# The Probabilistic $k$-Center Problem *

Marc Demange[1], Marcel Adonis Haddad[†1,2] and Cécile Murat[2]

[1]School of Science RMIT University, Melbourne, Vic., Australia
marc.demange@rmit.edu.au
[2]Paris-Dauphine Université, PSL Research University, CNRS UMR 7243, Lamsade
75016 Paris, France
marcel-adonis.haddad@dauphine.eu, cecile.murat@lamsade.dauphine.fr

## Abstract

The $k$-Center problem on a graph is to find a set $K$ of $k$ vertices minimizing the radius defined as the maximum distance between any vertex and $K$. We propose a probabilistic combinatorial optimization model for this problem, with uncertainty on vertices. This model is inspired by a wildfire management problem. The graph represents the adjacency of zones of a landscape, where each vertex represents a zone. We consider a finite set of fire scenarios with related probabilities. Given a $k$-center, its radius may change in some scenarios since some evacuation paths become impracticable. The objective is to find a robust $k$-center that minimizes the expected value of the radius over all scenarios. We study this new problem with scenarios limited to a single burning vertex. First results deal with explicit solutions on paths and cycles, and hardness on planar graphs.



## 1 Introduction

Forest fires are becoming an increasing concern for the population across the globe, and with projected climate change, this upward trend seems set to continue. Operations Research methods are one of the tools used by wildfire managers to guide decision making [1, 2]. Evacuation planning and facility location under uncertainty are some of the considered problems. In this paper, we address the problem of locating a number $k$ of emergency gathering points in a forest, in anticipation of fires. The general objective is to minimize the risk of having people trapped by the flames, i.e., maximize their chance to reach a safe gathering point or shelter before the fire arrives. Rescue operations are deployed to evacuate people from the shelters but these shelters are also equipped to protect people against flames if rescue cannot be completed on-time.

In our model, the landscape is represented by an adjacency graph $G = (V, E)$. Each node corresponds to an area and two nodes are connected by an edge if the related areas are adjacent. A fire may ignite on a node and

*We acknowledge the support of GEO-SAFE, H2020-MSCA-RISE-2015 project #691161.
†Corresponding author



In: G. Di Stefano, A. Navarra Editors: Proceedings of the RSFF'18 Workshop, L'Aquila, Italy, 19-20-July-2018, published at http://ceur-ws.org

then can spread on all neighboring nodes at a fixed maximum distance called the *range of impact.* The *impacted zone* is then the set of vertices where the fire spreads. In a real case scenario, the fire can spread over very large areas and the impacted zone grows dynamically. However, it might be relevant to focus on a relatively short period after ignition or after the alert, seen as the time required for all people present in the area to reach an assembly point. This motivates us to consider small ranges of impact depending on the efficiency of the early warning system. We then assume that people are safe after reaching the assembly point; this hypothesis is relevant depending on the exact nature and design of the assembly points.

Our problem is a particular case of *location-allocation problems.* A solution describes not only the location of emergency assembly points, defined as a subset $K \subseteq V, |K| = k$ and known as a *k-center* (location problem), but it also includes a strategy people should follow in case of fire to reach one assembly point (evacuation process). The location can be addressed during the *preparedness phase* and may involve sophisticated solutions. The allocation however is mainly addressed during the *response phase*, which usually requires efficient and simple processes. Indeed, even though evacuation scenarios can be prepared in advance, they should remain straightforward in order to be followed by untrained people, without supervision and under high stress conditions. Here, we will consider that somebody is assigned to the closest assembly point with the constraint he/she will never go through the fire; this constraint will be precise later. For a real case implementation, this could be supported by simple early warning system deployed on site and indicating the direction and distance of the closest accessible shelter(s). The objective is then to minimize the maximum distance from each vertex $v$ to $K$, called the *radius* of the $k$-center $K$. It corresponds to the worst case situation for someone present in the risk zone when a fire starts. However high uncertainty lies on where a fire ignites and thus on the accessibility of some vertices or on the practicability of some paths during the event. We model it as different possible scenarios, where a scenario is given by an ignition vertex and a range of impact. Then, the objective becomes the expected value of the radius in the new instance. We will restrict ourselves to a simple fire outbreak scenarios. This hypothesis is supported by our choice to focus on a relative short period after ignition. To simplify the problem in this first study we will consider a fixed range of impact. This corresponds to the situation where the considered region is homogeneous in terms of topography, wind and fuel load conditions.

The framework of *Probabilistic combinatorial optimization* [3–5], also known as *a priori optimization*, is particularly suitable to model such situations, where one has to deal with destructions or obstructions on a usual set-up that make an original global strategy potentially unfeasible. In this approach, a problem is decomposed in two phases. The whole instance is considered during the first phase (seen as the preparedness phase), while in a second time (response phase) only a part of the infrastructure remains available after an uncertain event has occurred. A *modification strategy* is identified beforehand to "automatically" transform a solution on the whole instance (called an *a priori* solution) to a feasible solution on a partial available instance (an *a posteriori* solution). To represent uncertainty, probabilities are assigned to the different possible partial instances. More precisely it is usually based on a probability distribution on the different components of the initial instance representing how likely they could be affected. The objective is then to compute an a priori solution on the whole instance optimizing the expected value of the effective solution induced on the partial instance using the modification scheme. Given our allocation strategy that assigns a person at risk to the closest safety point, the modification scheme describes the new set of emergency points after the event and the new possible routes so as to recompute the new allocation and the corresponding objective value. In this paper, since we assume the emergency gathering points or shelters are safe and since it is not possible to define new locations during the response phase, we consider that the $k$-center is not modified. Indeed, any original $k$-center always remains feasible in the new graph. The modification scheme only corresponds to the new possible routes; roughly speaking these routes are paths that do not cross the impacted zone. Specific rules are considered for people in this impacted zone as detailed in the next section.

In the usual non-probabilistic case, also called *deterministic* case, our problem is known as the MINIMUM $k$-CENTER problem and to our knowledge this is the first time the Probabilistic $k$-center is defined and discussed [6]. However, restricted versions of routing and networking-design probabilistic minimization problems (in complete graphs) have been studied (see, e.g., [7–11] ). In [12–15], the analysis of the probabilistic minimum traveling salesman problem, originally presented in [3, 4], has been revisited in order to propose new efficient resolution. Several other combinatorial problems have been also handled in the probabilistic combinatorial optimization approach, with or without recourse, including minimum vertex cover and maximum independent set [16, 17], longest path [18], Steiner tree problems [19, 20], minimum spanning tree [9, 21], minimum dominating set [22] and some other general combinatorial graph problems [23].

In this paper we apply the probabilistic combinatorial optimization setting for the $k$-center problem and we use the following notations. Let $G = (V, E)$ be an undirected and connected graph with $V = \{0, \ldots, n\}$. For $v \in V$, $N(v)$ is the set of its neighbors. A path of length $m$ is a sequence $v_0, \ldots, v_m$ of pairwise distinct vertices where $\forall i \in \{0, \ldots, m-1\}, v_i v_{i+1} \in E$. For $x, y \in V$, we denote by $\delta(x, y)$ the distance between $x$ and $y$ (that is the length of the shortest path between them). Similarly for a set $K \subseteq V$, $\delta(x, K) = \min_{y \in K} \delta(x, y)$. Given a $k$-center $K \subseteq V, |K| = k$, its radius in $G$ is denoted by $r^K = \max_{v \in V} \delta(v, K)$. Finally, given a real-valued function $f$ with domain $A$, we denote $\arg\min_{x \in A} f(x)$ the set $\{x \in A : f(x) = \min_{y \in A} f(y)\}$ of the minimum(s) of $f$ in $A$.

The paper is organized as follows. In Section 2, we introduce and define the problem, namely Probabilistic $k$-Center. In Section 3, we give and prove an explicit optimal solution on path and cycle instances. Section 4 deals with hardness of approximation on planar graphs of degrees 2 or 3. Finally, we conclude in Section 5.

## 2 Definition of the problem

Probabilistic $k$-Center is characterized by a couple $(\mathbb{S}, \mathbb{M})$, where $\mathbb{S}$ is the set of scenarios in each instance, and $\mathbb{M}$ is a modification strategy. In this paper we study scenarios where the fire is limited to a simple node. Then $\mathbb{S} = \{s_i : i \in V\}$ with $s_i$ the scenario where vertex $i$ burns. With the considered modification strategy $\mathbb{M}_0$, the location of the centers are conserved and the vertices are affected to the nearest center in the available instance. This strategy involves the following: when a fire ignites, all the people in the forest try to escape to the closest shelter to get evacuated by the rescue. If there is a shelter in the ignition area, we assume that people in this area find refuge in that shelter. Otherwise the people in the ignition zone always flee in the opposite direction of the fire. Once at sufficient distance, they can choose their path rationally in order to get to the closest shelter.

An instance of our problem is then a couple $(G, \mathrm{P})$ with $G$ a graph and P a probability system associated with the vertices. It gives, for each vertex, the probability of fire on it. By default, we will consider a uniform distribution under the assumption that we have one fire outbreak at a time on a simple node. So, each scenario has probability $p = \frac{1}{|V|}$. So in this work, Probabilistic $k$-Center will refer to this set-up. For scenario $s_j$ and for any $x, y \in V$, we denote by $\delta_j(x, y)$ the distance between $x$ and $y$ in $G \setminus \{j\}$. Note that $\delta_j(x, x) = 0$ and $\delta_j(x, y) = +\infty$ if there is no path between $x$ and $y$. Then, for scenario $s_j$ and $K \subset V$, the *induced evacuation distance* of a node $x$ is given by:

$$D_j(x, K) = \begin{cases} \delta_j(x, K) & \text{if } (x \neq j) \text{ or if } (x = j \text{ and } j \in K) \\ \max_{x \in N(j)} \{1 + \min_{y \in K} \delta_j(x, y)\} & \text{if } x = j \text{ and } j \notin K \end{cases} \quad (1)$$

Then, the *disrupted radius* of $K$ for scenario $s_j$ induced by $\mathbb{M}_0$ is:

$$r_j^K = \max_{x \in V} D_j(x, K) \quad (2)$$

We illustrate these definitions for $k = 3$ on a path of nine vertices and the 3-center $K = \{0, 5, 8\}$, whose nodes are drawn as pentagons. Figure 1, gives, for each node $x$, the value of $\delta(x, K)$. So, the radius of this solution is 2. In Figure 2, we illustrate a fire occurring on node 1. We give under each node $x \in V$ the value of $D_1(x, K)$, and

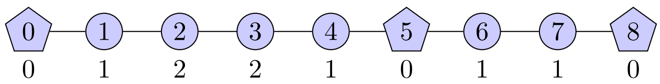


Figure 1: The distance of each node to the 3-centers {0,5,8} in the deterministic case.

underline it when it has changed. For example, node 2 is affected since the original evacuation path to shelter 0 is no more operational. Note that, in the worst case, people on node 1 escape to node 2 and then to shelter 5. The disrupted radius $r_1^K$ of scenario 1 is 4.

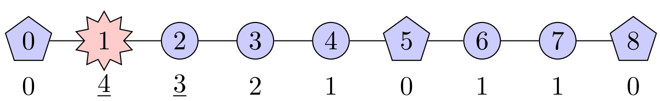


Figure 2: Evacuation distances in the scenario $s_1$.

Figure 3 illustrates a fire on node 5, with a shelter. It affects nodes 3,4 and 6.

We denote by $E$ the objective function of the PROBABILISTIC $k$-CENTER problem. It is the expectation value of the disrupted radius of $K$ :

$$E(K) = \sum_{j \in V} p r_j^K = p \cdot \sum_{j \in V} r_j^K \tag{3}$$

A solution $K$ is said feasible if its objective value is finite. We denote by $\mathcal{K}(G)$ the set of feasible solutions.

**Remark 1.** *Denote by $\Psi(G)$ the set of connected components of a graph $G$. Then $K$ is a feasible solution if $\forall j \in V, \forall W \in \Psi(V \setminus \{j\})$, $K \cap W \neq \emptyset$.*

An optimal solution for our version of the PROBABILISTIC $k$-CENTER problem is then a $k$-center $K^*$ verifying $K^* \in \arg\min_{K \in \mathcal{K}}\{E(K)\}$.

We introduce the following decomposition of the objective function:

$$E(K) = E_s(K) + E_{\bar{s}}(K) \quad s.t. \ : E_s(K) = p \sum_{j \in K} r_j^K, E_{\bar{s}}(K) = p \sum_{j \in G \setminus K} r_j^K \tag{4}$$

In other words, $E_s(K)$ is the contribution of scenarios for which fires occur in $K$ (called the *skeleton*), and $E_{\bar{s}}(K)$ is the contribution of scenarios for which fires occur on the other vertices (called the *body*). We can treat these components as two different problems. In particular, $E_{\bar{s}}(K)$ corresponds to a version of our problem where the nodes of $K$ are *fortified* [24], which means they are immune to fire. Note that, if a solution is optimal for both components, then it is optimal for the whole problem.

Moreover, we propose another evaluation function, denoted $\widehat{E}$, corresponding to the expected evacuation distance in the local area of the ignition node $j$ defined as the close neighborhood $(N(j) \cup \{j\})$ of $j$. The *locally induced radius* is then denoted $\widehat{r}_j^K = \max_{x \in N(j) \cup \{j\}} D_j(x, K)$ :

$$\widehat{E}(K) = p \cdot \sum_{j \in V} \widehat{r}_j^K \tag{5}$$

Obviously, for a feasible $K$ and $\forall j \in V$, we have:

$$\widehat{E}(K) \leq E(K) \tag{6}$$

We similarly decompose $\widehat{E}(K)$ in $\widehat{E}_s(K)$ and $\widehat{E}_{\bar{s}}(K)$.

## 3 An explicit solution on Paths and Cycles

Denote $P_{n+1}$ a path on $n+1$ vertices, $C_n$ a cycle on $n$ vertices, and $H$ a graph that is either a path or cycle with $n$ edges. Given the feasibility condition seen in Remark 1, for $K = (v_1, \dots, v_k) \in \mathcal{K}(P_{n+1})$, we necessarily have $v_1 = 0$ and $v_k = n$. Thus, $K$ induces $k-1$ segments $\mu_i$ of length $\lambda_i$, $i = 1, \dots, k-1$. On a cycle $C_n$ and for $K \in \mathcal{K}(C_n)$, $K$ would induces $k$ segments. In the following we deal with path and cycles simultaneously, unless specified otherwise. We define $\kappa = k$ if $H$ is a cycle, and $\kappa = k-1$ if $H$ is a path. We denote $E^H(K)$ the value of the solution $K$ in $H$. Denoting $\lambda = (\lambda_1, \dots, \lambda_\kappa)$, we can consider equivalently $E^H(K)$, $E_s^H(K)$ and $E_{\bar{s}}^H(K)$ as functions of $\lambda(K)$. We say that $\lambda = (\lambda_1, \dots, \lambda_\kappa)$ is *not decreasing* if $\lambda_1 \leq \ldots \leq \lambda_\kappa$. We define $\lambda_{\leq} = (\lambda_{i_1}, \dots, \lambda_{i_\kappa})$, where $i : \{1, \dots, \kappa\} \to \{1, \dots, \kappa\}$ is a permutation such that $\lambda_{i_1} \leq \ldots \leq \lambda_{i_\kappa}$, a non decreasing solution induced by $\lambda$. The possible $\lambda$s corresponding to a feasible $k$-center are all vectors, $(\lambda_1, \dots, \lambda_\kappa) \in \mathbb{N}^\kappa$ such that $\sum_{i=1}^{\kappa} \lambda_i = n$. We denote by $\Lambda(H)$ their set. The aim of this section is to give an explicit optimal solution for PROBABILISTIC $k$-CENTER on paths and cycles. In both cases, the balanced solution (see Definition 1), optimal in the deterministic case, reveals also to be optimal in the probabilistic case. However, the proof is non-trivial. For paths, we will show, in a first step, that this solution minimizes the contribution of the skeleton and in a second step, that it minimizes also the contribution of the body. We then derive the case on cycle by a reduction to the case on paths.

### 3.1 Expression of the disrupted radius and balanced solution

The results in this subsection hold for both cycles and paths with $n$ edges. We define $\mu_0 = \mu_{\kappa+1} = \emptyset$ in the case of paths. We recall that the objective function of the PROBABILISTIC $k$-CENTER problem requires the disrupted radius of $K$ induced for each scenario (see Equations (2), (3)). For scenario $s_j$, let $\mu_i$ be a segment such that $j \in \mu_i, i \in \{1, \ldots, \kappa\}$. We have:

$$r_j^K = \max\{\max_{x \in \mu_i}\{D_j(x,K)\}, \max_{x \in H \setminus \mu_i}\{D_j(x,K)\}\}$$

Observe that, $\forall x \in H \setminus \mu_i$, the evacuation path is not modified. On $\mu_q \notin \mu_i$, the evacuation path of any node is shorter than the evacuation path of the middle node(s) of $\mu_q$. Then, $\forall x \in \mu_q, D_j(x,K) \leq \lfloor \frac{\lambda_q}{2} \rfloor$ and $\max\limits_{x \in H \setminus \mu_i}\{D_j(x,K)\} = \max\limits_{q \in \{1,\ldots,\kappa\}: \mu_q \notin \mu_i}\{\lfloor \frac{\lambda_q}{2} \rfloor\}$. Now we look at $\max\limits_{x \in \mu_i}\{D_j(x,K)\}$. The evacuation paths on segment $\mu_i$ is always shorter than one of the evacuation paths from $j$ or - if $j \in K$ - from one of his neighbor. We distinguish then two cases:

If $j \in K$, $\max\limits_{x \in \mu_i \cup \mu_{i+1}}\{D_j(x,K)\} = \max\{D_j(j-1,K), D_j(j+1,K),\}$ since $D_j(j,K) = 0$. Based on Equation (1) we obtain:

$$r_j^K = \max\{\lambda_i - 1, \lambda_{i+1} - 1, \max_{q=1,\ldots,\kappa; j \notin \mu_q} \lfloor \tfrac{\lambda_q}{2} \rfloor\} \tag{7}$$

For example, in Figure 3, $r_5^K = \max\{\lambda_1 - 1, \lambda_2 - 1, \lfloor \frac{\lambda_3}{2} \rfloor\} = \max\{4, 2, 5\} = 5$.

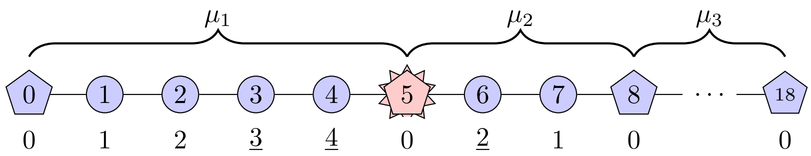


Figure 3: Evacuation distances to the 4-center $\{0, 5, 8, 18\}$ in the scenario $s_5$.

If $j \notin K$, $\max\limits_{x \in \mu_i}\{D_j(x,K)\} = D_j(j,K)$. Using Equation (1) we then have $D_j(j,K) = \max\{j - v_i, v_{i+1} - j\}$, and therefore:

$$r_j^K = \max\{j - v_i, v_{i+1} - j, \max_{q=1,\ldots,\kappa; j \notin \mu_q} \lfloor \tfrac{\lambda_q}{2} \rfloor\} \tag{8}$$

**Definition 1.** *A solution is called* balanced *if* $\forall i, j \in \{1, \ldots, \kappa\}, |\lambda_i - \lambda_j| \leq 1$, *and it will be called non-decreasing if* $\forall i, j \in \{1, \ldots, \kappa\}, i < j, \lambda_i \leq \lambda_j$. *We denote* $K^B \in \mathcal{K}(H)$ *the solution such that* $K^B$ *is a non-decreasing balanced* $k$*-center.*

We denote by $\lambda^B$ the related vector of distances. Thus, we have $\lambda_i^B \in \{\lfloor \frac{n}{\kappa} \rfloor, \lceil \frac{n}{\kappa} \rceil\}$ and $\lambda_1^B \leq \ldots \leq \lambda_\kappa^B$ for $i = 1, \ldots, \kappa$. In what follows, we show that the Probabilistic $k$-Center problem has an optimal balanced solution on paths and cycles. For $k \geq \frac{n}{2}$, we have the following result (the proof is given in Appendix):

**Proposition 1.** *For* $k \geq \frac{n}{2}$, $K^B$ *is an optimum solution for* Probabilistic $k$-Center *on paths and cycles.*

So, in what follows, we assume $k < \frac{n}{2}$. In this case, we show in Appendix the Lemma 1 :

**Lemma 1.** $E_s^H(\lambda^B) = \widehat{E}_s^H(\lambda^B)$ , $E_{\bar{s}}^H(\lambda^B) = \widehat{E}_{\bar{s}}^H(\lambda^B)$ *and* $E^H(\lambda^B) = \widehat{E}^H(\lambda^B)$.

In subsection 3.2, we show that $K^B$ minimizes $E_s^{P_{n+1}}(K)$. In order to prove that $K^B$ minimizes also $E_{\bar{s}}^{P_{n+1}}(K)$, in subsection 3.3, we establish a more general result, by showing that $K^B$ minimizes $E_{\bar{s}}^H(K)$. So, we can obtain Theorem 1 for paths and in subsection 3.4, we consider cycles.

### 3.2 The skeleton part on paths

In this subsection, we focus on the skeleton part on paths. We first show that some balanced solution minimizes $\widehat{E}_s^{P_{n+1}}(K)$ and then, that it minimizes as well $E_s^{P_{n+1}}(K)$. Based on Equations (5), (4) and (7) we have:

$$\widehat{E}_s^{P_{n+1}}(\lambda) = p \cdot \left(\lambda_1 + \sum_{i=2}^{\kappa} \max(\lambda_{i-1}, \lambda_i) + \lambda_\kappa\right) - p(\kappa + 1) \tag{9}$$

**Lemma 2.** $\widehat{E}_s^{P_{n+1}}(\lambda) \geq \widehat{E}_s^{P_{n+1}}(\lambda_{\leq})$

*Proof.* Consider $\lambda = (\lambda_1, \dots \lambda_\kappa)$ with $\lambda_i \geq 0, i = 1, \dots, \kappa$. If $\lambda_1 = \max\{\lambda_i, i = 1, \dots \kappa\}$, we define $r = 0$, else let $r$ be the maximum index in $\{1, \dots, \kappa\}$ such that $\lambda_1 \leq \dots \leq \lambda_r$ and $\forall j \geq r, \lambda_j \geq \lambda_r$. If $r = \kappa$, we have $\lambda = \lambda_{\leq}$ and nothing needs to be shown. Else, $r < \kappa - 1$ (the value $r = \kappa - 1$ is not possible), we consider $s \in \arg\min\{\lambda_j, j = r+1, \dots, \kappa\}$: by definition, we have $\lambda_{r+1} \geq \lambda_r$ if $r > 0$, $s > r+1$ and $\lambda_r \leq \lambda_s < \lambda_{r+1}$. We then consider $\lambda'$ the vector obtained from $\lambda$ by moving the $s^{\text{th}}$ coordinate at the position $r+1$:

$$\left| \begin{array}{llllll} \lambda'_i & = & \lambda_i, & i = 1, \dots, r & \lambda'_{r+1} & = \lambda_s, \\ \lambda'_i & = & \lambda_{i-1}, & i = r+2, \dots, s & \lambda'_i & = \lambda_i, \quad i > s \end{array} \right.$$

We claim that $\widehat{E}_s^{P_{n+1}}(\lambda') \leq \widehat{E}_s^{P_{n+1}}(\lambda)$. Indeed, suppose first $s < \kappa$ and consider the expression of $\widehat{E}_s^{P_{n+1}}(\lambda')$ and $\widehat{E}_s^{P_{n+1}}(\lambda)$, as sums of $\kappa + 1$ terms (see Equation (9)). The three terms $\max(\lambda_r, \lambda_{r+1}), \max(\lambda_{s-1}, \lambda_s)$ and $\max(\lambda_s, \lambda_{s+1})$ in the expression of $\widehat{E}_s^{P_{n+1}}(\lambda)$ are replaced in the expression of $\widehat{E}_s^{P_{n+1}}(\lambda')$ by $\max(\lambda_r, \lambda_s), \max(\lambda_s, \lambda_{r+1})$ and $\max(\lambda_{s-1}, \lambda_{s+1})$ and all the other terms are identical in both expressions. We conclude:

$$\begin{aligned} \widehat{E}_s^{P_{n+1}}(\lambda) - \widehat{E}_s^{P_{n+1}}(\lambda') = \; & p\,(\max(\lambda_r, \lambda_{r+1}) + \max(\lambda_{s-1}, \lambda_s) + \max(\lambda_s, \lambda_{s+1}) \\ & - [\max(\lambda_r, \lambda_s) + \max(\lambda_s, \lambda_{r+1}) + \max(\lambda_{s-1}, \lambda_{s+1})]) \\ = \; & p\,(\lambda_{r+1} + \lambda_{s-1} + \lambda_{s+1} - \lambda_s - \lambda_{r+1} - \max(\lambda_{s-1}, \lambda_{s+1}))) \\ = \; & p\,(\lambda_{s-1} + \lambda_{s+1} - \lambda_s - \max(\lambda_{s-1}, \lambda_{s+1})) \geq 0 \end{aligned}$$

where the last inequality holds since $\lambda_s \leq \min(\lambda_{s-1}, \lambda_{s+1})$. Suppose now $s = \kappa$, then a similar approach without the terms involving $s+1$ leads to:

$$\begin{aligned} \widehat{E}_s^{P_{n+1}}(\lambda) - \widehat{E}_s^{P_{n+1}}(\lambda') = \; & p \quad (\max(\lambda_r, \lambda_{r+1}) + \max(\lambda_{\kappa-1}, \lambda_\kappa) - [\max(\lambda_r, \lambda_\kappa) + \max(\lambda_\kappa, \lambda_{r+1})]) \\ = \; & \quad p\,(\lambda_{r+1} + \lambda_{\kappa-1} - \lambda_\kappa - \lambda_{r+1}) \geq 0 \text{ (since } \lambda_{\kappa-1} \geq \lambda_\kappa) \end{aligned}$$

Note that these arguments hold also if $r = 0$. The proposition is deduced by induction, repeatedly replacing the current vector $\lambda$ by $\lambda'$. □

We are now ready to show the main result of this part:

**Proposition 2.** $\lambda^B \in \arg\min\{E_s^{P_{n+1}}(\lambda), \lambda \in \Lambda(P_{n+1})\}$.

*Proof.* First we claim that $\lambda^B \in \arg\min\{\widehat{E}_s^{P_{n+1}}(\lambda), \lambda \in \Lambda(P_{n+1})\}$. Indeed in order to minimize $\widehat{E}_s^{P_{n+1}}(\lambda)$ for $\lambda \in \Lambda(P_{n+1})$, Lemma 2 ensures we can restrict ourselves to non decreasing solutions. But for any non decreasing $\lambda$ in $\Lambda(P_{n+1})$, we have $\widehat{E}_s^{P_{n+1}}(\lambda) = p(n+\lambda_\kappa) - p(\kappa+1)$ (Equation (9)) and consequently a non decreasing solution minimizing $\widehat{E}_s^{P_{n+1}}(\lambda)$ is obtained by solving:

$$\left\{ \begin{array}{ll} \min & \lambda_\kappa \\ & \lambda_1 \leq \dots \leq \lambda_\kappa \\ & \sum\limits_{i=1}^{\kappa} \lambda_i = n, \ \lambda \in \mathbb{N}^\kappa \end{array} \right.$$

This admits $\lambda^B$ as unique solution. Therefore, by Lemma 1 and Equation (6), we have: $E_s^{P_{n+1}}(\lambda^B) = \widehat{E}_s^{P_{n+1}}(\lambda^B) = \min\limits_{\lambda \in \Lambda(P_{n+1})} \widehat{E}_s^{P_{n+1}}(\lambda) \leq \min\limits_{\lambda \in \Lambda(P_{n+1})} E_s^{P_{n+1}}(\lambda)$ and consequently $\lambda^B \in \arg\min\{E_s^{P_{n+1}}(\lambda), \lambda \in \Lambda(P_{n+1})\}$. □

### 3.3 The body part

Here, we focus on ignition outside the skeleton, the "body" part, on paths and cycles and show that there is a balanced solution minimizing $E_{\bar{s}}^H(K)$. Our proof for the body part immediately applies for both cases, which is worth to be noted since, as already mentioned, the problem restricted to the body has its own interest in applications. First, we highlight some properties of a balanced solution. For the following let us denote $K = (v_1, \dots, v_k)$.

**Lemma 3.** *Given an initial solution $\lambda \in \Lambda(H)$, with $\lambda_a + \lambda_b = m, \lambda_a \geq \lambda_b$ and $|\lambda_a - \lambda_b| \geq 2$ for some $1 \leq b \leq a \leq \kappa$, we define a solution $\lambda'$ such that $\lambda'_a + \lambda'_b = m, |\lambda'_a - \lambda'_b| \leq 1$, $\lambda'_a \geq \lambda'_b$ and $\lambda'_i = \lambda_i, \forall i = 1, \dots, \kappa, i \neq a, b$. Then we have $\widehat{E}_{\bar{s}}^H(\lambda') \leq \widehat{E}_{\bar{s}}^H(\lambda)$.*

*Proof.* Let us denote $\alpha(\lambda_a) = \sum\limits_{j \in \mu_a \setminus K} \max\{j - v_a, v_{a+1} - j\}$ the contribution of $\mu_a$ to the value of $\widehat{E}_{\bar{s}}(\lambda)$. Then $\widehat{E}_{\bar{s}}^H(\lambda) = p \sum\limits_{i=1,\dots,\kappa} \alpha(\lambda_i)$. Denote also $A$ a logical proposition, and by $\mathbb{1}_A$ the boolean function such that

$\mathbb{1}_A = 1$ if $A$ is true, and $\mathbb{1}_A = 0$ if $A$ is false. For $j \in \lambda_1$, we have $\alpha(\lambda_a) = \sum_{j \in \mu_1 \setminus K} \max\{j - v_a, v_{a+1} - j\} = \sum_{j \in \mu_a \setminus K} \max\{j, \lambda_a - j\} = 2 \cdot \sum_{j=\lceil \frac{\lambda_a}{2} \rceil}^{\lambda_a - 1} (j) - \mathbb{1}_{(\lambda_a \text{ even})}(\frac{\lambda_a}{2}) = \lfloor \frac{\lambda_a}{2} \rfloor (\lambda_a + \lceil \frac{\lambda_a}{2} \rceil - 1) - \mathbb{1}_{(\lambda_a \text{ even})}(\frac{\lambda_a}{2})$. The previous result applies also to $\alpha(\lambda_b)$. As $\lambda_b = m - \lambda_a$, we can express $\alpha(\lambda_a) + \alpha(\lambda_b)$ as a function of $\lambda_a$:

$$\alpha(\lambda_a) + \alpha(\lambda_b) = \begin{array}{l} \lfloor \frac{\lambda_a}{2} \rfloor (\lambda_a + \lceil \frac{\lambda_a}{2} \rceil - 1) - \mathbb{1}_{(\lambda_a \text{ even})}(\frac{\lambda_a}{2}) \\ + \lfloor \frac{m - \lambda_a}{2} \rfloor (m - \lambda_a + \lceil \frac{m - \lambda_a}{2} \rceil - 1) - \mathbb{1}_{((m-\lambda_a) \text{ even})}(\frac{m - \lambda_a}{2}) \end{array}$$

If we study the different combinations of parities of $\lambda_a$ and $m$, we get:

$$\tfrac{3}{2}\lambda_a^2 - \tfrac{3n}{2}\lambda_a + \tfrac{3}{4}m^2 - m \leq \alpha(\lambda_a) + \alpha(\lambda_b) \leq \tfrac{3}{2}\lambda_a^2 - \tfrac{3m}{2}\lambda_a + \tfrac{3}{4}m^2 - m + \tfrac{1}{2}$$

Defining the function $f(\lambda_a) = \frac{3}{2}\lambda_a^2 - \frac{3m}{2}\lambda_a$, we have:

$$\alpha(\lambda_a) + \alpha(\lambda_b) - \frac{1}{2} \leq f(\lambda_a) + \frac{3}{4}m^2 - m \leq \alpha(\lambda_a) + \alpha(\lambda_b) \tag{10}$$

and the same holds for $(\lambda'_a, \lambda'_b)$. $\frac{3}{4}m^2 - m$ is fixed and $f(\lambda_a)$ increases for $\lambda_a > \frac{m}{2}$. Since $\lambda'_a \leq \lambda_a$, $f(\lambda'_a) \leq f(\lambda_a)$. Since $\alpha(\lambda_a) + \alpha(\lambda_b)$ and $\alpha(\lambda'_a) + \alpha(\lambda'_b)$ are integers and $\lambda'_a \leq \lambda_a$, we deduce from Equation (10) $\alpha(\lambda'_a) + \alpha(\lambda'_b) \leq \alpha(\lambda_a) + \alpha(\lambda_b)$. As $\alpha(\lambda_i) = \alpha(\lambda'_i), i = 1, \dots, k-1, i \neq a, b$, we conclude $\widehat{E}_{\bar{s}}(\lambda') \leq \widehat{E}_{\bar{s}}(\lambda)$. □

**Lemma 4.** $\widehat{E}^H_{\bar{s}}(\lambda) \geq \widehat{E}^H_{\bar{s}}(\lambda^B)$

*Proof.* Note first that $\forall \lambda \in \Lambda(H), \widehat{E}^H_{\bar{s}}(\lambda) = \widehat{E}^H_{\bar{s}}(\lambda_{\leq})$, and assume $\lambda$ is a non decreasing solution. We look at the pair of intervals $(\lambda_1, \lambda_\kappa)$ with the largest length difference. If $|\lambda_1 - \lambda_\kappa| \leq 1$, then $\lambda$ is a balanced solution and the lemma is verified. Otherwise, we create a new solution $\lambda'$ by replacing the extreme intervals by a new pair of intervals of size $\lambda'_1 = \lfloor \frac{n - \sum_{i=2}^{\kappa-1} \lambda_i}{2} \rfloor$ and $\lambda'_\kappa = \lceil \frac{n - \sum_{i=2}^{\kappa-1} \lambda_i}{2} \rceil$. As $\lambda_1 + \lambda_\kappa = \lambda'_1 + \lambda'_\kappa$, by Lemma 3, we deduce $\widehat{E}^H_{\bar{s}}(\lambda) \geq \widehat{E}^H_{\bar{s}}(\lambda')$. We denote $\lambda'' = \lambda'_{\leq}$ the non decreasing solution induced by $\lambda'$. Then, $\widehat{E}^H_{\bar{s}}(\lambda') = \widehat{E}^H_{\bar{s}}(\lambda'')$. Consequently, note that $\forall j = 1, \dots, \kappa, \lambda_1 \leq \lambda''_j \leq \lambda_\kappa$ We can make two observations: first $\lambda''_1 \geq \lambda_1$ and $\lambda''_\kappa \leq \lambda_\kappa$, thus $\lambda''_\kappa - \lambda''_1 \leq \lambda_\kappa - \lambda_1$. It means that the maximum length difference between intervals in the newly created solution doesn't increase compared to the original solution. The second observation is that $\lambda'_1 > \lambda_1$ and $\lambda'_\kappa < \lambda_\kappa$. This ensures that after at least $\frac{n}{2}$ iterations, the maximum length difference strictly decreases. Therefore we can iterate this process with the new extreme intervals ($\lambda''_1$ and $\lambda''_\kappa$) until we get a solution whose extreme intervals lengths differ by at most 1, in which case all intervals differ by at most 1. This is then a balanced solution, hereby the proof is completed. □

**Proposition 3.** $\lambda^B \in \arg\min\{E^H_{\bar{s}}(\lambda), \lambda \in \Lambda(H)\}$.

*Proof.* Using Lemmas 1 and 4, Equation (6) and $\forall \lambda \in \Lambda(H)$, we get:
$E^H_{\bar{s}}(\lambda) \geq \widehat{E}^H_{\bar{s}}(\lambda) \geq \widehat{E}^H_{\bar{s}}(\lambda^B) = E^H_{\bar{s}}(\lambda^B)$. □

We then conclude:

**Theorem 1.** $K^B \in \arg\min\{E^{P_{n+1}}(K), K \in \mathcal{K}(P_{n+1})\}$

*Proof.* As $\lambda^B \in \arg\min\{E^{P_{n+1}}_s(\lambda), \lambda \in \Lambda(P_{n+1})\}$ (Proposition 2 ), and $\lambda^B \in \arg\min\{E^{P_{n+1}}_{\bar{s}}(\lambda), \lambda \in \Lambda(P_{n+1})\}$ (Proposition 3 ), then $\lambda^B \in \arg\min\{E^{P_{n+1}}(\lambda), \lambda \in \Lambda(P_{n+1})\}$ (see Equation (4)). As $\lambda^B$ corresponds to a unique solution $K^B$ in the case of paths, the proof is complete. □

### 3.4 The case of cycles

**Proposition 4.** $K^B \in \arg\min\{E^{C_n}(K), K \in \mathcal{K}(C_n)\}$

*Proof.* (Sketch) Given an instance $C_n$ of our problem with $k$ centers, we match an instance $P_{n+1}$ of Probabilistic $k$-Center with $k+1$ centers with the extremity nodes $(1, n+1)$ part of any feasible solution $K_P$ of $P_{n+1}$. Thus, to any solution $K_P \in \mathcal{K}(P_{n+1})$ we can match a solution $K_C \in \mathcal{K}(C_n)$ of size $k$. Since it is the lengths of the segments induced by $K_C$ that distinguish a solution from another on $C_n$, $K_C$ matches $K_P$ if they both induce a series of $k$ segments of same lengths. We assume then, without loss of generality for $C_n$, that

$\lambda_1$ is the length of the shortest segment. The values of the solutions $K_C$ and $K_P$ verify the following relations: $E^{C_n}(K_C) = \frac{n+1}{n}\left(E^{P_{n+1}}(K_P) - \min\{r_1^{K_P}, r_{n+1}^{K_P}\}\right)$ and $\widehat{E}^{C_n}(K_C) = \frac{n+1}{n}\left(\widehat{E}^{P_{n+1}}(K_P) - (\lambda_1 - 1)\right)$. By Theorem 1, we know that a balanced solution $K_P^B$ minimizes $\widehat{E}^{P_{n+1}}(K_P)$. In addition $K_P^B$ maximizes the value of the segment of minimum length, $\lambda_1^B$. Then $K_C^B$, the solution on $C_n$ matching $K_P^B$, minimizes $\widehat{E}^{C_n}(K), \forall K \in \mathcal{K}(C_n)$. As $E^{C_n}(K_C^B) = \widehat{E}^{C_n}(K_C^B)$, Lemma 1 and Equation (6) induce $E^{C_n}(K^B) = \widehat{E}^{C_n}(K^B) \leq \widehat{E}^{C_n}(K) \leq E^{C_n}(K)$ for all $K \in \mathcal{K}(C_n)$. It concludes the proof. □

## 4 Complexity result

The case of planar graphs and in particular with low degree, is very natural for our application. It motivates us investigating the complexity status of our problem in restricted classed of planar graphs to better discriminate polynomial cases and hard cases.

Minimum Dominating Set is shown NP-hard on planar graphs of maximum degree 3 in [25]. More precisely the authors refer to a private communication by David S. Johnson and give the reduction but not the complete proof. The following lemma allows to prove it and will be required later.

The reduction is from Minimum Vertex Cover in planar graphs of maximum degree 3 [**?**]. This does not imply immediately the hardness of our problem. Indeed, we defined the Probabilistic $k$-Center as the case with fixed uniform probabilities $\frac{1}{|V|}$, while $k$-Center can only be seen as the specific case where the probabilities are all zeros. In this section, we will show an inapproximability result for Probabilistic $k$-Center on planar graphs of maximum degree 3 .

Given a planar graph $G = (V, E)$, one builds a graph $G' = (V', E')$ by replacing each edge $uv$ by a $C_4$ $a_{uv}b_{uv}c_{uv}d_{uv}$, linking $a_{uv}$ and $c_{uv}$ to $u$ and $v$, respectively.

**Lemma 5.** *For any $t \leq |V|$, $G$ has a vertex cover of size $t$ if and only if $G'$ has a dominating set of size $t + |E|$, i.e., a $(t + |E|)$-center of radius at most 1. Moreover, for each edge $uv \in E$ this $(t + |E|)$-center has exactly one vertex in $\{a_{uv}, c_{uv}\}$ and none in $\{b_{uv}, d_{uv}\}$. The transformation is polynomial.*

*Proof.* Consider $U \subset V$ a vertex cover of $G$ of size $t$. We define the set $U_E$ as follows. For every edge $uv \in E$, if $v \notin U$, we add $c_{uv}$ to $U_E$. If both $u$ and $v$ are in $U$, then we add either $a_{uv}$ or $c_{uv}$ to $U_E$. Then, $U \cup U_E$ is a dominating set of size $t + |E|$ in $G'$.

Assume conversely that $G'$ has a dominating set $K' \subset V'$ of cardinality $t + |E|$. For every edge $uv \in E$, we necessarily have $K' \cap \{a_{uv}, b_{uv}, c_{uv}, d_{uv}\} \neq \emptyset$ to cover vertices $b_{uv}$ and $d_{uv}$. In the meanwhile, it is never interesting to take $b_{uv}$ or $d_{uv}$, since it would always be possible to modify a solution using $b_{uv}$ and/or $d_{uv}$ into a solution using none them and of the same size. Thus we can always transform $K'$ into a dominating set satisfying $\forall uv \in E, |K' \cap \{a_{uv}, c_{uv}\}| = 1, K' \cap \{b_{uv}, d_{uv}\} = \emptyset$. It remains to prove that $K' \cap V$ is a vertex cover of $G$. Consider an edge $uv \in E$ and the unique vertex $x \in K' \cap \{a_{uv}, c_{uv}\}$. Since $K'$ is a dominating set, if $x = a_{uv}$, then $v \in K' \cap V$ and if $x = c_{uv}$, then $u \in K' \cap V$. This completes the proof. □

Consider a planar graph $G$ and assume it does not have pending vertices (vertices of degree 1), then using the previous construction, $G'$ has no pending vertex. Assume $G$ has a minimal vertex cover of size $t$, we define $k_t = t + |E|$ and consider a dominating set $K'$ of size $k_t$ in $G'$, with exactly one vertex in $\{a_{uv}, c_{uv}\}$ and none in $\{b_{uv}, d_{uv}\}$ for any edge $uv \in E$. Seeing $K'$ as a $k_t$-center, we now evaluate the related probabilistic radius for any scenario in the graph $G'$.

**Lemma 6.** *Using the above notations we have:*

$$r_u^{K'} = \begin{cases} 1 & \Leftrightarrow \quad u \in V \text{ and not protected} \\ 2 & \quad\quad \textit{else} \end{cases}$$

*Proof.* Since $G'$ is triangle-free with no pending vertex, and $K'$ is a dominating set, then we have: $\forall u \in V', r_u^{K'} \leq 2$. Consider first $u \in V$ and denote by $v_1, \ldots, v_d$ all neighbors of $u$ in $G$. Suppose $u$ is not protected ($u \notin K'$). By definition of $K'$, $\{v_1, \ldots, v_d\} \subset K'$ and thus $\{a_{uv_1}, \ldots, a_{uv_d}\} \subset K'$, so whatever the escaping direction, people located on $u$ will reach a center at distance 1. Since $K'$ is a dominating set in $G'$, it remains a dominating set in $G' \setminus \{u\}$, which proves $r_u^{K'} = 1$.

Suppose now $u \in V \cap K'$. Since the considered vertex cover is minimal, there is $j \in \{1, \ldots, d\}$ such that $v_j \notin K'$ and thus $K' \cap \{a_{uv_j}, b_{uv_j}, c_{uv_j}, d_{uv_j}\} = \{c_{uv_j}\}$. Then, the evacuation distances of $a_{uv_j}$ is 2, and thus $r_u^{K'} = 2$.

Suppose now $u \in \{b_{vw}, d_{vw}\}$ for $vw \in E$. Then $u \notin K'$ and only one neighbor of $u$ is in $K'$, inducing an evacuation distance of 2 for $u$.

Suppose finally $u \in \{a_{vw}, c_{vw}\}$ for $vw \in E$. If $u \notin K'$ exactly one of its three neighbors is in $K'$ and its evacuation distance is also 2. If finally $u \in K'$, the evacuation distance of $b_{vw}$ and $d_{vw}$ becomes 2. So, in all cases but the first one $r_u^{K'} = 2$ and the proof is complete. □

Our proof will require another transformation and the property mentioned in the next lemma. Consider a planar graph, $G = (V, E)$ of degree at most 3. For a given $q \in \mathbb{N}, q \geq 2$, we transform $G$ into $\widetilde{G}_q = (\widetilde{V}_q, \widetilde{E}_q)$ as follows. We choose randomly an orientation of edges in $E$ and we replace every edge $uv \in E$ oriented from $u$ to $v$ by the path $\widetilde{P}_{uv}^q = \{u, x_{uv}^1, x_{uv}^2, \ldots, x_{uv}^{2q}, v\}$. Note that $|\widetilde{V}_q| = |V| + 2q|E|$ and $\widetilde{E}_q = (2q+1)|E|$.

**Lemma 7.** *For any $t \leq |V|$, $G = (V, E)$ has a minimum vertex cover of size $t$ if and only if $\widetilde{G}_q = (\widetilde{V}_q, \widetilde{E}_q)$ has a minimum vertex cover of size $t + q|E|$.*

*Proof.* Assume first $U \subset V$ is a minimum vertex cover of size $t$ in $G$: $\forall uv \in E, \{u, v\} \cap U \neq \emptyset$. Then we build $\widetilde{U}_q \subset \widetilde{V}_q$ in $\widetilde{G}_q$ as follows. We initialize $\widetilde{U}_q$ with all vertices of $U$. Then, for every edge $uv \in E$, if $u \in U$, we add vertices $x_{uv}^{2i}, 1 \leq i \leq q$ to $\widetilde{U}$. If $u \notin U$ (then $v \in U$), we add vertices $x_{uv}^{2i+1}, 0 \leq i \leq q-1$. In both cases we have added exactly $q$ vertices and all edges of $\widetilde{P}_{uv}^q$ are covered. Thus, $|\widetilde{U}_q| = t + q|E|$. $\widetilde{U}_q$ is minimum because we need the $t$ vertices of the set $U$ to cover at least $|E|$ edges $ux_{uv}^1$, and we need at least $q|E|$ different vertices to cover $\{x_{uv}^1, x_{uv}^2, \ldots, x_{uv}^{2q} : \forall uv \in E\}$.

Assume now that $\widetilde{G}_q$ has a minimum vertex cover $\widetilde{U}_q$ of size $t + q|E|$. For every $uv \in E$, $\widetilde{P}_{uv}^q$ is covered by at least $q+1$ vertices. If $u, v \notin \widetilde{U}_q$, we can transform $\widetilde{U}_q$ into $\widetilde{U'}_q$ such that $u$ or $v$ is in $\widetilde{U'}_q$. Then $|(\widetilde{V}_q \setminus V) \cap \widetilde{U'}_q| = q|E|$ and thus $|V \cap \widetilde{U'}_q| = t$. It remains to prove that $U = V \cap \widetilde{U'}_q$ is a minimum vertex cover. $U$ is a vertex cover because if $u \in \widetilde{U'}_q$, then $\widetilde{P}_{uv}^q$ is covered in $\widetilde{G}_q$, thus $uv$ is covered in $G$. If $U$ is not minimum, then suppose $U^*$ is a minimum vertex cover of size $t^* < t$. Then by the transformation given in paragraph 1 we can get a vertex cover for $\widetilde{G}_q$ of size $t^* + q|E| < |\widetilde{U}_q|$ which contradicts our initial hypothesis. Thus $U$ is minimum. □

The last Lemma we will need for our hardness proof is certainly a known remark but we show it since we did not find any reference for it.

**Lemma 8.** Minimum Vertex Cover *is NP-hard in planar graphs with vertices of degree 2 or 3.*

*Proof.* The decision version of Minimum Vertex Cover is known to be NP-complete on planar graphs of maximum degree 3 [25]. Consider a planar graph $G$ of maximum degree 3 and with a pending vertex $v$. Consider the graph $G'$ obtained from $G$ by adding a triangle and linking one of its vertices with $v$ ($v$ is then of degree 2 in $G'$). $G'$ is planar with maximum degree 3 and one pending vertex less than $G$. Moreover, $G$ has a vertex cover of size $t$ if and only if $G'$ has a minimum vertex cover of size $t + 2$, which concludes the proof. □

We now are ready to prove the main result of this section. Recall that we consider in this paper Probabilistic $k$-Center under a uniform probability distribution. Note that if $k$ was a fixed constant, the number of $k$-centers would be polynomial and the problem itself could be polynomially solved on any graph. So, we assume that $k$ is part of the instance.

**Theorem 2.** *There is no polynomial time approximation for* Probabilistic $k$-Center *guaranteeing a ratio less than $\frac{20}{19}$ for planar graphs of degrees 2 or 3, unless P=NP.*

*Proof.* The proof is by contradiction. Let $\rho$ satisfy $1 < \rho < \frac{20}{19}$. Consider $\varepsilon > 0$ such that $\rho < \frac{20+2\varepsilon}{19+2\varepsilon} < \frac{20}{19}$. Take $q \in \mathbb{N}$, such that $\frac{5}{q} \leq \varepsilon$ and $q \geq 2$.

We suppose we have a polynomial approximation algorithm `A` for Probabilistic $k$-Center, admitting as argument a planar graph $H$ of degrees 2 or 3 and the number $k$ of centers, and guaranteeing the approximation ratio $\rho$. We then will show how to use this algorithm to solve the Minimum Vertex Cover problem on planar graphs with vertex degrees 2 or 3. Lemma 8 will give the contradiction, unless P=NP.

Consider a planar graph $G = (V, E)$ with vertex degrees in $\{2, 3\}$, instance of Minimum Vertex Cover. Consider the graph $\widetilde{G}_q = (\widetilde{V}_q, \widetilde{E}_q)$ and then $\widetilde{G}'_q = (\widetilde{V}'_q, \widetilde{E}'_q)$. We have in particular:

$$
\begin{array}{rcl}
|\widetilde{V}_q| & = & |V| + 2q|E| \\
|\widetilde{E}_q| & = & (2q+1)|E| \\
|\widetilde{V}'_q| & = & |\widetilde{V}_q| + 4|\widetilde{E}_q| \\
& = & |V| + (10q+4)|E|
\end{array} \tag{11}
$$

Denoting by $\tau(H)$ and $\eta(H)|$ the minimum size of a vertex cover and a dominating set in a graph $H$, respectively, we deduce from Lemmas 5 and 7:

$$
\begin{array}{rcl}
\tau(\widetilde{G}_q) & = & \tau(G) + q|E| \\
\eta(\widetilde{G}'_q) & = & \tau(\widetilde{G}_q) + |\widetilde{E}_q| \\
& = & \tau(G) + (3q+1)|E|
\end{array} \tag{12}
$$

We apply the hypothetical approximation algorithm `A` on $\widetilde{G}'_q$ for different values of $k$, starting with $k = 1$ and augmenting it.

Suppose first we use $k < \eta(\widetilde{G}'_q)$ centers and the algorithm computes a center $\widetilde{K}''$. Then, the non-probabilistic radius is at least 2 since $\widetilde{K}''$ cannot be a dominating set in $\widetilde{G}'_q$ and consequently:

$$
\forall u \in \widetilde{V}'_q, r_u^{\widetilde{K}''} \geq 2 \Longrightarrow E(\widetilde{K}'') \geq 2 \tag{13}
$$

Suppose now we use $k = \eta(\widetilde{G}'_q) = \tau(\widetilde{G}_q) + |\widetilde{E}_q|$ centers. Using Lemma 5 on graph $\widetilde{G}_q$, there is a $k$-center $\widetilde{K}'$ of graph $\widetilde{G}'_q$ satisfying the conditions of Lemma 6 and then, this Lemma ensures:

$$
|\widetilde{V}'_q|E(\widetilde{K}') = 2|\widetilde{V}'_q| - (|\widetilde{V}_q| - \tau(\widetilde{G}_q))
$$

We deduce, using Relations 11 and 12:

$$
\begin{array}{rcl}
|\widetilde{V}'_q|E(\widetilde{K}') & = & |V| + (19q+8)|E| + \tau(G) \\
& < & 2|V| + (19q+8)|E|
\end{array}
$$

where the last inequality holds because $\tau(G) < |V|$. So, we have:

$$
E(\widetilde{K}') < \frac{2|V| + (19q+8)|E|}{|V| + (10q+4)|E|} = 2 - \frac{q|E|}{|V| + (10q+4)|E|}
$$

Since $G$ has vertices of degree at least 2, we have $V \leq |E|$, thus:

$$
E(\widetilde{K}') < 2 - \frac{q|E|}{(10q+5)|E|} = 2 - \frac{q}{10q+5} \leq \frac{19+2\varepsilon}{10+\varepsilon}
$$

where the last inequality holds since $\frac{5}{q} \leq \varepsilon$. As a consequence, since an optimal probabilistic solution $\widetilde{K}'^*$ will satisfy $E(\widetilde{K}'^*) \leq E(\widetilde{K}')$, the approximation algorithm `A` will determine an approximated center $K'$ in $\widetilde{G_q}'$ of value:

$$
\begin{array}{rcl}
E(\widetilde{K}') & \leq & \rho \times \frac{19+2\varepsilon}{10+\varepsilon} \\
& < & \frac{20+2\varepsilon}{19+2\varepsilon} \times \frac{19+2\varepsilon}{10+\varepsilon} = 2
\end{array} \tag{14}
$$

Note that, given a center $\widetilde{K}'$, computing its probabilistic radius can be done in polynomial time since, for any vertex $v \in \widetilde{V}'_q$, computing $r_v^{\widetilde{K}'}$ can be performed using any minimum path algorithm. So, we apply successively the approximation algorithm `A` on the graph $\widetilde{G}'_q$ for increasing values of $k$, starting with $k = 1$, until the computed center $\widetilde{K}'$ satisfies $E(\widetilde{K}') < 2$. Equations 12, 13 and14 ensure that, the algorithm stops after $k = \eta(\widetilde{G}'_q) = \tau(\widetilde{G}_q) + |\widetilde{E}_q|$ iterations and then, $\tau(G) = |\widetilde{K}'| - (3q+1)|E|$. Since constructing $\widetilde{G}'_q$ and evaluating $E(\widetilde{K}')$ can be done in polynomial time, and since algorithm `A` will be run less than $|V|$ times, the whole process is polynomial. This is a contradiction if P≠NP, and the proof is complete. □

## 5 Conclusion

To our knowledge, this paper introduces the first probabilistic version of Minimum $k$-Center. As illustrated in this work, defining the problem already leads to interesting discussions. So far, we have considered only single node disruption scenarios under a uniform distribution of probabilities and for a modification strategy that preserves the location of the centers but affects only the allocation. Even though relatively restrictive, this version is natural for our application and is already non-obvious. Then, we propose an explicit solution on paths and cycles. The main idea is to express the objective function as the sum of two parts (contribution of the skeleton and the body) and then prove independently that the solution minimizes simultaneously both terms. This approach might be used in a more general setting. Finally we prove that this restrictive version is already NP-hard to approximate it within a factor $\frac{20}{19}$ on planar graphs of bounded degree 3.

This motivates investigating the complexity status of this version on more restrictive classes of graphs, in particular on subgrids (subgraphs of grids), that correspond to some real case applications. Minimum $k$-Center is polynomial on trees [25]; it motivates studying the complexity of Probabilistic $k$-Center on trees, which, to our knowledge, remains open. The next question will be to design approximation algorithms for the hard cases and a third objective will be to consider larger ranges of impact and any probabilities. Indeed, our analysis on paths and cycles cannot be immediately extended to non uniform probability systems.

## References


[1] D. L. Martell, "A review of recent forest and wildland fire management decision support systems research," *Current Forestry Reports*, vol. 1, no. 2, pp. 128–137, 2015.

[2] J. P. Minas, J. W. Hearne, and J. W. Handmer, "A review of operations research methods applicable to wildfire management," *International Journal of Wildland Fire*, vol. 21, no. 3, pp. 189–196, 2012.

[3] D. Bertsimas, *Probabilistic combinatorial optimization problems.* PhD thesis, Massachusetts Institute of Technology, 1988.

[4] P. Jaillet, *Probabilistic traveling salesman problems.* PhD thesis, Massachusetts Institute of Technology, 1985.

[5] C. Murat and V. T. Paschos, *Probabilistic combinatorial optimization on graphs.* John Wiley & Sons, 2013.

[6] H. Calik, M. Labbé, and H. Yaman, "p-Center Problems," in *Location Science* (G. Laporte, S. Nickel, and F. Saldanha da Gama, eds.), pp. 79–92, Cham: Springer International Publishing, 2015.

[7] I. Averbakh, O. Berman, and D. Simchi-Levi, "Probabilistic a priori routing-location problems," *Naval Research Logistics (NRL)*, vol. 41, no. 7, pp. 973–989, 1994.

[8] D. J. Bertsimas, "Traveling salesman facility location problems," *Transportation science*, vol. 23, no. 3, pp. 184–191, 1989.

[9] D. J. Bertsimas, "The probabilistic minimum spanning tree problem," *Networks*, vol. 20, no. 3, pp. 245–275, 1990.

[10] P. Jaillet, "Shortest path problems with node failures," *Networks*, vol. 22, no. 6, pp. 589–605, 1992.

[11] P. Jaillet and A. R. Odoni, *The probabilistic vehicle routing problem*, vol. Vehicle Routing: Methods and Studies, pp. 293–318. North Holland, Amsterdam 1988.

[12] P. Balaprakash, M. Birattari, T. Stützle, and M. Dorigo, "Estimation-based metaheuristics for the probabilistic traveling salesman problem," *Computers & Operations Research*, vol. 37, no. 11, pp. 1939–1951, 2010.

[13] L. Bianchi, J. Knowles, and N. Bowler, "Local search for the probabilistic traveling salesman problem: Correction to the 2-p-opt and 1-shift algorithms," *European Journal of Operational Research*, vol. 162, no. 1, pp. 206–219, 2005.

[14] M. Birattari, P. Balaprakash, T. Stützle, and M. Dorigo, "Estimation-based local search for stochastic combinatorial optimization using delta evaluations: a case study on the probabilistic traveling salesman problem," *INFORMS Journal on Computing*, vol. 20, no. 4, pp. 644–658, 2008.

[15] A. M. Campbell, "Aggregation for the probabilistic traveling salesman problem," *Computers & Operations Research*, vol. 33, no. 9, pp. 2703–2724, 2006.

[16] C. Murat and V. T. Paschos, "A priori optimization for the probabilistic maximum independent set problem," *Theoretical Computer Science*, vol. 270, no. 1-2, pp. 561–590, 2002.

[17] C. Murat and V. T. Paschos, "The probabilistic minimum vertex-covering problem," *International Transactions in Operational Research*, vol. 9, no. 1, pp. 19–32, 2002.

[18] C. Murat and V. T. Paschos, "The probabilistic longest path problem," *Networks*, vol. 33, no. 3, pp. 207–219, 1999.

[19] V. T. Paschos, O. A. Telelis, and V. Zissimopoulos, "Steiner forests on stochastic metric graphs," *Lecture Notes in Computer Science*, vol. 4616, p. 112, 2007.

[20] V. T. Paschos, O. A. Telelis, and V. Zissimopoulos, "Probabilistic models for the steiner tree problem," *Networks*, vol. 56, no. 1, pp. 39–49, 2010.

[21] N. Boria, C. Murat, and V. Paschos, "On the probabilistic min spanning tree problem," *Journal of Mathematical Modelling and Algorithms*, vol. 11, no. 1, pp. 45–76, 2012.

[22] N. Boria, C. Murat, and V. T. Paschos, "The probabilistic minimum dominating set problem," *Discrete Applied Mathematics*, 2016.

[23] C. Murat and V. T. Paschos, "Probabilistic optimization in graph-problems," *Algorithmic Operations Research*, vol. 5, no. 1, pp. 49–64, 2010.

[24] R. L. Church and M. P. Scaparra, "Protecting critical assets: the r-interdiction median problem with fortification," *Geographical Analysis*, vol. 39, no. 2, pp. 129–146, 2007.

[25] O. Kariv and S. L. Hakimi, "An algorithmic approach to network location problems. I: The $p$-centers," *SIAM Journal on Applied Mathematics*, vol. 37, no. 3, pp. 513–538, 1979.

# Appendix

1. **Proof of Proposition 1**

   **Proposition 1** *For $k \geq \frac{n}{2}$, $K^B$ is an optimum solution for* Probabilistic $k$-Center *on paths and cycles.*

   *Proof.* In $H$, at least one vertex out of two consecutive vertices is in $K^B$. Therefore $\lambda_i^B \leq 2, \forall i \in 1, \ldots, \kappa$, and by Equations (7) and (8) we can state that $r_j^{K^B} = 1, \forall j \in V$. Then $E^H(K^B) = pn$ (Equation (3)). Suppose $K$ a non balanced but optimum solution on $H$. As $K$ is non balanced, $\exists i \in 1, \ldots, \kappa : \lambda_i \geq 3$. Then for $j \in \mu_i, r_j^K \geq 2$ (see Equations (7) and (8)). And for $j \in H \setminus \{\mu_i\}, r_j^K \geq \lfloor \frac{\lambda_i}{2} \rfloor \geq 1$. Then $\sum\limits_{j \in V} r_j^K > n$ and $E^H(K) > pn$, which is contradictory with the hypothesis. Thus a non balanced solution can't be optimum on $H$ for $k \geq \frac{n}{2}$ and the lemma is proven. □

2. **Proof of Lemma 1**

   **Lemma** 1 $E_s^H(\lambda^B) = \widehat{E}_s^H(\lambda^B)$ , $E_{\bar{s}}^H(\lambda^B) = \widehat{E}_{\bar{s}}^H(\lambda^B)$ and $E^H(\lambda^B) = \widehat{E}^H(\lambda^B)$.

   *Proof.* In $H$ we have: $\lambda_q^B \in \{\lfloor \frac{n}{\kappa} \rfloor, \lceil \frac{n}{\kappa} \rceil\}, q \in \{1, \ldots, \kappa\}$. As $k < \frac{n}{2}$, then $\max\limits_{q \in \{1,\ldots,\kappa\}} \lfloor \frac{\lambda_q^B}{2} \rfloor = \lfloor \frac{1}{2} \lceil \frac{n}{\kappa} \rceil \rfloor \leq \lfloor \frac{n}{\kappa} \rfloor - 1 \leq \lambda_q^B - 1, \forall q = 1, \ldots, \kappa$. The same holds for a balanced solution $\lambda^B$ in with $\lambda_q^B \in \{\lfloor \frac{n}{k} \rfloor, \lceil \frac{n}{k} \rceil\}$. Then given Equation (7), we obtain $r_j^{K^B} = \widehat{r}_j^{K^B} : \forall j \in K$, thus $E_s^H(\lambda^B) = \widehat{E}_s^H(\lambda^B)$. We also observe that $\forall i = 1, \ldots, \kappa, \forall j \in \mu_i^B \setminus K, \delta_j(j, K^B) \geq \max\{j - v_i, v_{i+1} - j\} \geq \left\lceil \frac{\lambda_i^B}{2} \right\rceil \geq \lceil \frac{1}{2} \lfloor \frac{n}{\kappa} \rfloor \rceil \geq \lfloor \frac{1}{2} \lceil \frac{n}{\kappa} \rceil \rfloor = \max\limits_{q \in \{1,\ldots,\kappa\}} \lfloor \frac{\lambda_q^B}{2} \rfloor$. Then given Equation(8), we obtain $r_j^{K^B} = \widehat{r}_j^{K^B} : \forall j \in H \setminus K$, thus $E_{\bar{s}}^H(\lambda^B) = \widehat{E}_{\bar{s}}^H(\lambda^B)$. Obviously then $E_s^H(\lambda^B) + E_{\bar{s}}^H(\lambda^B) = \widehat{E}_s^H(\lambda^B) + \widehat{E}_{\bar{s}}^H(\lambda^B)$, therefore $E^H(\lambda^B) = \widehat{E}^H(\lambda^B)$. □