where $x \in \mathbb{R}^n$ is the decision variable, and $f_0, \ldots, f_m$ define the objective and constraints. In order to develop a robust counterpart, it is assumed that these functions depend on a set of extra parameters, encoded in a vector $u \in \mathbb{R}^p$; we use the notation $f_i(x, u)$ to reflect this dependence. Further, we assume that this vector is only partially known, $u \in \mathcal{U}$ where $\mathcal{U}$ is a given subset of $\mathbb{R}^p$. The robust counterpart writes

$$\min_x \max_{u \in \mathcal{U}} f_0(x, u) \ : \ \forall\, u \in \mathcal{U}, \ \ f_i(x, u) \leq 0, \ \ i = 1, \ldots, m.$$

Note that the problem can be equivalently represented as

$$\min_x \max_{u \in \mathcal{U}} f_0(x, u) \ : \ \max_{u \in \mathcal{U}} f_i(x, u) \leq 0, \ \ i = 1, \ldots, m.$$

Hence, the problem's complexity depends on our ability to efficiently compute and represent functions of the form

$$\max_{u \in \mathcal{U}} f_i(x, u).$$

There are a few cases where the robust counterpart is tractable.

3.2 A few tractable cases

To illustrate, consider a linear program (LP) of the form

$$\min_x c^T x \ : \ a_i^T x \leq b_i, \ \ i = 1, \ldots, m.$$

Now assume that the cost vector c is only known to belong to a given set $\mathcal{U}$. The robust counterpart writes

$$\min_x \max_{c \in \mathcal{U}} c^T x \ : \ a_i^T x \leq b_i, \ \ i = 1, \ldots, m.$$

The new objective function, referred to as the worst-case objective function:

$$x \rightarrow \max_{c \in \mathcal{U}} c^T x$$

can be efficiently computed for a large number of sets $\mathcal{U}$. For example, if $\mathcal{U} = \{c^{(j)}, \ j = 1, \ldots, K\}$ is finite, where each $c^{(j)} \in \mathbb{R}^n$ is given and corresponds to a *scenario*, the worst-case objective function is polyhedral, and the robust counterpart can be solved as an LP:

$$\min_x \max_{j=1,\ldots,K} (c^{(j)})^T x \ : \ a_i^T x \leq b_i, \ \ i = 1, \ldots, m.$$

Another popular model involves element-wise bounds on the cost vector: with $\mathcal{U} = \{u \ : \ \|u - \hat{u}\|_\infty \leq \rho\}$, where $\hat{u} \in \mathbb{R}^n$ is a given "nominal" value and $\rho \geq 0$ is a measure of the component-wise uncertainty, the robust counterpart is also an LP:

$$\min_x \hat{c}^T x + \rho \|x\|_1 \ : \ a_i^T x \leq b_i, \ \ i = 1, \ldots, m.$$

The so-called ellipsoidal uncertainty model assumes that the coefficient vector a is only known to lie in a ellipse in $\mathbb{R}^n$. The ellipsoidal uncertainty model is

$$\mathcal{U} = \{a = \hat{a} + Ru \ : \ \|u\|_2 \leq 1\},$$

where R is a given matrix that determines the shape and orientation of the ellipse. Such a model can be constructed in a data-driven way, using likelihood region techniques [6]. The robust countepart writes as a second-order cone program (SOCP):

$$\min_x \hat{c}^T x + \|R^T x\|_2 \ : \ a_i^T x \leq b_i, \ \ i = 1, \ldots, m.$$

3.3 Affine recourse

In some decision problems subject to uncertainty, some of the parameters that are not fully known at optimization time are revealed when implementing the solution. This happens often for example when there is a time aspect to the problem, and the uncertainty is gradually revealed as time goes by. In the charging station context, the operator learns of the arrival of vehicles as they come into the station. It is possible to exploit this structure at optimization time, by implementing what amounts to a feedback loop between uncertainty and decision. Affine recourse is a methodology that implements this idea; see [16] for an example in the context of power generation.

To illustrate this approach, consider a nominal decision problem with two time steps:

$$\min_{x,y} c^T x \ : \ Ax + By \leq b$$

where x contains "here and now" (first stage) decision variables, and y contains "wait and see" (second stage) decision variables, which can depend on the uncertainty.

Now assume that the "recourse matrix" B is fixed, while A, b, c are subject to uncertainty: A, b, c depend affinely on an uncertain vector $u \in \mathcal{U}$, where the uncertainty set $\mathcal{U}$ is given. We denote this dependence via $A(u)$, $b(u)$, $c(u)$. An adujstable robust counterpart is

$$\min_{x,\, y(\cdot)} \max_{u \in \mathcal{U}} c(u)^T x \ : \ \forall\, u \in \mathcal{U} \ : \ A(u)x + By(u) \leq b(u)$$

Here the "wait and see" decision y is now a *function* of u. The above problem is generally intractable.

In the Affinely Adjustable Robust Counterpart (AARC) model, we force $y(u)$ to be an affine function of the form

$$y(u) = y + Yu,$$

where vector y and matrix Y are both variables. In practice, we may have to restrict the way y depends on u: y must depend only on the parts of u that are actually revealed before the y-decision has to be made. This often translates as linear constraints on the matrix Y; typically, if y, u are time-series then Y must be strictly lower triangular. We encode these constraints with a set $\mathcal{Y}$ of admissible recourse matrices.

The AARC writes

$$\min_{x,\, y,\, Y \in} \max_{u \in \mathcal{U}} c(u)^T x \ : \ \forall\, u \in \mathcal{U} \ : \ A(u)x + B(y + Yu) \leq b(u), \ \ Y \in \mathcal{Y}.$$

Since $A(u)$, $b(u)$, $c(u)$ are all affine in u, the above is amenable to standard robust counterparts, as discussed above.

There are tractable extensions to the case when the recourse matrix B is also subject to uncertainty; in general these more general problems are difficult, as they involve products of uncertain variables. One notable exception, which will prove relevant later, is the case of scenario models. When $u \in \mathcal{U} := \{u^{(k)}, k = 1, \ldots, K\}$, where $u^{(k)}$'s are scenarios, with u entirely revealed in the second stage, the AARC becomes

$$\min_{x,\, y,\, Y} \max_{1 \leq k \leq K} c(u^{(k)})^T x \ : \ A(u^{(k)})x + B(u^{(k)})(y + Yu^{(k)}) \leq b(u^{(k)}), \ \ k = 1, \ldots, K.$$

3.4 Distributional robustness

Some extensions of the basic examples above relate to *distributional robustness*, where the uncertainty is modeled as random, with a partially known distribution. To illustrate the concept, consider a single linear constraint on a decision variable $x \in \mathbb{R}^n$, of the form $a^T x \leq b$, where $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$. Assume that a is a random Gaussian variable with a distribution denoted π, having mean $\hat{a}$ and covariance matrix $C \succeq 0$; factorize the latter as $C = RR^T$ where $R \in \mathbb{R}^{n \times n}$. To cope with this random uncertainty, we consider a *chance constraint*, of the form

$$\mathbf{Prob}_\pi\{a \ : \ a^T x \leq b\} \geq 1 - \epsilon.$$

where $\epsilon < 0.5$ is a given, required probability level, and where $\mathbf{Prob}_\pi$ refers to the probability of an event under the Gaussian distribution denoted π. It turns out that the corresponding constraint is a second-order cone constraint:

$$\hat{a}^T x + \kappa \|R^T x\|_2 \leq 0. \tag{1}$$

where $\kappa := \Phi(\epsilon) > 0$ is a tabulated function of the required reliability level.

In distributional robustness, we relax the assumption that the distribution on a is fully known. Instead, we may assume for example that the mean and covariance matrix are known, but otherwise the distribution is arbitrary, in particular, not necessarily Gaussian. Let $\mathcal{P}$ be the class of distributions in which a's distribution lies. The corresponding distributionally robust counterpart is

$$\min_{\pi \in \mathcal{P}} \mathbf{Prob}_\pi \{a \ : \ a^T x \leq b\} \geq 1 - \epsilon.$$

In words: we are requiring the constraint to hold with high probability, irrespective of the distribution π of a in set $\mathcal{P}$. In our specific case, it turns out that the above constraint can be written as (1), where now

$$\kappa = \sqrt{\frac{1-\epsilon}{\epsilon}},$$

which is of course always higher than the value used in the Gaussian case. As expected, removing the Gaussian assumption makes the robust counterpart more demanding in terms of robustness.

Ditributional robustness offers a way to deal with scenarios in a less conservative manner than that described in Section 3.2. Consider for example the case of a single scalar constraint on a decision vector $x \in \mathbb{R}^n$, of the form $a^T x \leq b$, where $b \in \mathbb{R}$ is known, and $a \in \mathbb{R}^n$ is a random variable that obeys to a known distribution $p \in \mathbb{R}^K$. Specifically, we assume that a takes the given value $a^{(k)}$ with probability p_k, $k = 1, \ldots, K$. If p is known, we can enforce the constraint in average, leading to a simple linear inequality:

$$b \geq \mathbf{E}_p(a^T x) = \frac{1}{K} \sum_{k=1}^{K} p_k a^{(k)}.$$

We now assume that the vector p is only known to belong to a given subset $\mathcal{P}$ of the probability simplex. The robust counterpart writes

$$b \geq \max_{p \in \mathcal{P}} \mathbf{E}_p(a^T x).$$

If p is completely unknown, that is, the set $\mathcal{P}$ of admissible distributions is the entire probability simplex, we recover the robust counterpart of the scenario model described in Section 3.2:

$$b \geq \max_{p \geq 0, \, \mathbf{1}^T p = 1} \mathbf{E}_p(a^T x) = \max_{1 \leq k \leq K} (a^{(k)})^T x.$$

There are less conservative ways to model uncertainty on probability distributions. For example, assume that we know a reference probability distribution q; we consider the set

$$\mathcal{P} := \left\{ p \geq 0 \ : \ \mathbf{1}^T p = 1, \ \ D(p||q) := \sum_{k=1}^{K} p_k \log(\frac{p_k}{q_k}) \leq \gamma \right\}, \tag{2}$$

where $\gamma > 0$ is a given number; here, $D(p||q)$ is a relative entropy (Kullback-Leibler distance) metric. The above set is not trivially equal to the whole probability simplex only when $\gamma < \max_k \log(1/q_k)$, which we assume. The robust counterpart writes

$$b \geq \max_{p \in \mathcal{P}} \sum_{k=1}^{K} p_k (a^{(k)})^T x.$$

Thanks to duality, it is possible to rewrite the above constraint in a tractable form. Letting $w_i := (a^{(i)})^T x$, $i = 1, \ldots, K$, we have

$$\max_{p \in \mathcal{P}} \; p^T w = \min_{\lambda,\nu} \; g(\lambda,\nu), \text{ where } g(\lambda,\nu) := \max_{p \geq 0} \; p^T w + \nu(1 - \mathbf{1}^T p) + \lambda(\gamma - D(p||q)).$$

As seen in [9], we can express the dual problem with a single variable:

$$\max_{p \in \mathcal{P}} \; p^T w = \min_{\lambda \geq 0} \; \gamma\lambda + \lambda \log\left(\sum_{k=1}^{K} q_k \exp(\frac{w_k}{\lambda})\right).$$

We obtain that $p^T w \leq b$ for every $p \in \mathcal{P}$ if and only if there exist $\lambda \geq 0$ such that

$$\gamma\lambda + \lambda \log\left(\sum_{k=1}^{K} q_k \exp(\frac{(a^{(k)})^T x}{\lambda})\right) \leq b.$$

The above is jointly convex in x, λ.

4 Aggregated Model

In [14], the authors proposed a simplified model in which the aspect of the allocation of vehicles to the different charging stations is ignored, with a focus on an efficient online algorithm for the optimal charging of vehicles over time. The online algorithm is based on an offline model that is expressed as a convex quadratic program. In this section, we adapt the offline model to the case when there are uncertainties in the electricity costs and arrival and departure times. In practice, the resulting model could be applied in an online fashion, based on recomputing the whole problem at time step, in a "rolling horizon" mode.

4.1 Nominal model

The proposed model for N vehicles over a time horizon of T steps takes the form

$$p^* := \min_{Y} \sum_{t=1}^{T} f_t\left(y_t^0 + \sum_{i \,:\, t \in [a_i, d_i]} Y_{ti}\right) \;:\; Y \geq 0, \; \sum_{t=a_i}^{d_i} Y_{ti} = L_i, \; i = 1, \ldots, N. \tag{3}$$

In the above, $f_t : \mathbb{R} \to \mathbb{R}$, $t = 1, \ldots, T$ are convex increasing functions that encode the cost of production at time t; y_0 is a given T-vector that corresponds to a non-flexible consumption of the charging station; variable Y is a $T \times N$ matrix, such that Y_{ti} is the power allocated to vehicle i at time t; $L \in \mathbb{R}^N$ is a (required) load vector divided by a time step parameter; finally, for each i, (a_i, d_i) are (given) arrival and departure times for vehicle i.

Let us encode the arrival and departure times in a $T \times N$ binary matrix A, where $A_{ti} = 1$ if t is between the arrival and departure times of vehicle i, and zero otherwise. With this convention, we have, for every $N \times T$ matrix Y:

$$\forall\, i \;:\; \sum_{t=a_i}^{d_i} Y_{ti} = \sum_{t=1}^{T} A_{ti} Y_{ti} = (Ae_i)^T Y e_i = e_i^T A^T Y e_i, \;\; \forall\, t \;:\; \sum_{i \,:\, t \in [a_i, d_i]} Y_{ti} = \sum_{i=1}^{N} A_{ti} Y_{ti} = e_t^T A Y^T e_t,$$

This leads to an expression of the problem in matrix form:

$$p^* = \min_{Y} \; F(y^0 + \mathbf{diag}(AY^T)) \;:\; \mathbf{diag}(A^T Y) \geq L, \;\; Y \geq 0, \tag{4}$$

where $F : \mathbb{R}^T \to \mathbb{R}$ is the function with values for $z \in \mathbb{R}^T$ given by

$$F(z) := \sum_{t=1}^{T} f_t(z_t).$$

Note that we have replaced the equality load constraint in (3) by an inequality; this is done without loss of generality, since the costs are increasing with increasing loads. The interpretation is that it is not optimal to produce in excess of the required load.

Sohet's thesis [13] discusses several choices for the production costs, for example (constant over time) quadratic costs, where $f_t(\xi) = \xi^2$ for every t. We may also consider time-varying linear costs: $f_t(\xi) = c_t\xi$, where $c_t > 0$ for every t. In that case, the problem writes

$$p^* = \min_{Y} \; c^T(y^0 + \mathbf{diag}(AY^T)) \;:\; \mathbf{diag}(A^TY) \geq L, \;\; Y \geq 0. \tag{5}$$

For later reference, we express the above in terms of the columns of A, Y, denoted a_i, y_i, $i = 1, \ldots, N$ respectively. We have

$$c^T \mathbf{diag}(AY^T) = c^T \mathbf{diag}\left(\sum_{i=1}^{N} a_i y_i^T\right) = c^T \left(\sum_{i=1}^{N} \mathbf{diag}(a_i y_i^T)\right) = \sum_{i=1}^{N} c^T(a_i \circ y_i) = \sum_{i=1}^{N} (c \circ y_i)^T a_i. \tag{6}$$

Problem (7) can be written as

$$p^* = \min_{y_1,\ldots,y_N} \; \sum_{i=1}^{N} (c \circ y_i)^T a_i \;:\; a_i^T y_i \geq L_i, \;\; y_i \geq 0, \;\; i = 1, \ldots, N. \tag{7}$$

It turns out that in the linear cost case, the optimal loading matrix corresponds to a somewhat trivial solution, which is to select times such that the cost is the smallest, and put all the charging load at that time step. This can be proven by taking the dual to the above; in contrast to the quadratic cost model, which requires the water-filling algorithm, the dual to the above has a closed-form solution.

To illustrate, consider a toy problem where $T = 4$, $N = 3$, $y^0 = 0$, and

$$c = \begin{pmatrix} 26 \\ 25 \\ 20 \\ 29 \end{pmatrix}, \;\; L = \begin{pmatrix} 12 \\ 17 \\ 19 \end{pmatrix}, \;\; A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{8}$$

The optimal load matrix and cost are then

$$Y_{\text{lin}}^* = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 19 \\ 12 & 17 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We observe that the loading matrix indeed corresponds to the simple strategy of charging "all at once" at the best (lowest cost) time step. In contrast, with a constant quadratic cost, the solution does distribute the load over the entire charging intervals:

$$Y_{\text{quad}}^* = \begin{pmatrix} 0 & 0 & 12.0000 \\ 5.0000 & 0 & 7.0000 \\ 7.0000 & 5.0000 & 0 \\ 0 & 12.0000 & 0 \end{pmatrix}.$$

4.2 Robust counterparts: uncertainty on production costs and loads

Uncertainty in the cost vector. As a first step, we address uncertainties in the costs. Precisely, we consider the nominal problem above, with linear, time-varying costs encoded in a cost vector $c \in \mathbb{R}^T$. Here, we assume that c is only known up to a sphere: $\|c - \hat{c}\|_2 \leq \rho$, where the "nominal" cost $\hat{c} \in \mathbb{R}^T$ and the uncertainty level $\rho \geq 0$ are both known. The robust counterpart to (7) is the SOCP

$$p^* = \min_Y \ \hat{c}^T(y^0 + \mathbf{diag}(AY^T)) + \rho\|y^0 + \mathbf{diag}(AY^T)\|_2 \ : \ \mathbf{diag}(A^TY) \geq L, \ \ Y \geq 0. \tag{9}$$

We observe that when $\rho = 0$, we recover the nominal problem with fixed costs; while for ρ large, the solution approaches that of the problem with quadratic costs. We can see that the robust model accomplishes a trade-off between the nominal cost and a term that corresponds to the case with quadratic costs. As a result, the robust model does not suffer from the drawback observed with linear costs, where charging systematically occurs at a single time step.

Returning to the toy example above, with $\hat{c}$ set as c in (8), and $\rho = 11$, we obtain indeed a better spread in charging values:

$$Y^*_{\text{rob}} = \begin{pmatrix} 0.0000 & 0.0000 & 9.3003 \\ 2.3005 & 0.0000 & 9.6997 \\ 9.6995 & 15.8011 & 0.0000 \\ 0.0000 & 1.1989 & 0.0000 \end{pmatrix}.$$

As expected, when $\rho \to \infty$, the solution corresponding to quadratic costs, Y^*_{quad}, is recovered.

In practice, since c represents a time-series, it would be advisable to implement a more sophisticated uncertainty model, for example one that is based on an ARMA models and likelihood regions [6].

Uncertainty in the load vector. In practice, due to uncertainty on the initial charge of vehicles for example, the load vector L is uncertain. Assume for example that the load vector L is only known to satisfy $L \in [\underline{L}, \overline{L}]$, where the lower and upper bounds $\underline{L}, \overline{L}$ are known. The robust counterpart has a trivial expression, in the same form as the nominal model (4), but with values of L replaced with their worst-case (largest) values, $\overline{L}$. If the infrastructure of the charging station allows to measure, at charging time, the actual desired load, then an affine recourse strategy may improve the performance. Assume now that L depends on an uncertain vector $u \in [0,1]^N$, precisely $L(u) = (\mathbf{1} - u) \circ \underline{L} + u \circ \overline{L}$. We then assume an affine dependence of the decision variable on u. Let us fix $i \in \{1, \ldots, N\}$. The i-th column of Y will be of the form $y_i(u) = \hat{y}_i + R_i u$, where $\hat{y}_i \in \mathbb{R}^T$, $R_i \in \mathbb{R}^{T \times N}$ are both decision variables.

We first enforce $y_i(u) \geq 0$ for every $u \in [0,1]^N$, leading to the convex constraint $\hat{y}_i \geq (-R_i)_+^T \mathbf{1}$, where $(\cdot)_+$ stands for the positive part of its matrix input. Next, the robust counterpart of the i-th load constraint takes the form

$$\forall\, u \in [0,1]^N \ : \ a_i^T(\hat{y}_i + R_i u) \geq (1 - u_i)\underline{L}_i + u_i \overline{L}_i.$$

The above is equivalent to the convex constraint:

$$a_i^T \hat{y}_i \geq \underline{L}_i + \mathbf{1}^T((\overline{L}_i - \underline{L}_i)e_i - R_i^T a_i)_+.$$

Treating the worst-case cost in this fashion is similar; it entails computing

$$\max_{u \in [0,1]^N} (c \circ a_i)^T y_i(u) = \hat{y}_i^T(c \circ a_i) + \mathbf{1}^T(R_i^T(c \circ a_i))_+.$$

The AARC model then writes

$$\begin{aligned} \min_{Y,(R_i)_{i=1}^T} \quad & \sum_{i=1}^N \left(\hat{y}_i^T(c \circ a_i) + \mathbf{1}^T(R_i^T(c \circ a_i))_+\right) \\ \text{s.t.} \quad & a_i^T \hat{y}_i \geq \underline{L}_i + \mathbf{1}^T((\overline{L}_i - \underline{L}_i)e_i - R_i^T a_i)_+, \ \ \hat{y}_i \geq (-R_i)_+^T \mathbf{1}, \ \ i = 1, \ldots, N. \end{aligned}$$

In order to reduce the number of additional variables involved, it may be advisable to enforce some structure on R_i's, for example $R_i = \mathbf{1} r_i^T$ for some vector $r_i \in \mathbb{R}^N$. This will restrict the performance of the AARC model, at the benefit of a reduced number of variables.

4.3 Robust counterpart: uncertainty on arrival and departure times

Scenario uncertainty. In order to cope with uncertainties in the arrival times, we first assume that the matrix A is only known to be long to a finite set $\mathcal{A}$ of possible scenarios of arrival / departure times, denoted $A^{(k)}$, $k = 1, \ldots, K$. The corresponding robust counterpart is

$$p^* = \min_{Y} \max_{A \in \mathcal{A}} F(y^0 + \mathbf{diag}(A^{(k)} Y^T)) \ : \ \mathbf{diag}((A^{(k)})^T Y) \geq L, \ \ k = 1, \ldots, K, \ \ Y \geq 0. \tag{10}$$

The problem is convex and can be solved using standard convex optimization software, for a wide variety of cost functions F.

Returning to the toy example above, assuming (known) linear costs, let us assume that there are two scenarios ($K = 2$), with

$$A^{(1)} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \ \ A^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

We obtain

$$Y = \begin{pmatrix} 0 & 0 & 0 \\ 0.0000 & 0 & 19.0000 \\ 12.0000 & 17.0000 & 17.6186 \\ 0 & 0.0000 & 3.6408 \end{pmatrix}.$$

This approach may lead to overly conservative results, as it involves provisioning for every scenario.

Revisiting scenarios: distributionally robust approach. We may consider reducing this conservatism using distributional robustness and entropic bounds on the scenarios, as seen in Section 3.4. Focusing on a column-wise model, and a generic single constraint $a^T y \geq l$, with $\mathcal{A}$ the corresponding uncertainty set associated with a, as before. We assume that the set $\mathcal{A}$ is a finite set of scenarios $a^{(k)}$, $k = 1, \ldots, K$. Further, we posit that each scenario is attached to a probability p_k, and that p is only known to belong to the entropic level set (2), with q a known reference distribution. We then formulate the robust counterpart of the expected load constraint, which is

$$l \leq \min_{p \in \mathcal{P}} \sum_{k=1}^{K} p_k (a^{(k)})^T y.$$

As seen in (2), the above is equivalent to the explicit convex constraint in y and an additional non-negative variable λ:

$$\gamma \lambda \geq l + \lambda \log \left(\sum_{k=1}^{K} q_k \exp(\frac{-(a^{(k)})^T y}{\lambda}) \right).$$

We can deal with the worst-case average cost in a similar manner.

Column-wise uncertainty. A class of uncertainty models on the matrix A assumes that each column a_i, $i = 1, \ldots, N$ (vehicle) is independently perturbed. Precisely, we assume that $A \in \mathcal{M}$, where

$$\mathcal{M} := \left\{ \begin{pmatrix} a_1 & \ldots & a_N \end{pmatrix} \ : \ a_i \in \hat{a}_i + V_i u_i \ : \ u_i \in \mathcal{U}_i, \ \ i = 1, \ldots, N \right\},$$

where $\hat{a}_i$'s correspond to nominal values, while $\mathcal{U}_i$'s are given subsets of $\mathbb{R}^{p_i}$, and $V_i \in \mathbb{R}^{T \times p}$ are given. Under this model, worst-case constraints and load constraints simplify greatly.

Concerning the load constraints, they write as $a_i^T y_i \geq L_i$, $i = 1, \ldots, N$, with y_i the i-th column of Y. Under the column-wise uncertainty model, it thus suffice to consider the robust counterpart to a generic single constraint $a^T y \geq l$, where l is a given scalar load, and a, y represent generic columns of A, Y. We assume that $a \in \hat{a} + V\mathcal{U}$, where $\hat{a} \in \mathbb{R}^T$ and $V \in \mathbb{R}^{T\times p}$, $\mathcal{U} \subseteq \mathbb{R}^p$ are given. The robust counterpart is

$$l \leq \min_{u\in\mathcal{U}} (\hat{a} + Vu)^T y = \hat{a}^T y - \phi_{\mathcal{U}}(-V^T y),$$

where, for given $z \in \mathbb{R}^p$:

$$\phi_{\mathcal{U}}(z) := \max_{u\in\mathcal{U}} z^T a$$

is the support function of $\mathcal{U}$ evaluated at z.

Likewise, the worst-case cost, thanks to the representation (6), simplifies to

$$\max_{A\in\mathcal{M}} \sum_{i=1}^{N} (c \circ y_i)^T a_i = \sum_{i=1}^{N} \left((c \circ y_i)^T \hat{a}_i + \phi_{\mathcal{U}_i}(V_i^T (c \circ y_i)) \right). \tag{11}$$

We obtain the following robust counterpart for the column-wise uncertainty model:

$$p^* = \min_{y_1,\ldots,y_N} \sum_{i=1}^{N} \left((c \circ y_i)^T \hat{a}_i + \phi_{\mathcal{U}_i}(V_i^T (c \circ y_i)) \right) \;:\; \hat{a}_i^T y_i \geq L_i + \phi_{\mathcal{U}}{}_i(-V_i^T y_i), \;\; y_i \geq 0, \;\; i = 1, \ldots, N. \tag{12}$$

Thus, under a column-wise uncertainty model, it suffices to develop a tractable representation of the support function of the sets $\mathcal{U}_i$; precisely a tractable way to express a constraint of the form $\phi_{\mathcal{U}_i}(z) \leq \tau$, where $z \in \mathbb{R}^{p_i}$ and $t \in \mathbb{R}$ are given.

Interval uncertainty on arrival and departure times. As a more specific example of column-wise uncertainty, we assume that the corresponding arrival and departure times are only known up to intervals. We can model this situation using a scenario approach.

To clarify this, consider a specific vehicle with time horizon $T = 7$, and an uncertain arrival time $t_a \in \{2, 3, 4\}$, with a charging duration fixed at 3. The corresponding column of A writes

$$a(u) = \hat{a} + Vu, \;\; \hat{a} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \;\; V = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & -1 \\ 0 & 0 \\ -1 & 0 \\ 0 & 1 \end{pmatrix}, \;\; u \in \mathcal{U} := \{v \in \{0,1\}^2, \;\; v^T \mathbf{1} \leq 1\}. \tag{13}$$

The number of uncertain parameters is $p = 2$.

More generally, we assume that the i-th column of A is perturbed as $a_i(u) = \hat{a}_i + V_i u_i$ with $u_i \in \mathcal{U}_i := \{v \in \{0,1\}^{p_i}, \;\; v^T \mathbf{1} \leq 1\}$. In that case, the robust counterpart is a simple scenario model, with $K_i := p_i + 1$ scenarios associated with each vehicle. Precisely, we set the corresponding scenarios as $a_i^{(k)} = \hat{a}_i + V_i e_k$, $k = 0, \ldots, p$, where for $k \geq 1$ e_k is the k-th unit vector in $\mathbb{R}^p$, and zero otherwise. We obtain the robust counterpart

$$p^* = \min_{y_1,\ldots,y_N} \sum_{i=1}^{N} \max_{1\leq k\leq K_i} (c \circ y_i)^T (a_i^{(k)}) \;:\; y_i \geq 0, \;\; (a_i^{(k)})^T y_i \geq L_i, \;\; k = 1, \ldots, K_i, \;\; i = 1, \ldots, N. \tag{14}$$

Typically $K_i = 9$, corresponding to nominal, high and low arrival and departure times for example.

We observe here a benefit with respect to the scenario uncertainty model described before. If we were to apply the latter directly, the total number of scenarios would be $K_1 \times \ldots \times K_N$. With a column-wise approach, we do not deal with such a potentially very large number of scenarios.