A Supplementary Material

A.1 *Full Proofs of Propositions*

We now report the full proof of Proposition 4.1.

Proposition 4.1. *The optimization problem in* (12) *can be posed as the constrained LS problem*

$$\begin{aligned} \textit{minimize} \quad & \alpha \cdot \|Ax - t\|_2^2 + \beta \cdot \|v\|_2^2 + \gamma \cdot \|u\|_2^2, \\ \textit{subject to} \quad & Dx \geq d,\, Tx \leq \tau,\, Bx - b \leq u, \\ & Bx - b \geq v,\, u \geq 0,\, v \leq 0, \end{aligned} \tag{17}$$

where $u, v \in \mathbb{R}^{|P|}$ *are additional auxiliary variables.*

Proof. As a first step of our proof, we show that

$$\text{minimize} \quad \|\max(0, Bx - b)\|_2^2$$

can be represented as a constrained LS problem with an additional variable $u \in \mathbb{R}^{|P|}$, as follows:

$$\begin{aligned} \text{minimize} \quad & \|u\|_2^2, \\ \text{subject to} \quad & Bx - b \leq u, 0 \leq u. \end{aligned} \tag{18}$$

In a similar way we can obtain a constrained LS problem corresponding to the minimization of $\|\min(0, Bx - b)\|_2^2$:

$$\begin{aligned} \text{minimize} \quad & \|v\|_2^2, \\ \text{subject to} \quad & Bx - b \geq v, v \leq 0. \end{aligned} \tag{19}$$

We can now replace (18) and (19) with the corresponding terms within (12), obtaining the problem in (17). □

A.2 *Scalability Evaluation*

In our main set of experiments, we have considered real data provided by our institution, for which our approach can compute the optimal solution in less than a second. Therefore, to properly evaluate the scalability of our approach, we consider the methodology discussed hereafter to generate realistic problem instances[5] with a large number of projects and researchers, and then we measure the runtime needed to calculate an optimal solution.

More precisely, we generate synthetic RPP instances with "realistic" features in two phases. The first phase consists of randomly generating the project's structure, i.e., the starting date and its budget, as well as the number of work packages. For each work package of the project, we randomly assign its required dedication and duration. The second phase involves the generation of the set of researchers that are involved in the projects, along with their availability, cost, and target dedication for a given project. These features are generated according to probability distributions reported in Table 1. We verified that such distributions result in RPP instances whose features resemble the realistic data provided by our institution.

We then generate a set of RPP problem instances varying the number of projects $|P|$ and the maximum number of work packages per project, namely W_{max}, to evaluate the scalability of our approach. We consider all combinations of $|P| \in \{5, 10, 50, 100, 500, 1000\}$ and $W_{max} \in \{5, 10, 15, 20, 25\}$, generating 10 instances for each pair of parameters, for a total of 300 instances.

[5] RPP is a novel problem, hence no instance generators are readily available in the literature. Here we employ a generation procedure inspired by the methodology of Drexl et al. [5].

Table 1: Distributions used for RPP instance generation.

Distribution	Feature
$Exp(1.8 \cdot 10^{-1})$	Start date of project
$U(26.4, 44.3)$	Mean budget per hour per project
$U(1, W_{max})$	Number of work packages per project
$Exp(7.6 \cdot 10^{-4})$	Dedication of work packages
$N(38.1, 4.7)$	Duration of work packages
$U(120, 140)$	Availability of researchers
$U(17.2, 54.0)$	Cost per hour for researchers
$Bern(3.9 \cdot 10^{-1})$	Is a researcher in a work package?
$U(0, \tau(r, m))$	Target dedication for r during m

Notice that, while we generate test instances with a fixed number of projects, the total number of researchers $|R|$ depends on the dedication demand of such projects.

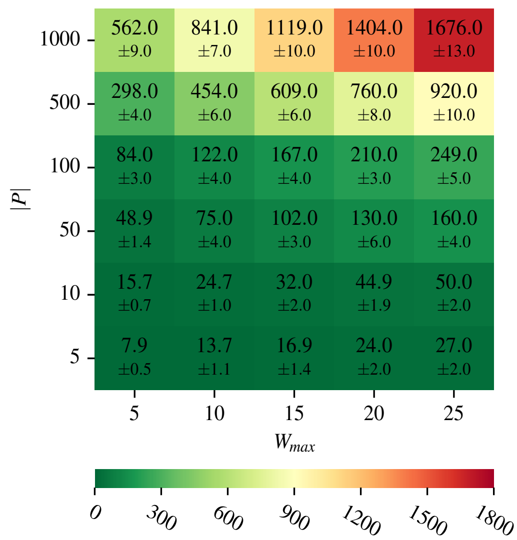


Figure 5: Number of researchers generated with respect to the number of projects $|P|$ and the maximum number of work packages per project W_{max} (best viewed in colors).

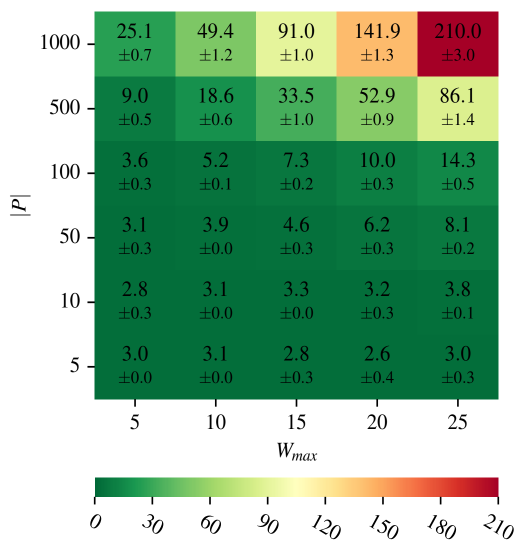
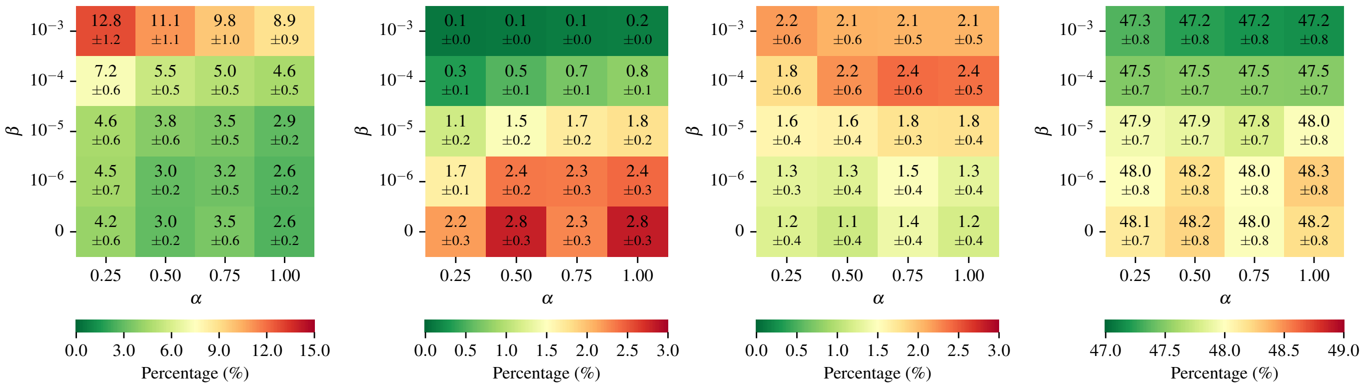


Figure 6: Total solution runtime. Each cell reports the average and the standard error of the mean over the 10 instances generated for each combination of $|P|$ and W_{max} (best viewed in colors).



(a) Target deviation

(b) Budget residual

(c) Dedication residual

(d) Contracted dedication

Figure 7: Results in the *sufficient budget* scenario for $\gamma = \infty$ considering uncertainty (best viewed in colors).

Figure 5 shows how the total number of researchers varies according to the number of projects $|P|$ and the maximum number of work packages per project W_{max}. That is, more researchers are generated to fulfill the demand required by a higher number of projects.

Figure 6 reports the total runtime of our approach (including both the runtime required for compiling the norm approximation problem and one required for solving it) with respect to the number of projects and the maximum number of work packages per project. Results show that instances with up to 100 projects are solved in at most a few tens of seconds. Moreover, for (arguably unrealistic) very large instances considering 1000 projects, a maximum of 25 work packages per project, and more than 1600 researchers, our approach is able, on average, to compute the optimal solution in 210 seconds.

This capability of scaling to very large instances is also thanks to the employment of *sparse matrices* to represent A, B, D, and T matrices, which allows us to achieve a dramatic reduction (up to 5 orders of magnitude) in terms of memory consumption with respect to the standard dense matrix representation. Overall, results show that our approach makes it viable to compute optimal solutions even for very large RPP instances with realistic features.

A.3 Plots of Experiments with Uncertainty

In Figure 7 we report the heatmaps visualizing the results of the experiments with uncertainty summarized in Section 5.2.