard. Indeed, it can be reduced to the resource-constrained scheduling problem; see, e.g., [10]. Differently, solving the subproblem in the direction of y corresponds to computing the value of the objective function (4a) having fixed a planning for the interventions x^{j+1}. This can be done in a polynomial time. Specifically, given a time instant $t \in \mathcal{T}$, we first need to compute cost values σ_t^s related to each scenario $s \in \mathcal{S}_t$ as

$$\sigma_t^s = \sum_{i \in \mathcal{I}} \sum_{t' \in \mathcal{T}(i,t)} \sigma_{t',t}^{i,s} (x^{j+1})_{t'}^i.$$

Then, we sort the scenarios in $\mathcal{S}_t$ by non-decreasing values of σ_t^s. The first p_t scenarios appearing in this order are those for which we set $(y^{j+1})_t^s = 1$. For the others, we set $(y^{j+1})_t^s = 0$. This procedure is repeated for each time period $t \in \mathcal{T}$.

4.2. **Application to the Portfolio Optimization Problem.** The OADM for the portfolio optimization problem is applied to the MILP (9), which makes use of variable vectors $x \in X$, $y = (y^s)_{s \in \mathcal{S}}$, and a variable q. It sequentially solves a subproblem in the x-direction and one in the y-direction. The variable q is part of both optimization tasks. Solving the subproblem in the x-directions corresponds to maximizing the portfolio's revenue, having fixed the scenario selection. This problem is modeled as a linear program (see the MILP (9)) and can thus be efficiently solved. Solving the subproblem in the y-direction corresponds to computing the value of the objective function (9a) knowing the portfolio's composition. This can be done in polynomial time.

The initial point computation and the iterative procedure are analogous to the one discussed for the MPP. Hence, we do not report the details here.

5. An Adaptive Scenario Clustering Approach

This section presents an adaptive scenario clustering algorithm (ASCA) for solving Problem (1) in the case of the function f being non-decreasing. Given $t \in \mathcal{T}$, let us first denote with $\bar{c}_t$ the average cost at t, i.e.,

$$\bar{c}_t = \frac{1}{|\mathcal{S}_t|} \sum_{s \in \mathcal{S}_t} c_t^s,$$

which allows to rewrite the objective function (1a) as

$$\min \quad \alpha \sum_{t \in \mathcal{T}} \bar{c}_t x + (1 - \alpha) \sum_{t \in \mathcal{T}} f\left(\mathbb{Q}[c_t^\top x]\right). \tag{26}$$

Since a large number of scenarios can make the resolution of the problem computationally hard, a way to approximate the general problem is to reduce the size of each $\mathcal{S}_t$ by clustering its scenarios. This allows to heuristically find feasible solutions of good quality quickly.

More precisely, let $\mathcal{C}_t$ be a partition of $\mathcal{S}_t$ into $K_t \leq |\mathcal{S}_t|$ nonempty clusters. Each cluster $\gamma \in \mathcal{C}_t$ has a cost vector c_t^γ and a probability p_t^γ. The ASCA consists in solving a sequence of instances of Problem (1)—each defined over a clustered scenario set $\mathcal{C}_t$ instead of the original set $\mathcal{S}_t$. The probability p_t^γ of cluster $\gamma \in \mathcal{C}_t$ is given by

$$p_t^\gamma = \sum_{s \in \gamma} p_t^s,$$

which satisfies

$$\sum_{\gamma \in \mathcal{C}_t} p_t^\gamma = 1. \tag{27}$$

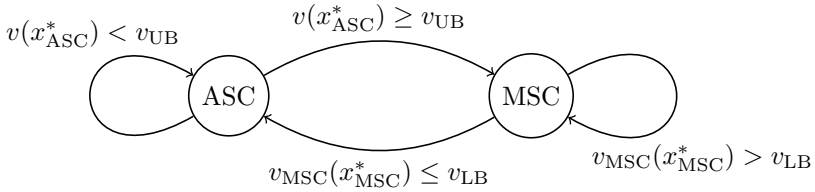

FIGURE 1. State diagram of the adaptive scenario clustering algorithm.

We now present two strategies to associate a cost vector to each cluster of scenarios. First, the average scenario clustering (ASC) associates with each cluster $\gamma \in \mathcal{C}_t$ a cost vector c_t^γ defined as follows:

$$(c_t^\gamma)_i = \frac{1}{|\gamma|} \sum_{s \in \gamma} (c_t^s)_i, \quad i \in [n],$$

where $[n] := \{1, \dots, n\}$. Let us indicate with $\mathcal{C}_t^{\text{ASC}}$ the corresponding clustering of scenarios.

Further, the minimum scenario clustering (MSC) associates with each cluster $\gamma \in \mathcal{C}_t$ a cost vector c_t^γ defined as

$$(c_t^\gamma)_i = \min \left\{ (c_t^s)_i \colon s \in \gamma \right\}, \quad i \in [n].$$

Let us indicate with $\mathcal{C}_t^{\text{MSC}}$ the corresponding clustering of scenarios.

Proposition 7. *Let x^* be an optimal solution of Problem (1) and let x^*_{ASC} as well as x^*_{MSC} denote optimal solutions of Problem (1) solved on the scenarios set $\mathcal{C}_t^{ASC}$ and $\mathcal{C}_t^{MSC}$, respectively. For a vector $x \in X$, let $v(x)$ and $v_{MSC}(x)$ denote the objective value w.r.t. x and (1a) defined over $\mathcal{S}_t$ and $\mathcal{C}_t^{MSC}$, respectively. Then,*

$$v_{MSC}(x^*_{MSC}) \leq v(x^*) \leq \min\{v(x^*_{ASC}), v(x^*_{MSC})\}$$

holds.

Proof. We get the first inequality by construction of the MSC, hence $v_{\text{MSC}}(x^*_{\text{MSC}}) \leq v_{\text{MSC}}(x^*) \leq v(x^*)$. The second inequality is due to x^* being optimal for Problem (1). □

Thus, both ASC and MSC allow to compute a bound on $v(x^*)$ while solving Problem (1) over clustered scenario sets. More specifically, solving the ASC corresponds to applying a heuristic tailored to finding good quality primal solutions for Problem (1). On the other side, the MSC will always yield a valid lower bound and may, depending on the considered instance and the underlying sets of clusters $\mathcal{C}_t$, also yield a good quality primal solution.

Corollary 1. *Let x^* be an optimal solution of Problem (1) and let x^*_{ASC} as well as x^*_{MSC} denote optimal solutions of Problem (1) solved for the scenarios set $\mathcal{C}_t^{ASC}$ and $\mathcal{C}_t^{MSC}$, respectively. For a vector $x \in X$, let $v(x)$ and $v_{MSC}(x)$ denote the objective value w.r.t. x and (1a) defined over $\mathcal{S}_t$ and $\mathcal{C}_t^{MSC}$, respectively. Then, if*

$$v_{MSC}(x^*_{MSC}) = \min\{v(x^*_{ASC}), v(x^*_{MSC})\}$$

holds, x^*_{ASC} or x^*_{MSC} is an optimal solution of Problem (1).

Additionally, solving Problem (1) with the MSC can be improved using the valid inequality (15).

Algorithm 2: Adaptive Scenario Clustering Algorithm.

Input : Threshold parameter $\varepsilon \in (0,1)$.

Set $\mathcal{C}_t^0 \leftarrow \{\mathcal{S}_t\}$ for all $t \in \mathcal{T}$, i.e., we start with a single cluster per index t.
Set $j \leftarrow 0$, $v_{\text{UB}} \leftarrow +\infty$, $v_{\text{LB}} \leftarrow -\infty$, and $\kappa \leftarrow$ true.
while $(v_{UB} - v_{LB})/v_{UB} \geq \varepsilon$ **do**
  if κ **then**
    Solve Problem (1) on $\mathcal{C}_t^{\text{ASC},j}$, let x^*_{ASC} denote the optimal solution.
    if $v(x^*_{ASC}) < v_{UB}$ **then**
      set $x_{\text{UB}} \leftarrow x^*_{\text{ASC}}$ and $v_{\text{UB}} \leftarrow v(x^*_{\text{ASC}})$
    else
      set $\kappa \leftarrow$ false.
    end
  else
    Solve Problem (1) on $\mathcal{C}_t^{\text{MSC}^+,j}$, let $x^*_{\text{MSC}^+}$ denote the optimal solution.
    if $v_{MSC^+}(x^*_{MSC^+}) > v_{LB}$ **then**
      set $x_{\text{LB}} \leftarrow x^*_{\text{MSC}^+}$ and $v_{\text{LB}} \leftarrow v_{\text{MSC}^+}(x^*_{\text{MSC}^+})$
    else
      set $\kappa \leftarrow$ true.
    end
    if $v(x^*_{MSC^+}) < v_{UB}$ **then**
      set $x_{\text{UB}} \leftarrow x^*_{\text{MSC}^+}$ and $v_{\text{UB}} \leftarrow v(x^*_{\text{MSC}^+})$.
    end
  end
  Update $j \leftarrow j+1$ and refine $\mathcal{C}_t^{j-1}$, yielding $\mathcal{C}_t^j$.
end
return x_{UB}

Proposition 8. *Given $\mathcal{C}_t^{MSC}$ for a specific $t \in \mathcal{T}$. Let γ be a cluster in $\mathcal{C}_t^{MSC}$ such that $p_t^\gamma > 1-\tau$ holds. Then, in Problem (2) defining the quantile of t, it holds that*

$$y_t^\gamma = 1,$$

and the resulting inequality (2b) is dominated by the valid inequality (15) for $\bar{\mathcal{S}} = \mathcal{S}_t \setminus \gamma$.

Proof. Since $p_t^\gamma > 1-\tau$, (2c) and (2d) imply $y_t^\gamma = 1$. Furthermore, the definition of $b_i(\bar{\mathcal{S}}_t)$ with $\bar{\mathcal{S}} = \mathcal{S}_t \setminus \gamma$ implies

$$\frac{b_i(\mathcal{S}_t \setminus \gamma)}{\tau - p(\mathcal{S}_t \setminus \gamma)} \geq (c_t^\gamma)_i, \quad i \in [n], \tag{28}$$

so that (15) dominates (2b). □

Proposition 8 hence tells us that given the set

$$\mathcal{B} := \{\gamma \in \mathcal{C}_t^{\text{MSC}} : p_t^\gamma > 1-\tau\},$$

we can replace the quantile constraint (2b) associated to $\gamma \in \mathcal{B}$ by the valid inequality (15) applied on the set $\mathcal{S}_t \setminus \gamma$ in the MSC. Furthermore, Constraint (2c) reduces to

$$\sum_{\gamma \in \mathcal{C}_t^{\text{MSC}}} y_t^\gamma p_t^\gamma \geq \tau - \sum_{\gamma \in \mathcal{B}} p_t^\gamma.$$

We denote by MSC^+ the resulting optimization problem.

Due to the inclusion of Inequality (15) derived for every element of $\mathcal{B}$, the set of feasible solutions of the MSC$^+$ problem is a subset of the MSC problem. As a consequence, we have that

$$v_{\text{MSC}}(x^*_{\text{MSC}}) \leq v_{\text{MSC}^+}(x^*_{\text{MSC}^+}) \leq v(x^*)$$

holds.

For what follows, we define v_{UB} and v_{LB} as the current best upper and lower bound on $v(x^*)$. Additionally, $\mathcal{C}_t^j$ stands for the clustering of $\mathcal{S}_t$ for $t \in \mathcal{T}$ in iteration j of the algorithm. We indicate with $\mathcal{C}_t^{\text{ASC},j}$ and $\mathcal{C}_t^{\text{MSC}^+,j}$ the instance of $\mathcal{C}_t^j$ where the cost vectors are calculated using ASC and MSC$^+$, respectively.

Algorithm 2 states the pseudo-code of our adaptive algorithm and Figure 1 shows a state diagram of the inner while loop of the algorithm. The algorithm solves a sequence of problems of Type (1) defined over $\mathcal{C}_t^{\text{ASC}}$ and $\mathcal{C}_t^{\text{MSC}^+}$ to improve the lower or upper bound on $v(x^*)$ as shown in Proposition 7. After each resolution, the algorithm refines the previous clustering $\mathcal{C}_t^{j-1}$, yielding $\mathcal{C}_t^j$ for all $t \in \mathcal{T}$. The resulting clustering $\mathcal{C}_t^j$ allows to have a better representation of the scenario set $\mathcal{S}_t$. This makes the next resolution of Problem (1) defined over $\mathcal{C}_t^{\text{ASC}}$ or $\mathcal{C}_t^{\text{MSC}^+}$ harder to solve. However, the new clustering of scenarios will likely allow to compute a better lower or upper bound on $v(x^*)$. The sequence of resolutions over $\mathcal{C}_t^{\text{ASC}}$ is then interrupted and switched to the resolution over $\mathcal{C}_t^{\text{MSC}^+}$ when $v(x^*_{\text{ASC}}) \geq v_{\text{UB}}$. Similarly, the sequence of resolution of Problem (1) over $\mathcal{C}_t^{\text{MSC}^+}$ is interrupted and switched to $\mathcal{C}_t^{\text{ASC}}$ when $v_{\text{MSC}^+}(x^*_{\text{MSC}^+}) \leq v_{\text{LB}}$. Finally, the algorithm terminates once we achieve a relative gap smaller than a prescribed tolerance ε, i.e.,

$$\frac{v_{\text{UB}} - v_{\text{LB}}}{v_{\text{UB}}} \leq \varepsilon.$$

Theorem 1. *Let x^* be an optimal solution of Problem (1) and let $v(x^*)$ be its value. Moreover, let $\mathcal{C}_t^j$ be the clustering of $\mathcal{S}_t$ for index $t \in \mathcal{T}$ in iteration j. Suppose further that there exists an index $t \in \mathcal{T}$ (that depends on the iteration j) such that*

$$|\mathcal{C}_t^j| > |\mathcal{C}_t^{j-1}|, \tag{29}$$

for all iterations j. Then, Algorithm 2 terminates after a finite number of cluster refinements with a point $x \in X$ such that

$$\frac{v(x) - v(x^*)}{v(x)} \leq \varepsilon. \tag{30}$$

Proof. Due to Inequality (29), each iteration of Algorithm 2 increases the size of $\mathcal{C}_t^j$ for at least one $t \in \mathcal{T}$. Therefore, if the termination criterion $(v_{\text{UB}} - v_{\text{LB}})/v_{\text{UB}} \leq \varepsilon$ is never satisfied, $\mathcal{C}_t^j$ will increase in size over the iterations until being equal to $\mathcal{S}_t$. If $\mathcal{C}_t^j$ equals $\mathcal{S}_t$ for all $t \in \mathcal{T}$, then $|\gamma| = 1$ for all $\gamma \in \mathcal{C}_t^j$.

The cost vectors c_t^γ for each $\gamma \in \mathcal{C}_t^j$ correspond to the cost vector associated with the single scenario in γ for both ASC and MSC clustering strategies. Otherwise, if $(v_{\text{UB}} - v_{\text{LB}})/v_{\text{UB}} \leq \varepsilon$ is satisfied, we know that x_{ASC} also satisfies Inequality (30) by Proposition 7. □

Remark 1. *We close this section with the discussion of two features of the ASCA.*

- *The cluster refinement step in Line 21 of Algorithm 2 is done using kernel density estimation (KDE). KDE allows to estimate the probability density function of a random variable by using a set of samples of this random variable [27]. The local minima of the estimated probability density function then yield a splitting of the random variable samples. In our case, for a given $t \in \mathcal{T}$ and $\gamma \in \mathcal{C}_t$ as well as a point $x \in X$, we compute the KDE of $(c_t^s)^\top x$ using all $s \in \gamma$, which results in a splitting of γ. We use x^*_{ASC} or x^*_{MSC}*

depending on whether the previous iteration $j-1$ of the ASCA used $\mathcal{C}_t^{ASC,j-1}$ or $\mathcal{C}_t^{MSC,j-1}$. Additionally, only a subset of time steps $\overline{\mathcal{T}} \subseteq \mathcal{T}$ is selected for re-clustering after each iteration of the ASCA. Given a parameter $\Theta \in (0,1)$, we compute this refinement set $\overline{\mathcal{T}}$ to be the minimal subset of $\mathcal{T}$ satisfying

$$\sum_{t\in\overline{\mathcal{T}}} \left|\mathbb{Q}[c_t^\top x; \mathcal{C}_t^{j-1}] - \mathbb{Q}[c_t^\top x]\right| > \Theta \sum_{t\in\mathcal{T}} \left|\mathbb{Q}[c_t^\top x; \mathcal{C}_t^{j-1}] - \mathbb{Q}[c_t^\top x]\right|.$$

Here, $\mathbb{Q}[\cdot]$ is the original quantile, whereas $\mathbb{Q}[\cdot; \mathcal{C}_t^{j-1}]$ denotes the quantile's approximation based on the clustering $\mathcal{C}_t^{j-1}$. Hence, we select the subset of indices $t \in \mathcal{T}$ that have the biggest difference between their clustered and their real quantile value for a chosen x.

- *The solution processes in Lines 5 and 12 of Algorithm 2 can benefit from a series of improvements. First, every valid inequality from MSC$^+$ obtained through Proposition 8 is kept in the subsequent MSC$^+$ and ASC problems even if the associated cluster is split in the refinement step. Second, the OADM discussed in Section 4 is applied first, as a heuristic, for every MILP. Here, we use x_{LB} and x_{UB} (see Algorithm 2) as initial iterates for the OADM provided that we are in the corresponding clustered problem. Third, we use the valid inequalities in (15) to speed up the MILP solution process.*

6. Numerical Results

In this section, we present and discuss the results obtained by testing our methods on the maintenance planning and on the portfolio optimization problem.

To this end, we make use of the following notation. When assessing the performance of our MILP models, we consider the following three configurations. We write MILP meaning that we solve Model (4) when considering the MPP and Model (9) when considering the POP, both without valid inequalities. We write MILP_{VI} and $\text{MILP}_{\text{VI}^*}$ meaning that we solve Model (4) enriched with Inequalities (18) and Inequalities (19), when considering the MPP. In the case of the POP, MILP_{VI} and $\text{MILP}_{\text{VI}^*}$ stand for Model (9) enriched with Inequalities (24) and (25), respectively. Furthermore, we add the superscript OADM to the notation yet introduced, when the corresponding configuration is warm-started with an initial solution found by OADM as described in Section 4. For example, in the case of the MPP, $\text{MILP}_{\text{VI}^*}^{\text{OADM}}$ means that Model (4) enriched with Inequalities (19) is warm-started with an initial solution found by the OADM. Let $\mathcal{M}$ denote the set of considered methods. For $m \in \mathcal{M}$ we denote by v_{LB}^m and v_{UB}^m the best lower and upper bound values found by method m on a given problem instance. For the sake of simplicity, we will additionally use v_{LB} and v_{UB} when discussing the upper or lower bound without referring to a specific method.

In the following, we make some further comments regarding the tests we will discuss in Sections 6.1 and 6.2. In configurations MILP_{VI} and $\text{MILP}_{\text{VI}^*}$, we separate valid inequalities only at the root node of the branch-and-bound tree, as preliminary results showed that this is the best strategy. Indeed, separating them at each node of the tree significantly reduces the time left to explore the tree itself, leading to poor primal bounds—in particular in the case of the MPP. Moreover, the valid inequalities considered in $\text{MILP}_{\text{VI}^*}$ dominate the ones considered in MILP_{VI}; see Proposition 4. In the case of the MPP, this is reflected in the results obtained by performing some preliminary tests. Hence, in Section 6.1 we only discuss the results obtained by $\text{MILP}_{\text{VI}^*}$. This behavior does not occur when considering the POP. Thus, in Section 6.2, we discuss the results of both configurations. Finally, we remark that the solver does not struggle to provide feasible solutions of good quality on the POP instances when solving MILP, MILP_{VI}, or $\text{MILP}_{\text{VI}^*}$; see Section 6.2.

Consequently, we do not consider the inclusion of OADM in the solution procedure for this problem.

The MPP computations have been executed on a remote server composed of eleven separate machines. Each machine has 512 GB RAM and two Intel Xeon Gold 6258R processors with 28 cores and 2.70 GHz. To reduce the time to produce the MPP results all tests are run in parallel with a maximum of 95 GB RAM and four processor cores. The POP computations have been executed on a remote server with 64 GB RAM and an AMD Opteron 6176 SE processor with 12 cores and 2.30 GHz. The techniques presented in this paper have been coded in C++14 and are compiled using g++ version 9.3.0. All MILP models are solved using Gurobi 9.1.0. The time limit of the overall approach, i.e., including the seperation of the valid inequalities or the OADM if applied, is set to 90 min for which the time for reading the instance is ignored.

6.1. **Numerical Results for the Maintenance Planning Problem.** We first present the numerical results of the proposed methods when applied to the ROADEF/EURO 2020 challenge instances. Table 1 shows the main characteristics of these instances. Their names start with a capital letter that, based on their alphabetical order, correspond to a different phase of the ROADEF/EURO 2020 challenge. Hence, instances with a name starting with a letter appearing later in the alphabet are more likely to be computationally challenging. Note that we omit to list those instances that are trivial due to a very small number of scenarios.

The size of some instances makes the time to complete multiple OADM iterations too large w.r.t. the total computational time allowed. Thus, we set the stopping criterion of the OADM so that only the first iteration is applied. Here, the OADM gap is set to 0.05 for the x-direction (Line 3) and to 0 for the z-direction (Line 4). Finally, the computational parameters of ASCA are given in Table 2. All parameters are chosen based on preliminary numerical tests.

The detailed results obtained on the ROADEF/EURO instances with our methods are reported in Tables 8 and 9 in Appendix A. However, because of the large amount of ROADEF/EURO instances we use the bar plots in Figures 2–5 to visually compare the different methods. Let $\mathcal{M}'$ be the subset of methods of $\mathcal{M}$ considered in the plot. Then, for all $m \in \mathcal{M}'$ the plots show the re-scaled value of v^m_{LB} (left side) and v^m_{UB} (right side) for all instances. The bars' lengths are determined as follows. The left side of the bar plots is equal to $\min_{m' \in \mathcal{M}'}\{v^{m'}_{\text{LB}} \colon v^{m'}_{\text{LB}} > 0\}/v^m_{\text{LB}}$. The numerator $\min_{m' \in \mathcal{M}'}\{v^{m'}_{\text{LB}} \colon v^{m'}_{\text{LB}} > 0\}$ takes the value of the smallest lower bound obtained by the methods in $\mathcal{M}'$ while ignoring a lower bound if the method m' does not improve on $v^{m'}_{\text{LB}} = 0$. We apply a similar rule for the right side of the bar plots using $v^m_{\text{UB}}/\max_{m' \in \mathcal{M}'}\{v^{m'}_{\text{UB}} \colon v^{m'}_{\text{UB}} < \infty\}$. Here, $\max_{m' \in \mathcal{M}'}\{v^m_{\text{UB}} \colon v^{m'}_{\text{UB}} < \infty\}$ takes the value of the largest upper bound obtained by the two compared methods without considering $v^{m'}_{\text{UB}}$ if no incumbent is found during the solution process of m'. The rescaling of v_{LB} or v_{UB} is constant for a specific instance and therefore allows to easily visualize how they compare for two different methods. Hence, bars close to zero mean that the corresponding method performs better compared to the other method on the respective bound. However, we would like to draw attention to the fact that the spacing between the sides of the bars is not representative of the relative gap given by $(v_{\text{UB}} - v_{\text{LB}})/v_{\text{LB}}$.

In addition, in Appendix B, Tables 10 and 11 contain summary statistics that characterize the performance of the methods proposed in this paper when applied to the ROADEF/EURO challenge 2020 instances. These tables numerically underpin the key take-aways of the next paragraphs that are based on Figures 2–5.

TABLE 1. Characteristics of all the ROADEF/EURO 2020 challenge instances.

ID	$\lvert\mathcal{I}\rvert$	$\lvert\mathcal{R}\rvert$	$\lvert\mathcal{T}\rvert$	$\sum_{t\in\mathcal{T}} \frac{\lvert\mathcal{S}_t\rvert}{\lvert\mathcal{T}\rvert}$	$\lvert\mathcal{D}\rvert$
A02	89	9	90	120.00	1869
A05	180	9	182	120.00	6791
A08	18	9	17	645.59	29
A11	54	9	53	639.53	96
A14	108	10	53	160.30	438
A15	108	10	53	320.06	438
B01	100	9	53	191.45	553
B02	100	9	53	191.45	404
B03	706	9	53	63.49	23674
B04	706	9	53	63.49	23674
B05	706	9	53	63.49	27276
B06	100	9	53	255.42	404
B07	250	9	53	191.45	3787
B08	119	9	42	254.07	550
B09	120	9	42	127.40	730
B10	398	9	25	192.40	3231
B11	100	9	53	191.45	679
B12	495	9	102	63.91	20205
B13	99	9	102	159.51	148
B14	297	9	191	95.50	14448
B15	495	9	250	63.38	61786
C01	120	9	53	191.45	1080
C02	120	9	53	191.45	828
C03	706	9	53	63.49	24260
C04	706	9	53	63.49	23638
C05	706	9	53	63.49	27276
C06	280	9	53	191.45	3404
C07	120	9	42	126.76	578
C08	426	9	25	192.88	3405
C09	110	9	53	191.45	718
C10	522	9	102	63.24	26250
C11	89	9	102	191.05	1474
C12	298	9	191	95.21	13996
C13	505	9	230	63.40	44384
C14	465	9	220	95.34	53628
C15	528	9	300	50.69	69715
X01	120	9	53	191.45	917
X02	706	9	53	63.49	24464
X03	280	9	53	191.45	3299
X04	426	9	25	188.84	4509
X05	467	9	220	95.30	48595
X06	528	9	300	50.64	79180
X07	209	9	300	63.52	8873
X08	209	9	300	63.60	6032
X09	548	9	30	156.97	8942
X10	460	9	35	159.54	7083
X11	521	9	131	63.35	35112
X12	522	9	131	63.92	35241
X13	336	9	212	95.27	19978
X14	613	9	180	63.73	57762
X15	613	9	180	63.32	64400

TABLE 2. ASCA parameters.

Θ	OADM gap (Line 3)	OADM gap (Line 4)	Solver gap (Lines 5,12)
0.25	0.0025	0	0.025

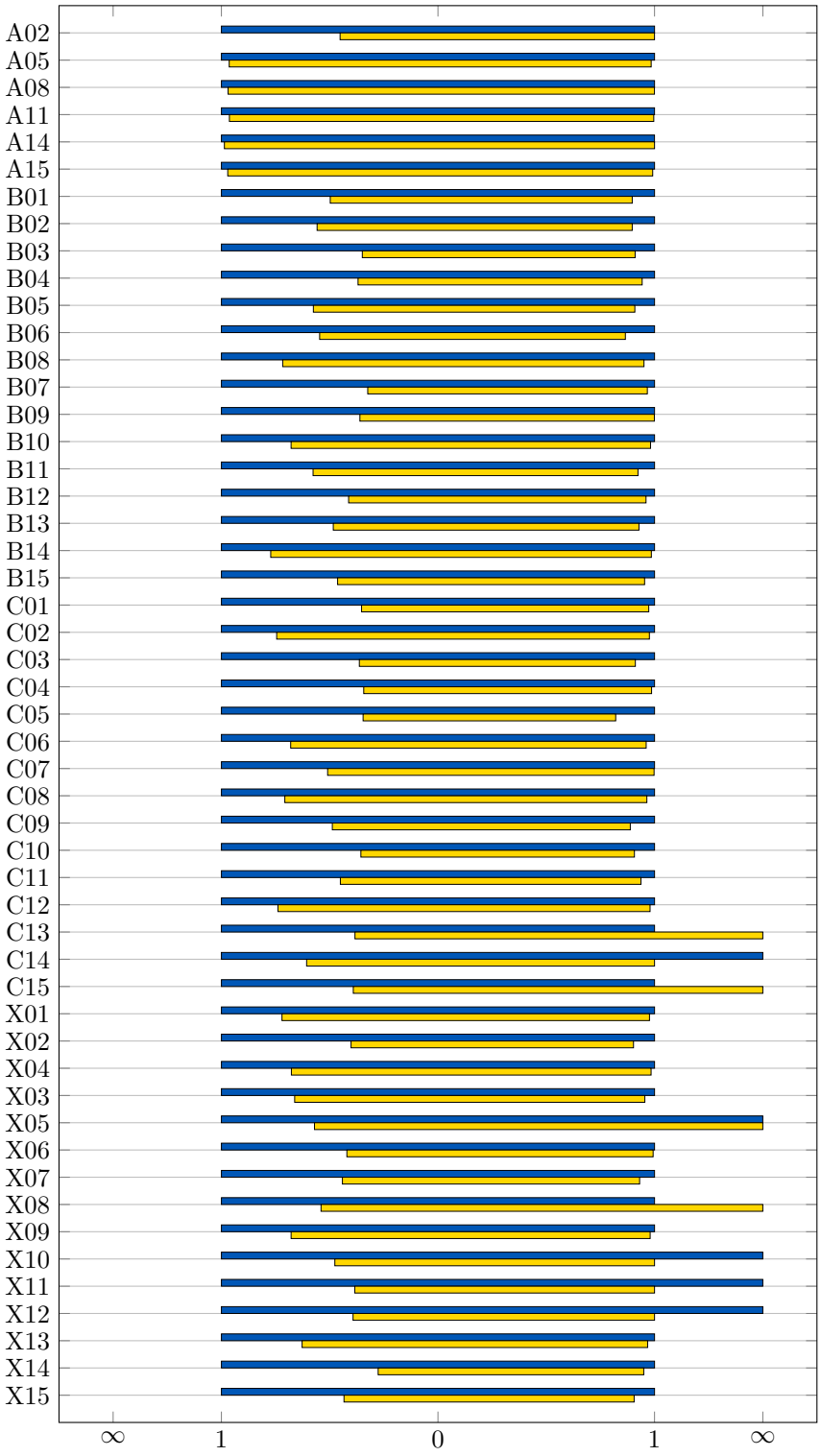

FIGURE 2. Bar plot for $\min_{m' \in \mathcal{M}'}\{v_{\mathrm{LB}}^{m'} \colon v_{\mathrm{LB}}^{m'} > 0\}/v_{\mathrm{LB}}^{m}$ (left side) and $v_{\mathrm{UB}}^{m}/\max_{m' \in \mathcal{M}'}\{v_{\mathrm{UB}}^{m'} \colon v_{\mathrm{UB}}^{m'} < \infty\}$ (right side), where $\mathcal{M}'$ is composed of MILP (blue) and $\mathrm{MILP}_{\mathrm{VI}^*}$ (yellow).

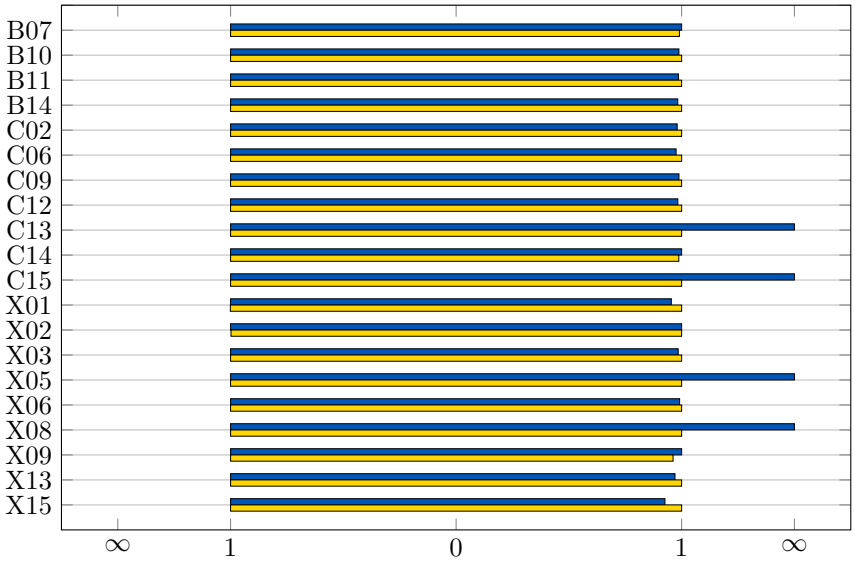

FIGURE 3. Bar plot for $\min_{m' \in \mathcal{M}'}\{v_{\mathrm{LB}}^{m'} \colon v_{\mathrm{LB}}^{m'} > 0\}/v_{\mathrm{LB}}^{m}$ (left side) and $v_{\mathrm{UB}}^{m}/\max_{m' \in \mathcal{M}'}\{v_{\mathrm{UB}}^{m'} \colon v_{\mathrm{UB}}^{m'} < \infty\}$ (right side), where $\mathcal{M}'$ is composed of $\mathrm{MILP}_{\mathrm{VI}^*}^{\mathrm{OADM}}$ (yellow) and $\mathrm{MILP}_{\mathrm{VI}^*}$ (blue). Only the instances with a significant difference in the results for v_{LB} and v_{UB} are displayed.

Figure 2 shows a comparison between the results obtained by MILP with those obtained by $\mathrm{MILP}_{\mathrm{VI}^*}$. We observe that $\mathrm{MILP}_{\mathrm{VI}^*}$ outperforms MILP on most of the instances, thus yielding an improvement on both v_{LB} and v_{UB}, where the lower bound improves more significantly than the upper bound. However, we remark that $\mathrm{MILP}_{\mathrm{VI}^*}$ fails to provide an incumbent solution for some instances within the time limit for which MILP succeeds in doing so; see, e.g., C13. Most likely, this has two reasons. First, additional time is needed to separate violated inequalities in the first node of the branch-and-bound tree. Second, the model's relaxations are a bit harder to solve after adding valid inequalities due to their increased size.

To counteract this effect, we also tested to apply the OADM described in Section 4 to the MILP model with valid inequalities for the MPP as a primal root node heuristic. Figure 3 shows the comparison of $\mathrm{MILP}_{\mathrm{VI}^*}$ and $\mathrm{MILP}_{\mathrm{VI}^*}^{\mathrm{OADM}}$. In the latter configuration, the point to warm-start is the result of a single iteration of the OADM. The resulting bar plot only displays the instances with a significant difference in v_{LB} or v_{UB}. Note that a single iteration of the OADM allows to obtain an incumbent for all the instances for which the resolution of $\mathrm{MILP}_{\mathrm{VI}^*}$ fails to do so.

For the remaining instances, we see that activating a single OADM step mostly performs worse in terms of v_{UB}. These results suggest that for large instances one should use the OADM as a primal heuristic to warm-start the resolution of Model (4) as a guarantee to obtain an incumbent solution. Figure 4 shows the results obtained by the ASCA presented in Section 5 when compared to the results obtained by MILP. One can observe that ASCA mostly provides better results in terms of both v_{LB} and v_{UB}. However, it fails to get a strictly positive value for v_{LB} for some of the more computationally challenging instances. As it is the case for $\mathrm{MILP}_{\mathrm{VI}^*}$, for both v_{LB} and v_{UB} we see that ASCA outperforms MILP. For those instances for which ASCA fails to obtain a strictly positive value for v_{LB}, ASCA keeps improving v_{UB} using the ASC problem during the initial iterations of the

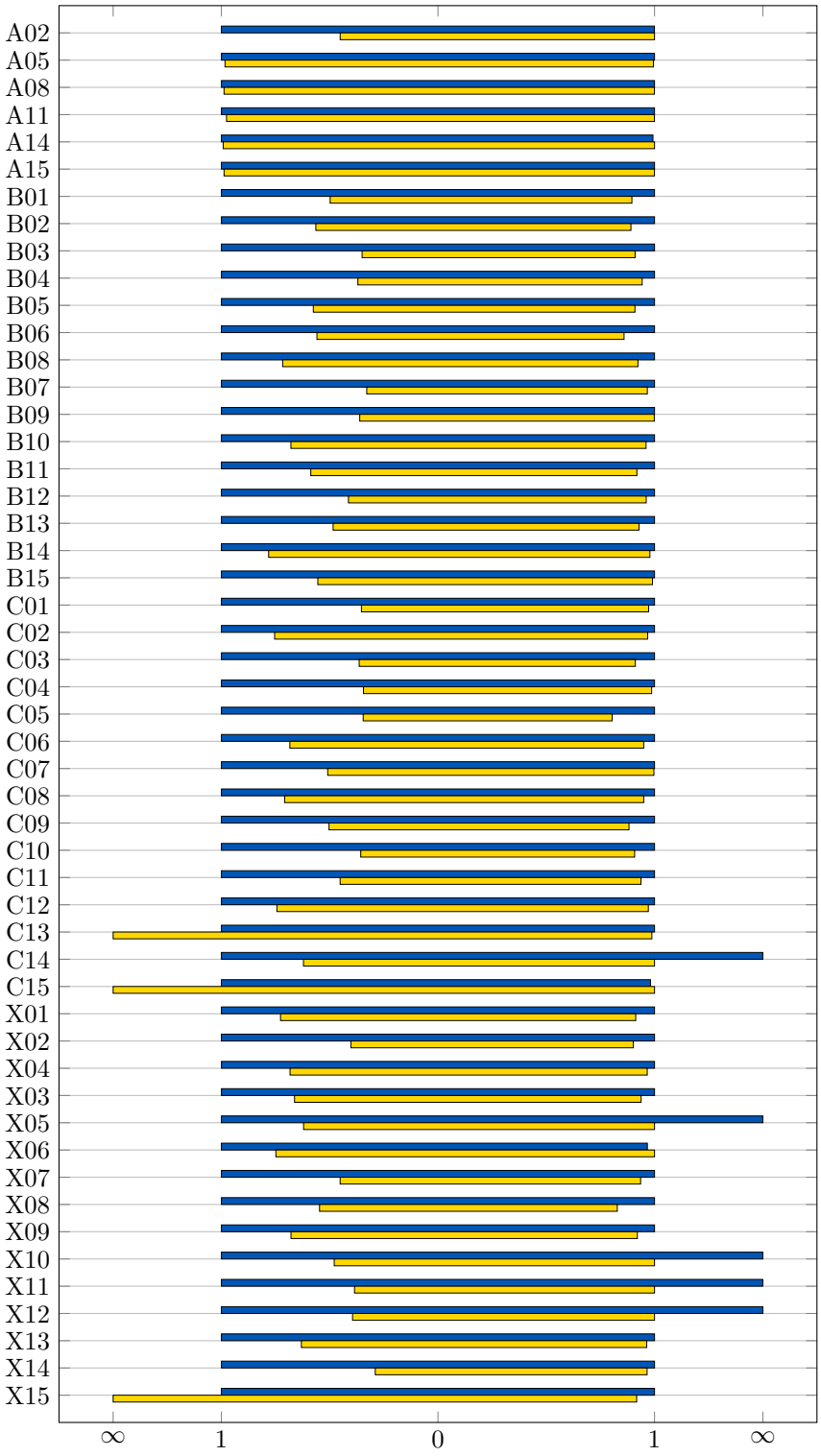

FIGURE 4. Bar plot for $\min_{m' \in \mathcal{M}'}\{v_{\mathrm{LB}}^{m'} \colon v_{\mathrm{LB}}^{m'} > 0\}/v_{\mathrm{LB}}^{m}$ (left side) and $v_{\mathrm{UB}}^{m}/\max_{m' \in \mathcal{M}'}\{v_{\mathrm{UB}}^{m'} \colon v_{\mathrm{UB}}^{m'} < \infty\}$ (right side), where $\mathcal{M}'$ is composed of MILP (blue) and ASCA (yellow).

algorithm. Since v_{UB} keeps decreasing in every iteration, ASCA never enters the MSC^+ problem before reaching the given time limit. Thus, it never improves on the $v_{\mathrm{LB}} = 0$ lower bound. Similarly, MILP fails to compute incumbent solutions for some of the more computationally challenging instances and, hence, does not decrease the $v_{\mathrm{UB}} = \infty$ bound. On the contrary, ASCA always finds an improved incumbent solution since it is designed to start solving the ASC problem to decrease the value of v_{UB}.

We close the comparison of the methods with Figure 5, which compares ASCA and $\mathrm{MILP}^{\mathrm{OADM}}_{\mathrm{VI}^*}$, i.e., it compares the methods that perform best in terms of v_{UB} and v_{LB}. Considering the values of v_{LB}, we see that $\mathrm{MILP}^{\mathrm{OADM}}_{\mathrm{VI}^*}$ always outperforms ASCA except for the instances A02, B05, C05, and C07. The opposite situation occurs when the two configurations are compared w.r.t. the values of v_{UB}, i.e., ASCA outperforms $\mathrm{MILP}^{\mathrm{OADM}}_{\mathrm{VI}^*}$ on most of the instances. Based on these observations, we recommend using ASCA if a primal solution of good quality is requested. In contrast, we recommend using $\mathrm{MILP}^{\mathrm{OADM}}_{\mathrm{VI}^*}$ if one wants to focus on closing the optimality gap.

6.2. **Numerical Results for the Portfolio Optimization Problem.** We consider synthetic instances as well as instances based on real-world data to test our methods on the portfolio optimization problem. More precisely, we build a first test set of 24 synthetic instances following the procedure used in [24] to generate instances for the chance-constrained portfolio optimization model. These instances contain $n \in \{20, 200\}$ equities and $|\mathcal{S}| = 200$ equiprobable scenarios with $p^s = 1/|\mathcal{S}|$, $s \in \mathcal{S}$. We draw the components of return vectors r^s, $s \in \mathcal{S}$, using an independent uniform distribution on the interval $[80, 150]$. This means, we generate equities whose returns range between a 20 % loss and a 50 % profit on the investment. The minimum expected return ρ is set to 110, forcing an average portfolio return of 10 % w.r.t. the invested budget. We consider the VaR confidence levels of 7.5 %, 15 %, and 22.5 %, i.e., $\tau \in \{0.075, 0.15, 0.225\}$ holds. The second test set consists of 64 instances based on the four real-world datasets considered in [23], for which we report the names and the main characteristics:

- **FTSE100 (Financial Times Stock Exchange, UK):** 83 equities, 625 price observations from 19/04/2004 to 04/04/2016;
- **SP500 (Standard & Poor's, USA):** 442 equities, 573 price observations from 18/04/2005 to 04/04/2016;
- **EUROSTOXX50 (Europe's leading blue-chip index, EU):** 49 equities, 729 price observations from 22/04/2002 to 04/04/2016;
- **DJIA (Dow Jones Industrial Average, USA):** 28 equities, 1353 price observations from 07/05/1990 to 04/04/2016.

Each dataset contains the weekly price observations of the equities and of the market index. The instances built from the same dataset are characterized by the number of equities in the dataset and a number of scenarios equal to the number of periods in the dataset minus one. Scenarios are considered to be equiprobable. Given an equity i and a scenario $s \in \mathcal{S}$, we compute the return of i in s as $r_i^s = (o_i^s - o_i^{s-1})/o_i^s$, where o_i^{s-1} and o_i^s are the price observations of i in periods $s-1$ and s, respectively. The minimum expected return ρ is set to the average market index return as suggested in [23]. Observe that we re-scale all returns to have losses represented by values lower than 100 and profits by values greater than 100, as for the synthetic instances. Here, the VaR confidence levels are 0.5 %, 1 %, 5 % and 10 %, i.e., $\tau \in \{0.005, 0.01, 0.05, 0.1\}$. Note that the first two values are also often used in practice [11]. Finally, for both types of instances, the objective function weight α takes values in set $\{0, 0.25, 0.5, 0.75\}$. Smaller values of α favor

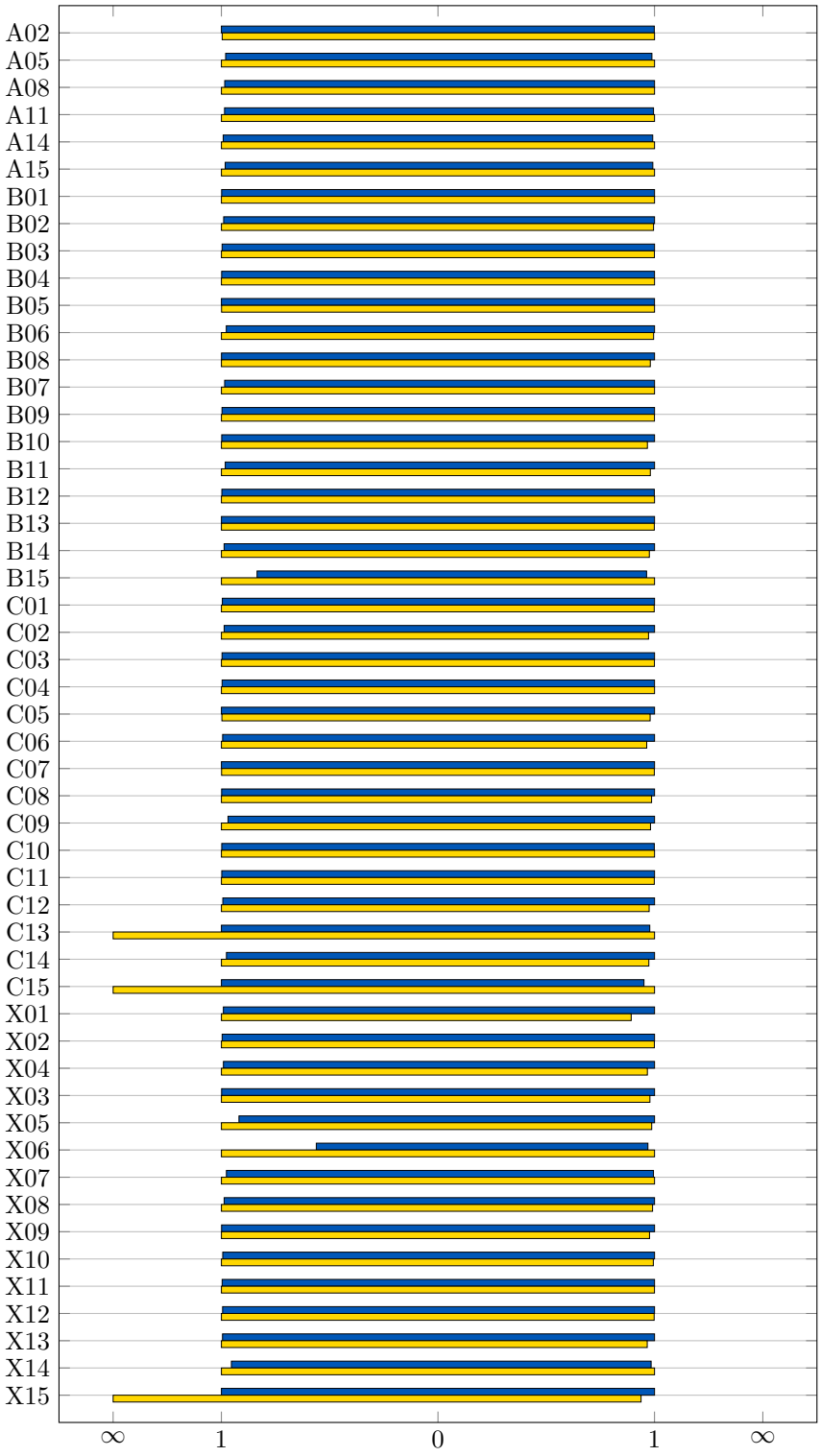

FIGURE 5. Bar plot for $\min_{m' \in \mathcal{M}'}\{v_{\mathrm{LB}}^{m'} \colon v_{\mathrm{LB}}^{m'} > 0\}/v_{\mathrm{LB}}^{m}$ (left side) and $v_{\mathrm{UB}}^{m}/\max_{m' \in \mathcal{M}'}\{v_{\mathrm{UB}}^{m'} \colon v_{\mathrm{UB}}^{m'} < \infty\}$ (right side), where $\mathcal{M}'$ is composed of $\mathrm{MILP}_{\mathrm{VI}^*}^{\mathrm{OADM}}$ (blue) and ASCA (yellow).

risk minimization over return maximization—larger values of α favor the opposite behavior.

First, we discuss the results obtained by solving the 24 synthetic instances by means of MILP, MILP_{VI}, $\text{MILP}_{\text{VI}^*}$, and ASCA. The results are given in Table 3.

TABLE 3. Results obtained by solving the instances of the portfolio optimization problem built following the procedure used in [24] with configurations MILP, MILP_{VI}, $\text{MILP}_{\text{VI}^*}$, and ASCA within a time limit of 90 minutes.

	Parameters					MILP			MILP_{VI}			$\text{MILP}_{\text{VI}^*}$			ASCA		
ID	Assets	$\lvert\mathcal{S}\rvert$	α	ρ	τ	v_{LB}	v_{UB}	t_{solve}(s) gap(%)	v_{LB}	v_{UB}	t_{solve}(s) gap(%)	v_{LB}	v_{UB}	t_{solve}(s) gap(%)	v_{LB}	v_{UB}	t_{solve}(s) gap(%)
W1	20	200	0	110	0.075	111.05	111.05	934.52s	111.05	111.05	1058.42s	111.05	111.05	885.98s	106.36	111.61	4.94%
W2	20	200	0	110	0.15	112.63	114.83	1.95%	112.63	114.77	1.90%	112.63	114.84	1.96%	107.66	120.20	11.65%
W3	20	200	0	110	0.225	113.84	119.22	4.72%	113.72	118.83	4.49%	113.82	120.51	5.88%	108.97	127.52	17.02%
W4	20	200	0.25	110	0.075	112.56	112.56	370.78s	112.56	112.56	380.56s	112.56	112.56	1226.77s	111.15	113.02	1.68%
W5	20	200	0.25	110	0.15	113.33	115.17	1.62%	113.26	115.41	1.90%	113.33	115.39	1.82%	110.33	118.96	7.82%
W6	20	200	0.25	110	0.225	115.05	119.07	3.49%	115.05	119.72	4.05%	115.01	119.54	3.95%	112.85	124.06	9.93%
W7	20	200	0.5	110	0.075	113.78	113.78	325.91s	113.78	113.78	113.78s	113.78	113.78	206.89s	112.65	113.79	1.00%
W8	20	200	0.5	110	0.15	114.35	115.71	1.19%	114.39	115.73	1.18%	114.38	116.21	1.60%	113.38	-	-
W9	20	200	0.5	110	0.225	114.96	117.76	2.44%	114.91	118.12	2.79%	114.88	118.10	2.81%	112.69	121.51	7.83%
W10	20	200	0.75	110	0.075	114.31	114.31	113.19s	114.31	114.31	92.86s	114.31	114.31	126.28s	114.31	114.31	2862.10s
W11	20	200	0.75	110	0.15	114.64	115.36	0.63%	114.68	115.27	0.52%	114.68	115.33	0.57%	114.13	-	-
W12	20	200	0.75	110	0.225	115.13	115.81	0.60%	115.13	115.96	0.72%	115.13	115.87	0.65%	113.71	117.82	3.61%
av^{20}						113.80	115.39	4 1.39%	113.79	115.46	4 1.46%	113.80	115.62	4 1.60%	111.52	-	-
W13	400	200	0	110	0.075	116.55	120.70	3.56%	116.55	119.16	2.24%	116.61	119.17	2.19%	112.19	-	-
W14	400	200	0	110	0.15	117.09	125.63	7.29%	117.22	122.35	4.38%	117.23	122.36	4.38%	110.36	134.98	22.31%
W15	400	200	0	110	0.225	117.60	130.21	10.73%	117.81	125.78	6.76%	117.81	125.74	6.73%	112.28	139.45	24.19%
W16	400	200	0.25	110	0.075	116.74	119.85	2.67%	116.76	118.70	1.66%	116.75	118.71	1.68%	111.57	124.68	11.76%
W17	400	200	0.25	110	0.15	117.20	123.38	5.27%	117.23	121.10	3.30%	117.23	121.13	3.33%	113.38	130.55	15.14%
W18	400	200	0.25	110	0.225	117.78	126.68	7.56%	117.83	123.70	4.99%	117.72	123.73	5.11%	113.45	132.81	17.07%
W19	400	200	0.5	110	0.075	116.91	118.83	1.64%	116.90	118.11	1.04%	116.91	118.12	1.04%	111.85	122.12	9.18%
W20	400	200	0.5	110	0.15	117.38	121.24	3.29%	117.33	120.07	2.34%	117.33	120.09	2.35%	114.62	125.37	9.38%
W21	400	200	0.5	110	0.225	117.58	123.29	4.86%	117.68	121.52	3.26%	117.65	121.51	3.28%	115.83	-	-
W22	400	200	0.75	110	0.075	117.25	118.14	0.76%	117.24	117.78	0.46%	117.25	117.79	0.46%	115.51	119.13	3.14%
W23	400	200	0.75	110	0.15	117.20	119.01	1.55%	117.21	118.55	1.14%	117.22	118.52	1.10%	115.84	121.14	4.57%
W24	400	200	0.75	110	0.225	117.66	120.39	2.32%	117.68	119.70	1.71%	117.72	119.75	1.73%	116.92	-	-
av^{400}						117.25	122.28	0 4.29%	117.29	120.54	0 2.77%	117.29	120.55	0 2.78%	113.65	-	-
av^{all}						115.52	118.83	4 2.84%	115.54	118.00	4 2.12%	115.54	118.09	4 2.19%	112.58	-	-

Each row of the table corresponds to an instance for which its parameterization is summarized in columns two to six of the table. Then, for each approach, we include three columns. The first and second columns contain the lower and upper bounds (v_{LB} and v_{UB}), respectively. The third is a mixed time/gap (t_{solve}/gap) column reporting either the computational time if the corresponding method solves the instance to global optimality within the time limit or the relative optimality gap in percentage ($100\,(v_{\mathrm{UB}} - v_{\mathrm{LB}})/v_{\mathrm{LB}}$) otherwise. Rows labeled av^{20}, av^{400}, and $\mathrm{av}^{\mathrm{all}}$ report average results on the instances with 20 assets, 400 assets, and on all instances, respectively. Here, columns labeled v_{LB} and v_{UB} report the average values of the lower and upper bounds respectively, while for the mixed time/gap column we report the number of instances solved to optimality by each of the methods and the average gap (obviously considering a gap of zero for the instances solved to optimality). Note that since ASCA does not provide upper bounds on all instances, and thus we do not report averages for the v_{UB} and the mixed time/gap column.

In general, the methods MILP, $\mathrm{MILP}_{\mathrm{VI}}$, and $\mathrm{MILP}_{\mathrm{VI}^*}$ yield comparable results. They provide feasible points of comparable value (v_{LB}) on all the instances. As opposed to the MPP, such values are of good quality. Indeed, the average optimality gap already after the resolution of the root node is rather small. It is equal to 11.5 %, 6.1 %, and 5.6 % for configurations MILP, $\mathrm{MILP}_{\mathrm{VI}}$, and $\mathrm{MILP}_{\mathrm{VI}^*}$, respectively. In addition, the methods prove optimality for the same four instances (W1, W4, W7, W10) on average in 436, 411, and 611 seconds, for MILP, $\mathrm{MILP}_{\mathrm{VI}}$, and $\mathrm{MILP}_{\mathrm{VI}^*}$ respectively. The average optimality gap returned by the solver on the instances for which optimality is not proven within the time limit is 2.84 % for the plain MILP, 2.12 % for $\mathrm{MILP}_{\mathrm{VI}}$, and 2.19 % for $\mathrm{MILP}_{\mathrm{VI}^*}$. ASCA is outperformed by MILP, $\mathrm{MILP}_{\mathrm{VI}}$, and $\mathrm{MILP}_{\mathrm{VI}^*}$. Indeed, ASCA finds one optimal solution (instance W10) and fails to provide an upper bound for five instances. The average optimality gap of the remaining ones is 9.59 %. One possible reason for this behavior might be the structure of the synthetic instances since the distribution of the return values might prevent ASCA to identify clusters of good quality.

Next, we asses the impact of the different families of valid inequalities introduced in Section 3.2 for the POP, i.e., we compare the results obtained by MILP, $\mathrm{MILP}_{\mathrm{VI}}$, and $\mathrm{MILP}_{\mathrm{VI}^*}$ on the synthetic instances. Note that for $\mathrm{MILP}_{\mathrm{VI}^*}$, where the separation of the inequalities in (25) is NP-hard (see Proposition 6), we separate them by means of the same procedure employed for the inequalities in (24). By doing so, the time spent to separate and build the two families of valid inequalities differs by a factor of ten, which is due to the additional time spent to evaluate $b_i(\cdot)$. However, in both cases, this time remains negligible w.r.t. the computational time limit since it is always less than 4 s. For this analysis, we restrict ourselves to consider only the large instances W13 to W24, where the inequalities have the largest impact. Note that none of these instances is solved to optimality regardless of the configuration. In what follows, the upper bound improvement of one configuration vs. another one, e.g., MILP vs. $\mathrm{MILP}_{\mathrm{VI}}$, is computed as $(v_{\mathrm{UB}}^{\mathrm{MILP}} - v_{\mathrm{UB}}^{\mathrm{MILP}_{\mathrm{VI}}})/v_{\mathrm{UB}}^{\mathrm{MILP}}$, where $v_{\mathrm{UB}}^{\mathrm{MILP}}$ and $v_{\mathrm{UB}}^{\mathrm{MILP}_{\mathrm{VI}}}$ are the upper bounds returned by the solver for the configurations MILP and $\mathrm{MILP}_{\mathrm{VI}}$, respectively. The improvements regarding the optimality gap are computed analogously.

First, we observe that the introduction of either inequalities (24) or (25) yields benefits in terms of both decreasing the upper bound and reducing the optimality gap. Indeed, the upper bound and integrality gap improvements on average are equal to 1.39 % and about 34.5 % when comparing MILP to both $\mathrm{MILP}_{\mathrm{VI}}$ and $\mathrm{MILP}_{\mathrm{VI}^*}$ (see the lines labeled with W13–W24 in Table 4). The same trend emerges when analyzing the values of the upper bound and the optimality gap after the

TABLE 4. Average upper bound and integrality gap improvements (in percentage) for the synthetic instances

	Comparison	avg. v_{UB} impr. (%)	avg. gap impr. (%)
Root node resolution	MILP vs. MILP_{VI}	1.73	34.50
	MILP vs. $\text{MILP}_{\text{VI}^*}$	1.77	39.15
	MILP_{VI} vs. $\text{MILP}_{\text{VI}^*}$	0.05	4.85
W13 - W24	MILP vs. MILP_{VI}	1.39	34.48
	MILP vs. $\text{MILP}_{\text{VI}^*}$	1.39	34.47
	MILP_{VI} vs. $\text{MILP}_{\text{VI}^*}$	-0.01	-0.05

solution of the root node of the branch-and-bound tree; see the row "Root node resolution" in Table 4.

As mentioned in the introduction of this section, although Inequalities (24) are dominated by Inequalities (25), see Proposition 6, this relation is not reflected in the computational results. Indeed, when comparing MILP_{VI} to $\text{MILP}_{\text{VI}^*}$, the optimality gap tends to be smaller for MILP_{VI}. The latter provides an average optimality gap of 2.12 % vs. 2.19 % provided by $\text{MILP}_{\text{VI}^*}$, as can be seen from line av^{all} of Table 3. Conversely, a slight dominance of configuration $\text{MILP}_{\text{VI}^*}$ over configuration MILP_{VI} arises from the results obtained after the solution of the root node; see Table 4. To explain this behavior, we note that the separation procedure for Inequalities (25) is heuristic.

Finally, Tables 5 and 6 report the results obtained by MILP, MILP_{VI}, $\text{MILP}_{\text{VI}^*}$, and ASCA on the instances based on real-world data. Again, each row of the two tables corresponds to an instance. The meaning of the columns is the same as before. The average results on the instances built from the same dataset are reported in the rows labeled with av^{FTSE}, av^{SP}, av^{EURO}, and av^{DJIA}. In both tables, rows identified by av^{all} contain the average results on all the instances in the table.

As for the synthetic instances, MILP, MILP_{VI}, and $\text{MILP}_{\text{VI}^*}$ yield similar good results. The performances of ASCA improve on the instances derived from the real-world datasets. Indeed, a feasible point is found for all the instances. Its average objective function value is slightly better for MILP, MILP_{VI}, and $\text{MILP}_{\text{VI}^*}$ compared to the one of the feasible points found by ASCA. In addition, all four methods prove the optimality of the 32 instances characterized by the VaR confidence level equal to 0.5 % or 1 %. MILP and MILP_{VI} provide three and one additional optima, respectively. The average optimality gap on all instances is of 0.37 % for MILP and 0.29 % for both MILP_{VI} and $\text{MILP}_{\text{VI}^*}$. It increases to 0.97 % with ASCA.

TABLE 5. Results obtained by solving the instances of the portfolio optimization problem that originate from [22] with configurations MILP, MILP_{VI}, $\text{MILP}_{\text{VI}^*}$, and ASCA within a time limit of 90 minutes.

	Parameters				MILP			MILP_{VI}			$\text{MILP}_{\text{VI}^*}$			ASCA		
ID	Assets	$\lvert\mathcal{S}\rvert$	α	τ	v_{LB}	v_{UB}	t_{solve}(s) gap(%)	v_{LB}	v_{UB}	t_{solve}(s) gap(%)	v_{LB}	v_{UB}	t_{solve}(s) gap(%)	v_{LB}	v_{UB}	t_{solve}(s) gap(%)
FTSE100	83	624	0	0.005	96.05	96.05	1.16s	96.05	96.05	3.20s	96.05	96.05	3.87s	96.05	96.05	7.00s
FTSE100	83	624	0	0.01	96.60	96.60	5.39s	96.60	96.60	6.14s	96.60	96.60	4.99s	96.60	96.60	11.41s
FTSE100	83	624	0	0.05	98.06	98.66	0.62%	98.06	98.67	0.62%	98.05	98.64	0.61%	97.24	98.97	1.78%
FTSE100	83	624	0	0.1	98.68	99.93	1.27%	98.65	99.75	1.11%	98.62	99.81	1.20%	97.45	101.13	3.78%
FTSE100	83	624	0.25	0.005	97.11	97.11	1.05s	97.11	97.11	3.45s	97.11	97.11	3.48s	97.11	97.11	8.47s
FTSE100	83	624	0.25	0.01	97.51	97.51	5.40s	97.51	97.51	4.93s	97.51	97.51	5.78s	97.51	97.51	14.76s
FTSE100	83	624	0.25	0.05	98.61	99.02	0.42%	98.61	99.05	0.44%	98.61	99.08	0.48%	97.95	99.18	1.27%
FTSE100	83	624	0.25	0.1	99.04	100.07	1.04%	99.06	99.93	0.88%	99.04	99.90	0.86%	98.24	100.92	2.73%
FTSE100	83	624	0.5	0.005	98.16	98.16	1.11s	98.16	98.16	3.70s	98.16	98.16	3.43s	98.16	98.16	6.30s
FTSE100	83	624	0.5	0.01	98.42	98.42	6.64s	98.42	98.42	5.21s	98.42	98.42	8.47s	98.42	98.42	10.98s
FTSE100	83	624	0.5	0.05	99.16	99.43	0.28%	99.15	99.48	0.33%	99.16	99.46	0.30%	98.67	99.65	0.99%
FTSE100	83	624	0.5	0.1	99.47	100.13	0.67%	99.44	100.05	0.62%	99.45	100.05	0.60%	99.19	101.10	1.93%
FTSE100	83	624	0.75	0.005	99.24	99.24	1.50s	99.24	99.24	2.69s	99.24	99.24	2.51s	99.24	99.24	7.39s
FTSE100	83	624	0.75	0.01	99.34	99.34	5.63s	99.34	99.34	7.69s	99.34	99.34	7.50s	99.34	99.34	13.14s
FTSE100	83	624	0.75	0.05	99.74	99.87	0.13%	99.73	99.87	0.13%	99.74	99.84	0.10%	99.56	99.90	0.35%
FTSE100	83	624	0.75	0.1	99.90	100.21	0.31%	99.90	100.20	0.30%	99.90	100.20	0.30%	99.54	100.51	0.98%
av^{FTSE}					98.44	98.73	8 0.30%	98.44	98.71	8 0.28%	98.44	98.71	8 0.28%	98.14	98.99	8 0.86%
SP500	442	572	0	0.005	96.96	96.96	4.21s	96.96	96.96	6.45s	96.96	96.96	6.83s	96.96	96.96	38.78s
SP500	442	572	0	0.01	97.46	97.46	29.22s	97.46	97.46	26.13s	97.46	97.46	31.84s	97.46	97.46	104.39s
SP500	442	572	0	0.05	98.43	99.60	1.19%	98.44	99.20	0.78%	98.46	99.22	0.77%	96.92	100.24	3.43%
SP500	442	572	0	0.1	99.02	101.21	2.20%	99.02	100.25	1.25%	99.02	100.18	1.17%	96.92	101.66	4.89%
SP500	442	572	0.25	0.005	97.79	97.79	3.44s	97.79	97.79	6.02s	97.79	97.79	6.32s	97.79	97.79	33.82s
SP500	442	572	0.25	0.01	98.16	98.16	27.43s	98.16	98.16	20.20s	98.16	98.16	19.88s	98.16	98.16	84.75s
SP500	442	572	0.25	0.05	98.91	99.79	0.89%	98.90	99.50	0.60%	98.90	99.46	0.57%	97.70	100.38	2.75%
SP500	442	572	0.25	0.1	99.32	101.15	1.83%	99.33	100.22	0.89%	99.33	100.24	0.92%	97.06	101.42	4.50%
SP500	442	572	0.5	0.005	98.62	98.62	3.23s	98.62	98.62	5.64s	98.62	98.62	5.94s	98.62	98.62	30.44s
SP500	442	572	0.5	0.01	98.87	98.87	21.04s	98.87	98.87	14.83s	98.87	98.87	21.59s	98.87	98.87	109.42s
SP500	442	572	0.5	0.05	99.37	99.93	0.57%	99.37	99.77	0.40%	99.37	99.75	0.39%	98.73	100.28	1.58%
SP500	442	572	0.5	0.1	99.65	100.80	1.15%	99.65	100.30	0.64%	99.65	100.31	0.66%	98.63	101.11	2.52%
SP500	442	572	0.75	0.005	99.47	99.47	3.05s	99.47	99.47	4.63s	99.47	99.47	4.75s	99.47	99.47	27.81s
SP500	442	572	0.75	0.01	99.60	99.60	19.75s	99.60	99.60	19.58s	99.60	99.60	23.18s	99.60	99.60	72.37s
SP500	442	572	0.75	0.05	99.88	100.15	0.27%	99.89	100.09	0.20%	99.89	100.08	0.19%	99.34	100.34	1.01%
SP500	442	572	0.75	0.1	100.08	100.60	0.51%	100.08	100.44	0.37%	100.08	100.44	0.35%	99.62	100.85	1.23%
av^{SP}					98.85	99.39	8 0.54%	98.85	99.17	8 0.32%	98.85	99.16	8 0.31%	98.24	99.58	8 1.37%
av^{all}					98.65	99.06	16 0.42%	98.65	98.94	16 0.30%	98.64	98.94	16 0.30%	98.19	99.28	16 1.12%

TABLE 6. Results obtained by solving the instances of the portfolio optimization problem that originate from [22] with configurations MILP, MILP_{VI}, $\text{MILP}_{\text{VI}^*}$, and ASCA within a time limit of 90 minutes.

	Parameters				MILP			MILP_{VI}			$\text{MILP}_{\text{VI}^*}$			ASCA		
ID	Assets	$\lvert\mathcal{S}\rvert$	α	τ	v_{LB}	v_{UB}	t_{solve}(s) gap(%)	v_{LB}	v_{UB}	t_{solve}(s) gap(%)	v_{LB}	v_{UB}	t_{solve}(s) gap(%)	v_{LB}	v_{UB}	t_{solve}(s) gap(%)
EuroStoxx50	49	728	0	0.005	93.55	93.55	1.35s	93.55	93.55	4.20s	93.55	93.55	4.16s	93.55	93.55	4.76s
EuroStoxx50	49	728	0	0.01	94.75	94.75	2.00s	94.75	94.75	3.68s	94.75	94.75	3.39s	94.75	94.75	4.85s
EuroStoxx50	49	728	0	0.05	97.08	97.37	0.30%	97.10	97.39	0.30%	97.10	97.38	0.28%	96.52	97.37	0.89%
EuroStoxx50	49	728	0	0.1	98.06	99.09	1.06%	98.06	98.97	0.93%	98.07	98.96	0.91%	97.71	99.82	2.16%
EuroStoxx50	49	728	0.25	0.005	95.20	95.20	1.31s	95.20	95.20	5.47s	95.20	95.20	5.47s	95.20	95.20	4.61s
EuroStoxx50	49	728	0.25	0.01	96.11	96.11	1.69s	96.11	96.11	3.39s	96.11	96.11	3.16s	96.11	96.11	6.87s
EuroStoxx50	49	728	0.25	0.05	97.88	97.88	2322.74s	97.86	98.08	0.22%	97.88	98.10	0.23%	97.44	98.15	0.74%
EuroStoxx50	49	728	0.25	0.1	98.62	99.38	0.77%	98.58	99.30	0.73%	98.57	99.31	0.75%	97.92	100.35	2.49%
EuroStoxx50	49	728	0.5	0.005	96.86	96.86	1.15s	96.86	96.86	3.21s	96.86	96.86	3.18s	96.86	96.86	5.18s
EuroStoxx50	49	728	0.5	0.01	97.47	97.47	1.71s	97.47	97.47	3.48s	97.47	97.47	3.68s	97.47	97.47	6.32s
EuroStoxx50	49	728	0.5	0.05	98.65	98.65	1645.94s	98.65	98.72	0.07%	98.64	98.80	0.16%	98.48	98.72	0.25%
EuroStoxx50	49	728	0.5	0.1	99.14	99.64	0.51%	99.16	99.63	0.48%	99.14	99.61	0.47%	98.68	100.15	1.49%
EuroStoxx50	49	728	0.75	0.005	98.51	98.51	1.10s	98.51	98.51	4.62s	98.51	98.51	4.66s	98.51	98.51	4.60s
EuroStoxx50	49	728	0.75	0.01	98.84	98.84	1.63s	98.84	98.84	3.81s	98.84	98.84	3.70s	98.84	98.84	6.21s
EuroStoxx50	49	728	0.75	0.05	99.44	99.44	1623.69s	99.44	99.44	5184.55s	99.44	99.48	0.03%	99.34	99.53	0.19%
EuroStoxx50	49	728	0.75	0.1	99.70	99.95	0.25%	99.71	99.88	0.18%	99.70	99.90	0.19%	99.50	100.06	0.56%
av^{EURO}					97.49	97.67	11 0.18%	97.49	97.67	9 0.18%	97.49	97.68	8 0.19%	97.31	97.84	8 0.55%
DJIA	28	1352	0	0.005	95.10	95.10	3.65s	95.10	95.10	19.21s	95.10	95.10	19.15s	95.10	95.10	15.52s
DJIA	28	1352	0	0.01	95.80	95.80	55.13s	95.80	95.80	51.63s	95.80	95.80	54.78s	95.80	95.80	89.20s
DJIA	28	1352	0	0.05	97.47	98.51	1.07%	97.44	98.49	1.08%	97.44	98.49	1.08%	96.63	99.75	3.23%
DJIA	28	1352	0	0.1	98.07	99.84	1.81%	98.12	99.53	1.44%	98.11	99.49	1.40%	97.42	101.05	3.73%
DJIA	28	1352	0.25	0.005	96.38	96.38	3.69s	96.38	96.38	20.20s	96.38	96.38	20.28s	96.38	96.38	15.61s
DJIA	28	1352	0.25	0.01	96.91	96.91	56.67s	96.91	96.91	58.44s	96.91	96.91	52.04s	96.91	96.91	86.69s
DJIA	28	1352	0.25	0.05	98.18	98.99	0.82%	98.18	98.95	0.78%	98.17	98.92	0.76%	97.73	99.71	2.03%
DJIA	28	1352	0.25	0.1	98.63	99.90	1.29%	98.64	99.71	1.09%	98.64	99.71	1.09%	98.04	100.91	2.93%
DJIA	28	1352	0.5	0.005	97.67	97.67	3.12s	97.67	97.67	21.14s	97.67	97.67	21.21s	97.67	97.67	15.81s
DJIA	28	1352	0.5	0.01	98.02	98.02	59.04s	98.02	98.02	54.79s	98.02	98.02	58.00s	98.02	98.02	67.28s
DJIA	28	1352	0.5	0.05	98.84	99.44	0.61%	98.83	99.37	0.55%	98.86	99.37	0.51%	98.42	100.27	1.89%
DJIA	28	1352	0.5	0.1	99.16	100.00	0.84%	99.19	99.89	0.71%	99.17	99.91	0.74%	98.90	100.74	1.86%
DJIA	28	1352	0.75	0.005	98.96	98.96	2.19s	98.96	98.96	19.21s	98.96	98.96	19.41s	98.96	98.96	17.39s
DJIA	28	1352	0.75	0.01	99.13	99.13	54.76s	99.13	99.13	64.15s	99.13	99.13	65.87s	99.13	99.13	101.70s
DJIA	28	1352	0.75	0.05	99.56	99.85	0.29%	99.55	99.83	0.28%	99.57	99.83	0.27%	99.38	100.22	0.85%
DJIA	28	1352	0.75	0.1	99.72	100.17	0.46%	99.73	100.14	0.41%	99.72	100.11	0.39%	99.57	100.50	0.93%
av^{DJIA}					97.98	98.42	8 0.45%	97.98	98.37	8 0.40%	97.98	98.36	8 0.39%	97.75	98.82	8 1.09%
av^{all}					97.73	98.04	19 0.32%	97.73	98.02	17 0.29%	97.73	98.02	16 0.29%	97.53	98.33	16 0.82%

TABLE 7. Average upper bound and integrality gap improvements (in percentage) on the real-world data restricted to instances with $\tau \in \{0.05, 0.1\}$

	Comparison	avg. v_{UB} impr. (%)	avg. gap impr. (%)
$\tau \in \{0.05, 0.1\}$	MILP vs. MILP_{VI}	0.14	15.04
	MILP vs. $\text{MILP}_{\text{VI}^*}$	0.14	17.00
	MILP_{VI} vs. $\text{MILP}_{\text{VI}^*}$	0.00	2.09

We close this section with some final remarks about the impact of the valid inequalities (24) and (25). To do so, we compare the results obtained by configurations MILP, MILP_{VI}, and $\text{MILP}_{\text{VI}^*}$ on the 29 instances for which all the methods fail in proving optimality, i.e., instances characterized by a VaR confidence level equal to 5 % and 10 % minus the EUROSTOXX50 instances with $\tau = 0.05$ and $\alpha \in \{0.25, 0.5, 0.75\}$. We notice that the valid inequalities help in improving the upper bound and gap also on these instances. However, such improvements are smaller compared to the ones obtained for the synthetic instances. Indeed, as can be seen in Table 7, the upper bound improves on average about 0.14 % when comparing MILP with both MILP_{VI} and $\text{MILP}_{\text{VI}^*}$. The gap improves on average about 15 % and 17 % when comparing MILP with MILP_{VI} and $\text{MILP}_{\text{VI}^*}$, respectively. The dominance relation between Inequalities (25) and (24) is not clearly reflected in the numerical results. Indeed, the average improvement of the upper bound of configuration $\text{MILP}_{\text{VI}^*}$ vs. MILP_{VI} is negligible also on this test set and we report an average gap improvement about 2.09 % when comparing configuration $\text{MILP}_{\text{VI}^*}$ vs. MILP_{VI}.

To conclude, we observe that while valid inequalities may not be used to tackle small instances, they should be considered when dealing with large instances. The improvements they provide are indeed meaningful if considered in relative terms.

7. CONCLUSION

In this paper we considered several solution techniques for mixed-integer quantile minimization problems. We stated the problem in a very general form and developed techniques to strengthen the dual bound (via tailored valid inequalities), to find good primal solutions quickly (via the overlapping ADM), and to derive provably optimal solutions using a problem-specific approach (via the adaptive clustering method). Our numerical results on the maintenance planning problem of the ROADEF/EURO challenge 2020 and on the quantile-based version of the portfolio optimization problem show that the combination of these techniques significantly outperforms the application of general-purpose MILP solvers.

We briefly touched the field of chance constraints that is highly related to the quantile minimization problems discussed in this paper. Thus, a natural topic of future research will be to investigate on how to transfer our novel techniques to improve solution methods for chance-constrained problems. Moreover, other techniques from stochastic integer programming such as tailored decomposition methods could also be tried to be applied to the setting discussed in this paper in future work.

ACKNOWLEDGMENTS

Marius Roland and Martin Schmidt acknowledge the support by the German Bundesministerium für Bildung und Forschung within the project “EiFer”. Martin Schmidt thanks the DFG for their support within the projects A05 and B08 in CRC TRR 154.

References

[1] G. J. Alexander and A. M. Baptista. "A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model." In: *Management science* 50.9 (2004), pp. 1261–1273.

[2] G. J. Alexander and A. M. Baptista. "Economic implications of using a mean-VaR model for portfolio selection: A comparison with mean-variance analysis." In: *Journal of Economic Dynamics and Control* 26.7-8 (2002), pp. 1159–1193.

[3] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. "Coherent measures of risk." In: *Mathematical Finance* 9.3 (1999), pp. 203–228.

[4] S. Benati and R. Rizzi. "A mixed integer linear programming formulation of the optimal mean/Value-at-Risk portfolio problem." In: *European Journal of Operational Research* 176 (2007), pp. 423–434. DOI: 10.1016/j.ejor.2005.07.020.

[5] S. Boyd, N. Parikh, and E. Chu. *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc., 2011.

[6] P. F. Fischer. "An Overlapping Schwarz Method for Spectral Element Solution of the Incompressible Navier–Stokes Equations." In: *Journal of Computational Physics* 133.1 (1997), pp. 84–101. DOI: 10.1006/jcph.1997.5651.

[7] D. Gabay and B. Mercier. "A dual algorithm for the solution of nonlinear variational problems via finite element approximation." In: *Computers & Mathematics with Applications* 2.1 (1976), pp. 17–40.

[8] A. A. Gaivoronski and G. Pflug. "Finding optimal portfolios with constraints on value at risk." In: *Proceedings III Stockholm seminar on risk behavior and risk management*. 1999.

[9] A. A. Gaivoronski and G. Pflug. "Value-at-Risk in Portfolio Optimization: Properties and Computational Approach." In: *Journal of Risk* 7.2 (2005), pp. 1–31.

[10] M. R. Garey and D. S. Johnson. *Computers and intractability*. Vol. 174. San Francisco: freeman, 1979.

[11] N. Gatzert and H. Wesker. "A comparative assessment of Basel II/III and Solvency II." In: *The Geneva Papers on Risk and Insurance-Issues and Practice* 37.3 (2012), pp. 539–570.

[12] B. Geißler, A. Morsi, L. Schewe, and M. Schmidt. "Solving power-constrained gas transportation problems using an MIP-based alternating direction method." In: *Computers & Chemical Engineering* 82 (2015), pp. 303–317.

[13] B. Geißler, A. Morsi, L. Schewe, and M. Schmidt. "Solving Highly Detailed Gas Transport MINLPs: Block Separability and Penalty Alternating Direction Methods." In: *INFORMS Journal on Computing* 30.2 (2018), pp. 309–323. DOI: 10.1287/ijoc.2017.0780.

[14] R. Glowinski and A. Marroco. "Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires." In: *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique* 9.R2 (1975), pp. 41–76.

[15] T. Kleinert, M. Labbé, F. Plein, and M. Schmidt. "Closing the Gap in Linear Bilevel Optimization: A New Valid Primal-Dual Inequality." In: *Optimization Letters* 15 (2021), pp. 1027–1040. DOI: 10.1007/s11590-020-01660-6.

[16] T. Kleinert and M. Schmidt. "Computing Feasible Points of Bilevel Problems with a Penalty Alternating Direction Method." In: *INFORMS Journal on Computing* (2019). DOI: 10.1287/ijoc.2019.0945. Forthcoming.

[17] C.-C. Lin. "Comments on "A mixed integer linear programming formulation of the optimal mean/Value-at-Risk portfolio problem"." In: *European Journal of Operational Research* 194 (2009), pp. 339–341.

[18] R. Liu, Z. Lin, and Z. Su. "Linearized alternating direction method with parallel splitting and adaptive penalty for separable convex programs in machine learning." In: *Asian Conference on Machine Learning*. PMLR. 2013, pp. 116–132.

[19] R. Mansini, W. Ogryczak, and M. G. Speranza. "LP solvable models for portfolio optimization: A classification and computational comparison." In: *IMA Journal of Management Mathematics* 14.3 (2003), pp. 187–220.

[20] H. Markowitz. "Portfolio Selection." In: *The Journal of Finance* 7.1 (1952), pp. 77–91. DOI: 10.1111/j.1540-6261.1952.tb01525.x.

[21] S. Na, S. Shin, M. Anitescu, and V. M. Zavala. *On the Convergence of Overlapping Schwarz Decomposition for Nonlinear Optimal Control*. 2021.

[22] J. Puerto, F. Ricca, M. Rodríguez-Madrena, and A. Scozzari. "A combinatorial optimization approach to scenario filtering in portfolio selection." In: *Computers & Operations Research* 142 (2022), Paper No. 105701, 14. DOI: 10.1016/j.cor.2022.105701.

[23] J. Puerto, M. Rodríguez-Madrena, and A. Scozzari. "Clustering and portfolio selection problems: A unified framework." In: *Computers & Operations Research* 117 (2020), p. 104891. DOI: https://doi.org/10.1016/j.cor.2020.104891.

[24] F. Qiu, S. Ahmed, S. S. Dey, and L. A. Wolsey. "Covering linear programming with violations." In: *INFORMS Journal on Computing* 26.3 (2014), pp. 531–546.

[25] L. Schewe, M. Schmidt, and D. Weninger. "A decomposition heuristic for mixed-integer supply chain problem." In: *Operations Research Letters* 48.3 (2020), pp. 225–232. DOI: 10.1016/j.orl.2020.02.006.

[26] S. Shin, V. M. Zavala, and M. Anitescu. "Decentralized Schemes With Overlap for Solving Graph-Structured Optimization Problems." In: *IEEE Transactions on Control of Network Systems* 7.3 (2020), pp. 1225–1236. DOI: 10.1109/TCNS.2020.2967805.

[27] B. W. Silverman. *Density estimation for statistics and data analysis*. Routledge, 2018. DOI: 10.1201/9781315140919.

[28] Y. Song and J. R. Luedtke. "Branch-and-cut approaches for chance-constrained formulations of reliable network design problems." In: *Mathematical Programming Computation* 5.4 (2013), pp. 397–432.

[29] Y. Song, J. R. Luedtke, and S. Kücükyavuz. "Chance-constrained binary packing problems." In: *INFORMS Journal on Computing* 26.4 (2014), pp. 735–747.

[30] M. W. Tanner and L. Ntaimo. "IIS branch-and-cut for joint chance-constrained stochastic programs and application to optimal vaccine allocation." In: *European Journal of Operational Research* 207.1 (2010), pp. 290–296.

Appendix A. Detailed Results for the Instances of the ROADEF/EURO Challenge 2020

In Tables 8 and 9, lines labelled with av^A, av^B, av^C and av^X report the average gap(%) returned by the corresponding method on the corresponding group of instances. Please note that for the computation of the averages we only consider the instances for which all the methods were able to provide a valid lower and upper bound on the optimal value. Thus, we exclude from the computation of the average all the instances for which at least one of the methods fails to return a gap.

TABLE 8. Results for all methods applied to the A and B instances of the ROADEF/EURO challenge 2020.

	MILP			$\text{MILP}_{\text{VI}^*}$			$\text{MILP}_{\text{VI}^*}^{\text{OADM}}$				ASCA		
ID	v_{UB}	v_{LB}	gap (%)	v_{UB}	v_{LB}	gap (%)	$v_{\text{UB}}^{\text{OADM}}$	v_{UB}	v_{LB}	gap (%)	v_{UB}	v_{LB}	gap (%)
A02	4676.15	2083.50	55.44	4675.89	4608.65	1.44	7991.34	**4673.71**	4595.80	1.67	4678.91	**4612.49**	**1.42**
A05	648.18	593.72	8.40	638.16	**615.90**	3.49	659.91	**637.11**	615.79	**3.35**	644.94	603.74	6.39
A08	744.67	717.19	3.69	744.42	**739.97**	**0.60**	746.55	**744.29**	737.13	0.96	744.29	725.92	2.47
A11	497.91	459.75	7.66	496.19	477.31	3.80	518.34	**495.41**	**477.74**	**3.57**	497.73	470.84	5.40
A14	2271.19	2127.77	6.31	2271.65	2157.24	5.04	2291.29	**2268.71**	**2162.90**	**4.66**	2287.78	2145.27	6.23
A15	2297.10	2096.75	8.72	**2277.59**	2159.78	**5.17**	2317.34	2278.21	**2160.11**	5.18	2296.29	2122.53	7.57
av^{A}			15.04			3.26				3.23			4.91
B01	4451.12	1907.40	57.15	3992.93	3834.47	3.97	3999.52	**3986.20**	**3835.03**	**3.79**	3986.20	3830.89	3.90
B02	4859.51	2054.48	57.72	4359.30	**3680.72**	15.57	4411.48	4350.37	3680.17	**15.41**	**4330.27**	3641.75	15.90
B03	38771.13	11874.80	69.37	35296.93	33985.29	**3.72**	35330.74	35312.20	**33994.73**	3.73	**35293.58**	33889.08	3.98
B04	36945.42	12416.39	66.39	34834.64	**33635.76**	**3.44**	34939.14	**34831.98**	33599.68	3.54	34836.88	33564.50	3.65
B05	2637.10	1335.18	49.37	**2397.39**	**2320.31**	**3.22**	2474.54	2398.49	2318.00	3.36	2398.83	2319.43	3.31
B06	5015.20	2013.19	59.86	4340.10	3684.19	15.11	4349.30	4327.98	**3684.50**	**14.87**	**4309.50**	3601.40	16.43
B07	8195.35	4399.00	46.32	7796.01	6134.67	21.31	7723.64	7718.25	**6134.70**	20.52	**7567.91**	6132.68	**18.96**
B08	7691.52	2364.37	69.26	7436.24	7288.90	1.98	7440.06	7436.27	**7296.56**	**1.88**	**7435.99**	7187.81	3.34
B09	7502.89	2625.93	65.00	7496.88	7283.08	2.85	7847.58	7496.88	**7287.07**	**2.80**	**7494.42**	7267.69	3.03
B10	11048.19	5746.89	47.98	10844.68	8472.20	21.88	11005.20	10977.73	**8472.26**	22.82	**10616.97**	8461.01	**20.31**
B11	3963.75	1839.02	53.60	3660.97	**3188.13**	**12.92**	3717.19	3713.06	3187.54	14.15	**3641.00**	3131.08	14.00
B12	39140.68	15155.39	61.28	**37601.56**	36725.67	**2.33**	37667.20	37603.45	**36727.96**	2.33	37641.14	36656.72	2.62
B13	5414.18	2342.38	56.74	**5025.02**	**4844.50**	**3.59**	5030.40	5030.21	4837.91	3.82	5025.08	4837.25	3.74
B14	12207.73	7426.41	39.17	12028.83	**9615.46**	**20.06**	12241.84	12239.41	9615.44	21.44	**11946.94**	9496.75	20.51
B15	23652.51	9877.52	58.24	**22568.23**	**21309.97**	**5.58**	23812.01	22570.58	21309.77	5.59	23431.63	17801.63	24.03
av^{B}			57.16			9.17				9.34			10.51

TABLE 9. Results for all methods applied to the C and X instances of the ROADEF/EURO challenge 2020.

	MILP			$\text{MILP}_{\text{VI}^*}$			$\text{MILP}_{\text{VI}^*}^{\text{OADM}}$				ASCA		
ID	v_{UB}	v_{LB}	gap (%)	v_{UB}	v_{LB}	gap (%)	$v_{\text{UB}}^{\text{OADM}}$	v_{UB}	v_{LB}	gap (%)	v_{UB}	v_{LB}	gap (%)
C01	8754.26	2906.19	66.80	8521.05	8236.61	3.34	8540.41	8528.23	**8256.12**	**3.19**	**8515.90**	8219.94	3.48
C02	3669.06	2211.57	39.72	3582.18	**2969.22**	**17.11**	3658.81	3654.43	2969.12	18.75	**3552.86**	2931.82	17.48
C03	36781.76	11557.87	68.58	33525.65	31826.83	5.07	33588.69	33533.84	**31857.06**	**5.00**	**33518.56**	31767.32	5.22
C04	38115.53	12467.93	67.29	**37596.26**	**36366.72**	**3.27**	37664.53	37600.49	36352.53	3.32	37602.92	36238.94	3.63
C05	3860.26	1066.27	72.38	3168.34	3083.32	2.68	3186.90	3168.93	3080.38	2.79	**3104.70**	**3087.10**	**0.57**
C06	8889.88	4583.96	48.44	8547.15	**6739.07**	21.15	8778.44	8762.82	6739.06	23.09	**8447.05**	6700.88	**20.67**
C07	6102.69	3008.25	50.71	6090.10	5903.59	3.06	6293.30	6091.73	5903.97	3.08	**6085.66**	**5908.25**	**2.92**
C08	11812.39	6493.37	45.03	11394.01	**9180.80**	19.42	11618.12	11377.87	9180.74	19.31	**11224.33**	9171.45	**18.29**
C09	6391.26	2136.59	66.57	5676.04	**4379.36**	**22.84**	5767.11	5744.49	4378.94	23.77	**5638.02**	4245.87	24.69
C10	47781.75	14965.24	68.68	43345.46	**42034.45**	**3.02**	43437.28	**43344.44**	42026.25	3.04	43391.98	41964.03	3.29
C11	6135.58	2481.49	59.56	5751.22	**5504.78**	**4.29**	5757.33	5757.33	5500.69	4.46	**5750.90**	5492.30	4.50
C12	13176.78	6937.76	47.35	12904.93	**9389.22**	27.24	13179.90	13125.88	9389.19	28.47	**12800.16**	9328.34	**27.12**
C13	44036.01	15866.48	63.97	-	**41391.28**	-	42570.54	**42542.09**	41390.50	**2.71**	43496.46	-	-
C14	-	10699.95	-	28200.33	**17639.45**	37.45	28313.00	27848.40	17639.38	36.66	**27115.66**	17228.93	**36.46**
C15	41171.04	15164.85	63.17	-	**38782.32**	-	39934.64	**39900.47**	38778.58	**2.81**	41945.39	-	-
av^{C}			58.43			11.04				11.52			10.99
X01	4411.05	2068.03	53.12	4308.46	2870.76	33.37	4530.92	4514.52	**2870.80**	36.41	**4028.32**	2844.47	**29.39**
X02	35730.18	12267.71	65.67	32261.49	30558.33	5.28	32308.35	32259.72	**30613.06**	**5.10**	**32234.14**	30511.03	5.35
X03	8421.26	4391.73	47.85	8287.66	6487.80	21.72	8425.14	8416.38	**6488.21**	22.91	**8138.92**	6427.76	**21.02**
X04	12113.01	5963.90	50.76	11570.88	**9011.73**	22.12	11640.98	11609.69	9011.53	22.38	**11362.25**	9003.65	**20.76**
X05	-	10098.40	-	-	**17707.11**	-	23644.29	23640.86	17706.67	**25.10**	**23334.07**	16272.12	30.26
X06	47350.09	19324.99	59.19	**47044.52**	**45939.42**	**2.35**	49452.91	47473.88	45938.93	3.23	48971.01	25811.71	47.29
X07	14205.28	5610.97	60.50	**13227.00**	**12713.27**	**3.88**	13228.65	13227.88	12713.27	3.89	13293.00	12419.29	6.57
X08	16825.52	5349.20	68.21	-	9905.30	-	14070.05	14059.22	**9905.47**	**29.54**	**13933.08**	9781.10	29.80
X09	22025.67	8745.96	60.29	21580.38	12894.90	40.25	20791.99	20760.96	**12895.06**	37.89	**20281.50**	12885.20	**36.47**
X10	-	7515.01	-	17426.28	15778.72	**9.45**	17608.25	17513.64	**15781.16**	9.89	**17416.37**	15684.45	9.94
X11	-	14642.87	-	39131.74	38112.94	**2.60**	39150.58	39138.44	**38118.22**	2.61	**39128.22**	37990.10	2.91
X12	-	14724.24	-	**48008.97**	**37542.86**	**21.80**	54145.99	48130.56	37531.09	22.02	48047.70	37324.54	22.32
X13	16445.38	8812.37	46.41	15921.75	**14044.20**	**11.79**	16408.89	16404.04	14043.90	14.39	**15852.38**	13972.27	11.86
X14	83601.02	21153.47	74.70	**79446.98**	**76589.60**	**3.60**	79558.65	79470.78	76589.45	3.63	80737.89	73062.59	9.51
X15	50556.76	16831.18	66.71	**45793.90**	**38780.60**	**15.31**	50013.28	49461.03	38779.93	21.59	46384.15	-	-
av^{X}			57.61			16.04				16.65			20.91

Appendix B. Summary Statistics for the ROADEF/EURO Challenge 2020 Instances

Table 10. Average improvement of v_{LB}, v_{UB}, and gap. Each row compares one method $m \in \mathcal{M}$ and displays, from left to right, the average over all instances of $(v_{\mathrm{UB}}^{\mathrm{MILP}} - v_{\mathrm{UB}}^{m})/v_{\mathrm{UB}}^{\mathrm{MILP}}$, $(v_{\mathrm{LB}}^{m} - v_{\mathrm{LB}}^{\mathrm{MILP}})/v_{\mathrm{LB}}^{\mathrm{MILP}}$, and $(\mathrm{gap}^{\mathrm{MILP}} - \mathrm{gap}^{m})/\mathrm{gap}^{\mathrm{MILP}}$. All instances for which one method yields an unbounded value for v_{LB} or v_{UB} are excluded from the computation.

Method	avg. v_{UB} impr.(%)	avg. v_{LB} impr.(%)	avg. gap impr.(%)
$\mathrm{MILP}_{\mathrm{VI}^*}$	4.78	97.43	75.08
$\mathrm{MILP}_{\mathrm{VI}^*}^{\mathrm{OADM}}$	4.39	97.43	74.38
ASCA	5.19	92.58	69.02

Table 11. Number of instances for which v_{LB} and v_{UB} are bounded. Each row corresponds to a method and reports the amount of instances for which this method results in a bounded LB or a bounded UB after the 90 min time limit. There are 51 ROADEF/EURO instances in total.

Method	UB	LB
MILP	46	51
$\mathrm{MILP}_{\mathrm{VI}^*}$	47	51
$\mathrm{MILP}_{\mathrm{VI}^*}^{\mathrm{OADM}}$	51	51
ASCA	51	48

(D. Cattaruzza) Univ. Lille, CNRS, Centrale Lille, Inria, UMR 9189 - CRIStAL Lille, France

Email address: diego.cattaruzza@centralelille.fr

(M. Labbé) (A) Université Libre de Bruxelles, Department of Computer Science, Boulevard du Triomphe, CP212, 1050 Brussels, Belgium; (B) Inria Lille - Nord Europe, Parc scientifique de la Haute Borne, 40, av. Halley - Bât A - Park Plaza, 59650 Villeneuve d'Ascq, France

Email address: martine.labbe@ulb.be

(M. Petris) Univ. Lille, Inria, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France

Email address: matteo.petris@inria.fr

(M. Roland, M. Schmidt) Trier University, Department of Mathematics, Universitätsring 15, 54296 Trier, Germany

Email address: roland@uni-trier.de, martin.schmidt@uni-trier.de