The “sidekick” routing paradigm for VMT reduction and improved accessibility

July 2021

A Research Report from the Pacific Southwest Region University Transportation Center

John Gunnar Carlsson, University of Southern California

Bo Jones, University of Southern California

Pacific Southwest Region UTC

University Transportation Center

USC University of Southern California

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. PSR-MT-19-20	**2. Government Accession No.** N/A	**3. Recipient’s Catalog No.** N/A
4. Title and Subtitle The “sidekick” routing paradigm for VMT reduction and improved accessibility		**5. Report Date** 07/01/2021
		6. Performing Organization Code N/A
7. Author(s) John Gunnar Carlsson		**8. Performing Organization Report No.** TBD
9. Performing Organization Name and Address METRANS Transportation Center University of Southern California University Park Campus, RGL 216 Los Angeles, CA 90089-0626		**10. Work Unit No.** N/A
		11. Contract or Grant No. USDOT Grant 69A3551747109
12. Sponsoring Agency Name and Address U.S. Department of Transportation Office of the Assistant Secretary for Research and Technology 1200 New Jersey Avenue, SE, Washington, DC 20590		**13. Type of Report and Period Covered** Final report (01/01/2020 – 06/30/21)
		14. Sponsoring Agency Code USDOT OST-R
15. Supplementary Notes		

16. Abstract

This project has combined tools from geospatial analysis, mathematical optimization theory, and computational geometry to study a routing paradigm that we call *sidekick routing*. A sidekick routing scheme is a logistical framework in which a large vehicle, such as a truck or van, serves as a mobile base for a fleet of small vehicles (the “sidekicks”), such as autonomous ground vehicles (AGVs) or unmanned aerial vehicles (UAVs). Systems of this kind have significant potential to simultaneously reduce vehicle miles travelled (VMT) – because the sidekicks are not restricted to streets – and to improve accessibility to goods, because the so-called “last mile” cost of transporting those goods is reduced.

The sidekick paradigm has very recently seen use in many public and private sector organizations, both in California and elsewhere. However, although the requisite *physical* technology is reasonably mature, the requisite *management* technology (i.e. systems for determining efficient routing strategies) are relatively nascent. Moreover, the extent to which such services can provide a societal benefit are not yet understood, although the results from this report indicate that the potential is very high.

17. Key Words Vehicle routing; unmanned ground vehicles; UGVs; unmanned aerial vehicles; UAVs		**18. Distribution Statement** No restrictions.	
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages** 48	**22. Price** N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Contents

About the Pacific Southwest Region University Transportation Center

The Pacific Southwest Region University Transportation Center (UTC) is the Region 9 University Transportation Center funded under the US Department of Transportation’s University Transportation Centers Program. Established in 2016, the Pacific Southwest Region UTC (PSR) is led by the University of Southern California and includes seven partners: Long Beach State University; University of California, Davis; University of California, Irvine; University of California, Los Angeles; University of Hawaii; Northern Arizona University; Pima Community College.

The Pacific Southwest Region UTC conducts an integrated, multidisciplinary program of research, education and technology transfer aimed at *improving the mobility of people and goods throughout the region.* Our program is organized around four themes: 1) technology to address transportation problems and improve mobility; 2) improving mobility for vulnerable populations; 3) Improving resilience and protecting the environment; and 4) managing mobility in high growth areas.

U.S. Department of Transportation (USDOT) Disclaimer

Disclosure

Principal Investigator, Co-Principal Investigators, others, conducted this research titled, “The “sidekick” routing paradigm for VMT reduction and improved accessibility” at the Epstein Department of Industrial and Systems Engineering in the Viterbi School of Engineering at the University of Southern California. The research took place from 01/01/20 to 06/30/21 and was funded by a grant from the USDOT in the amount of $96793. The research was conducted as part of the Pacific Southwest Region University Transportation Center research program.

Acknowledgements

The authors gratefully acknowledge the support of METRANS for funding this research.

Abstract

This project has combined tools from geospatial analysis, mathematical optimization theory, and computational geometry to study a routing paradigm that we call *sidekick routing*. A sidekick routing scheme is a logistical framework in which a large vehicle, such as a truck or van, serves as a mobile base for a fleet of small vehicles (the “sidekicks”), such as autonomous ground vehicles (AGVs) or unmanned aerial vehicles (UAVs). Systems of this kind have significant potential to simultaneously reduce vehicle miles travelled (VMT) – because the sidekicks are not restricted to streets – and to improve accessibility to goods, because the so-called “last mile” cost of transporting those goods is reduced.

The sidekick paradigm has very recently seen use in many public and private sector organizations, both in California and elsewhere. However, although the requisite *physical* technology is reasonably mature, the requisite *management* technology (i.e. systems for determining efficient routing strategies) are relatively nascent. Moreover, the extent to which such services can provide a societal benefit are not yet understood, although the results from this report indicate that the potential is very high.

The “sidekick” routing paradigm for VMT reduction and improved accessibility

Executive Summary

One of the more novel recent innovations in the logistics world, both in theory and in practice, is the use of small autonomous vehicles to facilitate last-mile delivery. One particular scheme that has received considerable recent attention is the a “sidekick” scheme, in which a large cargo truck acts as a mobile “base” that deploys smaller vehicles, such as drones or unmanned ground vehicles (UGVs). The sidekicks alternate between visiting the truck to pick up items and visiting the customers, and the overall objective is to determine a coordinated set of routes for all vehicles in order to optimize system efficiency, such as minimizing the time to completion or the vehicle miles travelled (VMT).

Although the hardware for these systems is fairly mature, the problem of determining efficient routes has not been considered until very recently. From the perspective of routing these systems pose an exceptionally difficult challenge due to the need to synchronize multiple vehicles that can all be traveling at the same time and at different speeds. We cannot consider vehicles’ routes separately as we must include the possibility of vehicles carrying other vehicles for periods of time and the need for intermittent meetings of vehicles at the same position at the same point in time. Thus we see that individual vehicles’ routes are highly interdependent, and any reasonable objective will be impacted by this interdependence, making the optimization very hard. Furthermore, the high-level attributes of these systems are not at all clear: how much more efficient can they be? When are they useful? What are the trade-offs inherent in such a scheme? In this report, we develop a continuous approximation model that estimates the improvements to efficiency that such a system provides, in the asymptotic limit as many demand points are drawn from a continuous probability distribution. Our analysis indicates under what circumstances these sidekicks can offer the most benefit, as a function of their speeds and the number of sidekicks available.

1 Introduction

One logistical paradigm that has received considerable attention in recent years is the *sidekick routing scheme*. A sidekick routing scheme is a logistical framework in which a large vehicle, such as a truck or van, serves as a mobile base for a fleet of small vehicles (the “sidekicks”), such as autonomous ground vehicles (AGVs) or unmanned aerial vehicles (UAVs). The sidekicks alternate between visiting the truck to pick up items and visiting the customers, and the overall objective is to determine a coordinated set of routes for all vehicles in order to optimize system efficiency, such as minimizing the time to completion or the vehicle miles travelled (VMT). A sketch of such a system is shown in Figure 1.

Schemes of this kind have been deployed by many public and private sector organizations very recently, as described below and pictured in Figure 2:

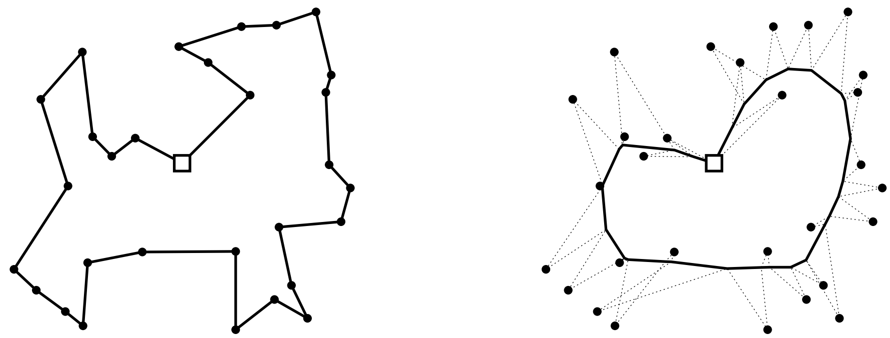
- The California-based startup companies Kiwi Campus, Dispatch (recently acquired by Amazon [29]), and Starship all use a hybrid system involving vans and ground vehicles to deliver food and groceries.

- The Ohio-based company AMP Electric Vehicles has introduced a system that they call the “horsefly” scheme, in which a drone flies back and forth with a delivery van. The California-based startup MatterNet has also partnered with Mercedes-Benz in designing a similar system for transporting blood samples; they call their system the “Vision Van”.

- The Swiss Post has been using a hybrid van-and-drone scheme to deliver goods from department stores free of charge within the Zürich city center since 2017 [9].

Although the hardware for these systems is fairly mature, the problem of determining efficient routes has not been considered until very recently. From the perspective of routing these systems pose an exceptionally difficult challenge due to the need to synchronize multiple vehicles that can all be traveling at the same time and at different speeds. We cannot consider vehicles’ routes separately as we must include the possibility of vehicles carrying other vehicles for periods of time and the need for intermittent meetings of vehicles at the same position at the same point in time. Thus we see that individual vehicles’ routes are highly interdependent, and any reasonable objective will be impacted by this interdependence, making the optimization very hard. Furthermore, the high-level attributes of these systems are not at all clear: how much more efficient can they be? When are they useful? What are the trade-offs inherent in such a scheme? We employ a continuous approximation analysis as a means of answering these questions.

This report is organized as follows. In Section 2 we provide an overview of related work. In Section 3 we formally define the sidekick routing problem. In Section 4 we introduce preliminary results that will be of use in our analysis. In Section 5 we derive our main result concerning the asymptotic behavior of the sidekick routing problem. In Section 6 we summarize the operational implications of this result. In particular, we consider how much improvement in efficiency can be gained by switching to the sidekick system and the tradeoffs that must be weighed in implementing the system. Finally, in Section 7 we consider the scaling behavior, and dependence on the configuration of the sidekicks, that our result tells us we should expect. We empirically demonstrate that actual tour times, obtained by heuristically solving the sidekick problem on simulated sets of customer points, corroborate our expectations.



(a) (b)

Figure 1: The figure on the left shows a *travelling salesman tour* of a set of client destinations and a central depot, that is, the shortest tour that visits a collection of points and starts and ends at the central depot. The figure on the right shows the solution to a “sidekick” problem in which the truck has a “helper” (such as a robot or a drone) that alternates between visiting the truck and visiting the customer locations. Note that the truck’s tour with helpers is about half the length of the tour with no helpers.

1.1 Remark on notational conventions

When it is necessary to specify that a map is injective, we use $\hookrightarrow$.

2 Related work

We begin by considering problems that pose similar fundamental questions to that of sidekick routing. We then look at applications of the continuous approximation paradigm to other combinatorial problems. Finally, we review past treatment of sidekick routing problems, comparing these papers’ model assumptions and contributions to those of this work.

One of the basic phenomena that is of interest to us is the trade-off between efficiency in transportation along a backbone network (in this case, the route of the truck) versus direct trips between locations (in this case, the direct trips taken by sidekicks); this is arguably one of the fundamental dichotomies in transportation and logistics [11, 12]. In this sense, our problem of interest is philosophically similar is [10], which asks whether small local retail stores are preferable to “big-box” retailers, with [81], which estimates the changes in net CO_2 emissions that result by introducing grocery delivery services, and with [82], which computes the optimal layout of a set of facility locations that are themselves connected with a backbone network.

this report is concerned with a continuous approximation model for a transportation problem, and is therefore philosophically similar to (for example) [8], which analytically determines trade-offs between transportation and inventory costs, [32], which shows how to route emergency relief vehicles to beneficiaries in a time-sensitive manner, and [33], which describes a simple geometric model for determining the optimal mixture of a fleet of vehicles that perform distribution. The basic premise of the continuous approximation paradigm is that one replaces combinatorial quantities that are difficult to compute with simpler mathematical formulas, which (under certain conditions) provide accurate estimations of the desired quantity. Such approximations exist for many combinatorial problems, such as the travelling salesman problem [3, 21], facility location [28, 30, 55], and any *subadditive Euclidean functional* such as a minimum spanning tree, Steiner tree, or matching [63, 68, 69]. Our particular usage of the continuous approximation paradigm is that we treat the customer demand as coming from a probability density function; in this sense our work also shares some commonality with *stochastic* vehicle routing problems [14, 24].

In 2018 Otto et al. compiled a comprehensive review, [54], of work on optimization approaches to systems employing drones for a wide range of applications, to include package delivery and specifically package delivery using drones as sidekicks for trucks. We cover much of the same territory to give a complete picture of the drone, and particularly sidekick, landscape and survey even more recent developments in the area.

We are primarily interested in prior work in the area of continuous approximation and theoretical results bounding the objective or characterizing the improvement due to sidekick introduction. For work on solving sidekick problems we concern ourselves principally with these papers' formulations of the problem. There are many variants, each differing in the assumptions that are made about the delivery system. Three critical questions, the answers to which change from model to model, that need to be posed are given below.

- Does the truck also deliver packages or are packages only delivered by the sidekicks?
- Can the truck carry multiple sidekicks capable of making simultaneous deliveries?
- Are the sidekick launch and pickup locations restricted to customer points, or otherwise to a discrete set of points that is specified a priori?

Table 1 provides a summary of how previous formulations have answered these questions and the work that was done on the resulting problem. We can see that our formulation has the least restrictive answers to these questions and thus addresses the problem in the greatest generality. That is, in our model we have the following.

- We allow for both the case that deliveries must be made by sidekicks and the case that the truck can also make deliveries.
- There can be any number of sidekicks on the truck and they are free to be launched and picked up in any order.
- The sidekick launch and pickup locations can be any point in the plane.

Other factors that distinguish the models surveyed here are the treatment of a restricted drone range and the way that the objective, be it completion time or some measure of energy consumed, is determined. This work assumes unlimited drone range and that drones make a single delivery per trip from the truck. We take as our objective completion time. We assume for simplicity that the time spent actually dropping a package at a customer node as well as the time spent capturing a sidekick and preparing it for relaunch are negligible. Both the truck and the sidekicks travel at fixed speeds along Euclidean distances. The specification of their relative speeds does however allow one to build some knowledge of the underlying network into the objective. One additional assumption that adds to the robustness of our formulation is that the sidekicks are allowed to be slower than the truck. We are thus able to accurately model systems like the truck-AGV schemes discussed in the introduction, whereas some papers surveyed require that the sidekicks be faster.

2.1 Other Autonomous Vehicle Delivery Schemes and Applications

A lot of work is being done in the area of optimal coordination of autonomous vehicles. Applications beyond just delivery include power network surveillance in response to extreme weather events [43], fire detection and extinguishing [76], facilitating intervehicle communications after disasters [35], and performing earth observation [80].

Other work has focused on drone-specific operational aspects of making deliveries. These include routing that considers the possibility of drones failing along their routes [73], routing that takes into account uncertain air temperature’s impact on battery duration [38], and depot location and route planning that takes into account battery consumption rates dependent on payload [74].

Examples of routing for delivery using a fleet of only autonomous vehicles can be found in [67] and [16]. Kim et al. [39] give a drone fleet routing problem for medical deliveries with the additional facet of depot location planning.

Heterogeneous vehicle delivery systems where the different types of vehicles work completely in parallel, both making deliveries from a depot rather than being synchronized, are also studied. Murray and Chu [51] give formulation of one such problem, the “Parallel Drone Scheduling Traveling Salesman Problem (PDTSP).” A truck makes a tour of some points while drones serve the others by single-customer trips from the depot. Murray and Chu provide a heuristic solution to this problem. Another solution approach can be found in [50]. Li et al. [42] use a continuous approximation model to derive the expected cost of a similar problem where trucks move first from a distribution center to depots that each serve a region. A PDTSP-like system is then used within the regions. In [71] drones provide an alternative delivery method to a cross-docking system. In [75] a Markov Decision Model is developed for the problem of meeting dynamic customer demand with either trucks or drones.

A problem that is similar to the sidekick problem in that it seeks to take advantage of the long range of a truck as well as the benefits of autonomous vehicles appears often in the literature as the Two Echelon Routing Problem. In it trucks start at a distribution center then take tours of, or direct routes to, secondary drone depots. Trucks bring packages or packages *and* drones. Drones then make deliveries from the secondary depots. In some cases trucks also visit customers on their tour. Variations of this system can be found in [66, 37, 53, 19].

There are a number of problems that involve tandem delivery that do not quite fall into our classification of a sidekick routing problem due to some restrictive assumptions. These assumptions are the truck moves along a linear course [64], the truck’s route is predetermined [4], the visit sequence is predetermined [23], and sidekicks must be picked up where they are launched [47, 20]. A nondelivery application of a sidekick system is presented in [72]. A ground robot and a drone are used in tandem to monitor nitrogen levels in an agricultural plot.

2.2 Exact Solutions to Sidekick Problems

In 2015 Murray and Chu [51] were the first to formally describe a sidekick system. They provide a mixed integer linear programming (MILP) formulation of the “Flying Sidekick Traveling Salesman Problem (FSTSP).” In it a single truck carries a single drone. The truck can, and for some customers must, make deliveries. Sidekick launch and pickup locations are limited to customer points and the depot. A time endurance range is imposed on the sidekicks. The objective is to minimize the time to complete all deliveries. Another similar, widely cited, integer programming (IP) formulation, the “Traveling Salesman Problem with Drone (TSP-D)” is given by Agatz et al. [1]. Ha et al. [27] introduce a cost based objective to these single-drone models. Jeong et al. [34] present a mixed integer programming (MIP) formulation extending this problem by letting drone range depend on payload and introducing no fly zones. Additional modifications to the MIP model can be found in [18] and [77]. Bouman et al. [5] give a dynamic programming solution to the problem.

Other extensions to the single-drone problem involve allowing for drone launch and pickup at a discrete set of points rather than, or in addition to, the customer points. Mathematical programming formulations of such problems can be found in [46] and [44]. Matthew et al. [49] present a solution to this problem without truck deliveries by way of a reduction to the generalized traveling salesman problem. Poikonen and Golden [57] go a step further in freeing launch and pickup locations. They formulate the “Mothership and Drone Routing Problem,” in which launches and pickups can occur anywhere in the plane and drones are allowed to make multiple deliveries per trip from the truck. The authors give a branch and bound method for solving this problem exactly.

Murray and Raj [52] extend the FSTSP model of [51] to allow for multiple heterogeneous sidekicks. They further develop a more sophisticated treatment of endurance that takes into account payloads. A MILP model is presented. Kitjacharoenchai et al. [40] give a MIP formulation for a more limited multi-drone problem in which only one drone launch or pickup can occur at each customer point. Karak and Abdelghany [36] formulate a MIP that extends the multi-sidekick model to allow for a discrete set of non-customer launch and pickup points. Boysen et al. [7] provide an MIP formulation for a multi-sidekick problem in which robots are launched from a predetermined launch sites and return to a predetermined set of robot depots. The truck can refill to its capacity of robots at any robot depot.

Another class of sidekick problems is sidekick vehicle routing problems (VRPs) involving fleets of multiple, often-capacitated, trucks and sidekicks. Wang et al. [78] provided the first formulation, the “Vehicle Routing Problem with Drones,” but their work is not concerned with exact solutions. Wang and Sheu [79] give a MIP formulation. MIP formulations of the problem with the addition of customer time windows can be found in [62] and [61].

All of these exact solutions are intractable on problems of practical size and are generally only able to solve problem instances up to size 10 customer points in a reasonable amount of time. Thus heuristic approaches are needed.

2.3 Heuristic Solutions to Sidekick Problems

Many of the papers given in the previous section also develop heuristic methods for the problems they pose. Murray and Chu [51] give a solution to the FSTSP that begins with a truck TSP tour and iteratively reassigns customers to drones. Likewise Agatz et al.’s [1] solution to the TSP-D begins with a TSP tour and then partitions the tour into truck customers and drone customers. Poikonen et al. [58] develop a heuristic method for solving the TSP-D that uses a branch and bound procedure to explore possible sequences of deliveries. Tang et al. [70] present a constraint programming formulation of the TSP-D. Other heuristic solutions to the versions of the problem with a single-drone and launch and pickup locations restricted to customer points can be found in [27, 26, 22, 45, 25, 60, 77, 34].

Heuristic solutions to the problem with the extension of launch and pickups allowed at a discrete, predetermined set of non-customer points are given in [46] and [44]. Marinelli et al. [48] provide a heuristic algorithm for the “en-route” truck-drone delivery system. In it the drone can be launched and picked up at any point along the lines that make up the truck’s tour of the customers that it visits. Poikonen and Golden [57] develop a heuristic for the Mothership and Drone Routing Problem which has totally free launch and pickup sites.

Murray and Raj give a heuristic solution to their FSTSP with multiple sidekicks model. Other heuristic solutions to multi-sidekick models with launch and pickup restricted to customer points can be found in [6] and [40].

Karak and Abdelghany [36] and Boysen et al. [7] heuristically solve their problems with multiple drones and a fixed set of non-customer launch and pickup sites. Poikonen [56] provides a heuristic solution to a new problem, the “Multi-Visit Drone Routing Problem.” A single truck can carry multiple drones, though their launch and pickup order is restricted to all must be launched then all must be picked up. Launch and pickup locations are a given discrete set. Most notably, a drone can deliver multiple packages per trip from the truck.

For heuristic algorithms to solve sidekick VRPs, Daknama and Kraus [17] assume launch and pickups at customers, Wang et al. [79] assume launch and pickups at discrete set in addition to customers, and Schermer et al. [65] assume launch and pickups at discrete locations along the direct paths the truck travels between truck-delivery points.

2.4 Theoretical Results and Continuous Approximation Models

Wang et al. [78] consider the Vehicle Routing Problem with Drones in which multiple vehicles each carry multiple drones. They derive upper bounds on the improvement to be gained over the optimal TSP and VRP solutions without drones as well as the improvement to be gained by introducing faster drones. Poikonen et al. [59] extend the model of [78]. A battery life (time limit) is imposed on the drones; the possibility of using different distance metrics for the truck and drone and the possibility of using cost rather than time based objectives are considered; and there is an extension to the close-enough vehicle routing problem. Their results are bounds on improvement due to introduction of drones and due to different drone configurations.

Agatz et al. [1] produce a result that is a generalization of the results of [78] when applied to the TSP-D. That is they give an upper bound on the improvement over just-truck routing allowing different distance metrics to be used for the truck and drone distances. The authors further give a lower bound to the TSP-D and an approximation algorithm using minimum spanning trees.

Campbell et al. [13] study a continuous approximation model for a sidekick problem with a truck carrying multiple drones. Demand is modeled as a continuous spatial density. Customer points are visited in rectangular swaths. The authors provide the expected cost of delivery in terms of the customer density and the truck and drone per-unit-distance and dropoff costs. Comparison is made to the expected cost without drones. Unlike in our model, drone launch and pickup locations are limited to customer points, and the sequence of deliveries is fixed to a truck delivery at which all drones are launched followed by another truck delivery at which all drones are picked up and relaunched.

In [15] Carlsson and Song consider the sidekick problem as formulated in this report except restricted to only one sidekick and assuming that the sidekick is faster than the truck. Using a continuous approximation model that assumes a smooth demand distribution they are able to derive the asymptotic behavior of the optimal tour as the number of customers goes to infinity. This then yields a characterization of the improvement to be gained by introducing a sidekick and how this improvement depends on the relative speeds of the truck and sidekick.

(a) BoxBot (b) Dispatch (c) Kiwi Campus

(d) HorseFly (e) Starship (f) Swiss Post (g) Mercedes Vision Van

Figure 2: Various hardware implementations of sidekick routing schemes.

Table 1: A summary of work on sidekick problems. Papers are classified by their model assumptions as well as their contributions.

Publication	TSP/VRP	Truck Delivers	Multiple Drones	Launch/Pickup Sites	Exact Solutions	Heuristic Algorithms	Continuous Approximation	Theoretical Results
[51], [27]	TSP	✓	-	Customer Points	MILP	✓	-	-
[25], [60], [26], [22], [45], [58]	TSP	✓	-	Customer Points	-	✓	-	-
[49]	TSP	-	-	Fixed Set	GTSP Reduction	-	-	-
[46]	TSP	-	-	Fixed Set	IP	✓	-	-
[57]	TSP	-	-	Completely Free	Branch & Bound	✓	-	-
[5]	TSP	✓	-	Customer Points	DP	-	-	-
[56]	TSP	-	✓	Fixed Set	-	✓	-	-
[48]	TSP	✓	-	Lines Between Truck Deliveries	-	✓	-	-
[1]	TSP	✓	-	Customer Points	IP	✓	-	Bounds on Improvement
[6]	TSP	-	✓	Customer Points	MIP	✓	-	-
[70]	TSP	✓	-	Customer Points	Constraint Programming	✓	-	-
[18]	TSP	✓	-	Customer Points	MILP	-	-	-
[52]	TSP	✓	✓	Customer Points	MILP	✓	-	-
[44]	TSP	-	-	Fixed Set	MIP	✓	-	-

Table 1: (continued)

Publication	TSP/VRP	Truck Delivers	Multiple Drones	Launch/Pickup Sites	Exact Solutions	Heuristic Algorithms	Continuous Approximation	Theoretical Results
[40]	TSP	✓	✓	Customer Points	MIP	✓	-	-
[36]	TSP	-	✓	Fixed Set	MIP	✓	-	-
[77], [34]	TSP	✓	-	Customer Points	MIP	✓	-	-
[78], [59]	VRP	✓	✓	Customer Points	-	-	-	Bounds on Improvement
[17]	VRP	✓	✓	Customer Points	-	✓	-	-
[62]	VRP	✓	✓	Customer Points	MIP	-	-	-
[79]	VRP	✓	✓	Fixed Set	MIP	✓	-	-
[65]	VRP	✓	✓	Path Between Truck Deliveries	-	✓	-	-
[61]	VRP	✓	1 Per Truck	Customer Points	MIQP	✓	-	-
[13]	TSP	✓	✓	Customer Points	-	-	✓	Expected Cost
[15]	TSP	✓	-	Completely Free	-	-	✓	Asymptotics, Improvement Over TSP
This Work	TSP	✓	✓	Completely Free	-	-	✓	Asymptotics, Improvement Over TSP

Pacific Southwest Region UTC
University Transportation Center

3 Problem definitions

We begin by formally defining the problem of sidekick routing with multiple sidekicks. We assume that a single, uncapacitated truck must provide service to a collection of n customers in the plane, using the assistance of k sidekicks having unit capacity, and that the goal is to minimize the time to completion. To simplify exposition, we will first formulate our problem with an additional constraint that the truck itself is not permitted to visit any customers:

Definition 1. Let $p_1, \ldots, p_n$ be a collection of points in the plane. Let k denote the number of sidekicks. Let ϕ_0 denote the speed of the truck, and let ϕ_1 denote the speed of each sidekick (ϕ_1 can be greater or less than ϕ_0). Let variables $x_1, \ldots, x_n$ be the launch points for the sidekicks, and let variables $y_1, \ldots, y_n$ be the pickup points for the sidekicks. That is, point p_i is visited by a sidekick that is launched at point x_i and is retrieved at point y_i.

Let variables z_j, $j \in \{1, \ldots, 2n\}$, be the location of the jth sidekick launch or pickup event, and let variables t_j, $j \in \{1, \ldots, 2n\}$, be the time of the jth sidekick launch or pickup event. The z_j's take the same values as the x_i's and the y_i's; we introduce them only to make indexing easier in the formulation. We let z_0 be the initial position of the truck and let z_{2n+1} be its final position. We require that the truck's tour be a loop, i.e. $z_{2n+1} = z_0$. We let t_0, equaling zero, be the time at which the truck starts its loop and let t_{2n+1} be the time at which the truck completes its loop.

Let $\sigma : \{1, \ldots, n\} \hookrightarrow \{1, \ldots, 2n\}$ map customer index i to the place of that customer's sidekick launch event in the ordering of all launch and pickup events. Similarly let $\pi : \{1, \ldots, n\} \hookrightarrow \{1, \ldots, 2n\}$ map customer index i to the place of that customer's sidekick pickup event in the ordering of all launch and pickup events. Let $\mathcal{F}$ be the set of all pairs of mappings (σ, π) that induce a valid sidekick tour. The conditions for inclusion in $\mathcal{F}$ are

$$
\begin{aligned}
&\sigma(i) < \pi(i) \quad \forall i \in \{1, \ldots, n\} && \text{(launches occur before corresponding pickups)}\\
&|\{i : \sigma(i) < j\}| - |\{i : \pi(i) < j\}| \leq k \quad \forall j && \text{(never more than } k \text{ sidekicks in use)}\\
&\sigma, \pi \text{ are injective} && \text{(a launch and a pickup for each customer)}\\
&\sigma(i) \neq \pi(i') \quad \forall i, i' \in \{1, \ldots, n\}. && \text{(one event per place in the ordering)}
\end{aligned}
$$

The last two conditions say that the maps σ and π have to jointly form a bijection between $\{1, \ldots, n\}$ and $\{1, \ldots, 2n\}$ (to be precise, the two maps actually form a bijection between the multiset $\{1, \ldots, n\} \uplus \{1, \ldots, n\}$ and $\{1, \ldots, 2n\}$, where $\uplus$ denotes the multiset union [41]).

The *sidekick problem* problem is then given by

$$
\underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{t},\sigma,\pi}{\text{minimize}} \quad t_{2n+1} \quad s.t. \tag{SK1}
$$

$$
t_j \geq t_{j-1} + \frac{1}{\phi_0}\|z_j - z_{j-1}\| \quad \forall j \in \{1, \ldots, 2n+1\} \tag{1}
$$

$$
t_{\pi(i)} \geq t_{\sigma(i)} + \frac{1}{\phi_1}\|x_i - p_i\| + \frac{1}{\phi_1}\|p_i - y_i\| \quad \forall i \in \{1, \ldots, n\} \tag{2}
$$

$$
\begin{aligned}
z_{\sigma(i)} &= x_i \quad \forall i \in \{1, \ldots, n\}\\
z_{\pi(i)} &= y_i \quad \forall i \in \{1, \ldots, n\}\\
t_0 &= 0\\
z_{2n+1} &= z_0\\
(\sigma, \pi) &\in \mathcal{F},
\end{aligned}
$$

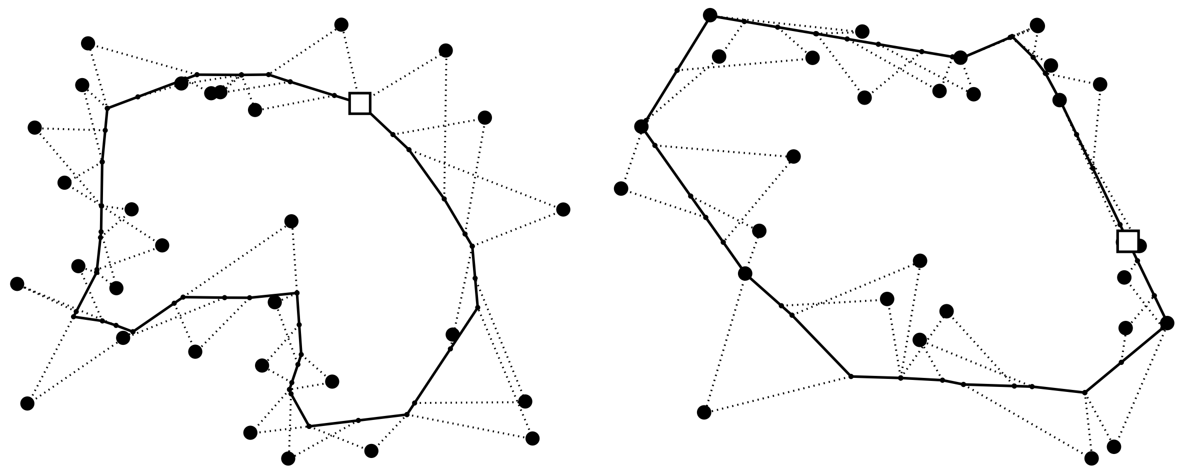


Figure 3: Solutions to the sidekick problem for $n = 30$ customers, $k = 3$ sidekicks, and a sidekick speed, ϕ_1, that is twice the speed, ϕ_0, of the truck. The solid line is the truck tour; dotted lines are the sidekicks' routes; the square represents the starting and finishing point of the truck. In the figure on the left all deliveries are made by the sidekicks (Problem SK1). In the figure on the right the truck is also allowed to make deliveries (Problem SK2).

where the objective value is the time at which the truck completes its loop, (1) captures the time needed for the truck to travel between launch and pickup points, and (2) captures the time needed for a sidekick to travel from its launch point, to a customer, and then to its pickup point.

To extend (SK1) to the case where the truck is permitted to visit customers, some additional notation is required:

Definition 2. We partition the set of customers into two sets $\mathcal{S} \subseteq \{1, \ldots, n\}$, representing those customers visited by a helper, and its complement $\mathcal{T} = \bar{\mathcal{S}}$, representing those customers visited by the truck (these sets are optimization variables because we can choose which customers are visited by the truck). The number of events is now equal to $m := 2|\mathcal{S}| + |\mathcal{T}|$ because a truck visiting a customer counts as only one event. This necessitates a third map $\theta : \mathcal{T} \hookrightarrow \{1, \ldots, m\}$, in addition to the maps $\sigma, \pi : \mathcal{S} \hookrightarrow \{1, \ldots, m\}$. Let $\mathcal{F}$ be the set of all (σ, π, θ) that induce a valid sidekick tour. We have the same conditions as in the previous problem that ensure σ and π do not pickup before launching or use more than k sidekicks. In addition, in this case we must require that each sidekick-visited customer has a launch and a pickup event and each truck-visited customer has a truck visit event, with each of these events being mapped to a unique place in the ordering of events. That is,

$$\sigma, \pi : \mathcal{S} \hookrightarrow \{1, \ldots, m\}$$
$$\theta : \mathcal{T} \hookrightarrow \{1, \ldots, m\}$$
$$\sigma(\mathcal{S}), \pi(\mathcal{S}), \theta(\mathcal{T}) \text{ are pairwise disjoint.}$$

Put another way, the maps σ, π, and θ have to jointly form a bijection between $\mathcal{S} \cup \mathcal{T}$ and $\{1, \ldots, m\}$ (to be precise, the three maps actually form a bijection between the multiset $\mathcal{S} \uplus \mathcal{S} \cup \mathcal{T}$ and $\{1, \ldots, m\}$).

The extension to (SK1) is then a natural one:

$$\begin{aligned}
\underset{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{t},\sigma,\pi,\theta}{\text{minimize}} \quad & t_{m+1} \quad s.t. && \text{(SK2)}\\
t_{j+1} &\geq t_j + \frac{1}{\phi_0}\|z_{j+1} - z_j\| \quad \forall j \in \{1, \ldots, m\} && (3)\\
t_{\pi(i)} &\geq t_{\sigma(i)} + \frac{1}{\phi_1}(\|x_i - p_i\| + \|p_i - y_i\|) \quad \forall i \in \mathcal{S} && (4)\\
z_{\sigma(i)} &= x_i \quad \forall i \in \mathcal{S}\\
z_{\pi(i)} &= y_i \quad \forall i \in \mathcal{S}\\
z_{\theta(i)} &= p_i \quad \forall i \in \mathcal{T}\\
t_0 &= 0\\
z_{m+1} &= z_0\\
(\sigma, \pi, \theta) &\in \mathcal{F},
\end{aligned}$$

where $\mathcal{S}$ is defined as the domain of σ and π and $\mathcal{T}$ is the domain of θ.

Figure 3 shows examples of solutions to the problems defined above for 30 customers with multiple sidekicks that are faster than the truck.

4 Preliminaries

Having defined two variants of sidekick routing, we now turn to some preliminary results that will be useful in our analysis of these problems. This section presents existing results from prior work as well as some additional analysis of our own.

4.1 Existing results from related work

The following classical theorem, originally stated in [3] and further developed in [68, 69], is one of the fundamental results of the continuous approximation paradigm; it relates the length of a TSP tour of a sequence of points to the distribution from which they were sampled:

Theorem 3 (BHH Theorem)**.** *Suppose that $X_1, X_2, \ldots$ is a sequence of random points i.i.d. according to an absolutely continuous probability density function f defined on a compact planar region $\mathcal{R}$. Then with probability one, the length $\mathsf{TSP}(X_1, \ldots, X_n)$ of the optimal travelling salesman tour through all X_i's satisfies*

$$\lim_{n\to\infty} \frac{\mathsf{TSP}(X_1, \ldots, X_n)}{\sqrt{n}} = \beta_{\mathsf{TSP}} \iint_{\mathcal{R}} \sqrt{f(x)}\, \mathrm{d}x$$

where β_{TSP} is a positive constant.

Although the exact value of β_{TSP} is unknown, it has been shown that $0.6250 \leq \beta_{\mathsf{TSP}} \leq 0.9204$; see [2, 3].

The concept of a *subadditive Euclidean functional* was introduced in [68], which provides a key insight that we will use in this report:

Definition 4. A function $L(\cdot)$ from the set of finite subsets of $\mathbb{R}^2$ to the non-negative real numbers is said to be a *monotone subadditive Euclidean functional on* $\mathbb{R}^2$ if it satisfies the following properties:

1. $L(\emptyset) = 0$.
2. *Homogeneity*: $L(\alpha x_1, \dots, \alpha x_n) = \alpha L(x_1, \dots, x_n)$ for all real $\alpha > 0$.
3. *Translation invariance*: $L(x_1 + x, \dots, x_n + x) = L(x_1, \dots, x_n)$ for all $x \in \mathbb{R}^2$.
4. *Monotonicity*: $L(x \cup A) \geq L(A)$ for any $x \in \mathbb{R}^2$ and finite subset $A \subset \mathbb{R}^2$.
5. *Geometric subadditivity*: There exists a constant $C > 0$, such that for all positive integers m, n and $\{x_1, \dots, x_n\} \in [0,1]^2$, we have

$$L(x_1, \dots, x_n) \leq \sum_{i=1}^{m^2} L(\{x_1, \dots, x_n\} \cap Q_i) + Cm$$

   where $\{Q_i\}$, $1 \leq i \leq m^2$ is the partition of $[0,1]^2$ into squares of edge length $1/m$.

Examples of subadditive Euclidean functionals include the TSP tour and the Steiner tree. The minimum spanning tree, the minimum matching, and the nearest neighbor graph are all “close” to being subadditive Euclidean functionals, but violate the monotonicity requirement (though it turns out that this can easily be overcome for all relevant applications). The monographs [69, 83] are devoted to more general settings for Theorem 3, with the most prominent generalization being the following:

Theorem 5 (basic theorem of subadditive Euclidean functionals)**.** *Suppose L is a monotone subadditive Euclidean functional defined on $\mathbb{R}^2$. If the random variables $\{X_i\}$ are independent with the uniform distribution on $[0,1]^2$, then with probability one, we have*

$$\frac{L(X_1, \dots, X_n)}{\sqrt{n}} \to \beta_L$$

as $n \to \infty$, where $\beta_L \geq 0$ is a constant.

We conclude with some additional problem definitions and convergence results that will also prove key to our analysis:

Definition 6 (Medians Problem)**.** Given a collection of points $x_1, \dots, x_n$ in $\mathbb{R}^2$ and a positive integer p, the the *p-medians problem* is given by

$$\text{PMed}(x_1, \dots, x_n; p) := \min_{\mathcal{S} \subset \{1,\dots,n\} : |\mathcal{S}| \leq p} \; \sum_{i=1}^{n} \min_{j \in \mathcal{S}} \|x_i - x_j\| \, ;$$

that is, the problem of selecting a subset $\mathcal{S} \subset \{1, \dots, n\}$ of *median* points such that $|\mathcal{S}| \leq p$, that minimizes the sum of the distances from all points to their nearest median.

Definition 7 (Balanced Medians Problem)**.** The *balanced medians problem* BMed$(x_1, \ldots, x_n; d)$ is a further-constrained variation of the p-medians problem. We can equivalently express p-medians as the problem of selecting a set of medians $\mathcal{S} \subset \{1, \ldots, n\}$ and an assignment of the points x_i to medians such that the sum of the distances from the points to their assigned medians is minimized. With no constraint on our assignment selection we have that in the p-medians problem the optimal assignment for any median set is simply to assign a point to its nearest median. The balanced medians problem imposes an additional constraint on the assignment selection, namely median $x_j \in \mathcal{S}$ can have at most $d \geq 2$ non-median points assigned to it. It is further required that each median is assigned to itself.

That is,

$$\text{BMed}(x_1, \ldots, x_n; d) := \min_{\substack{\mathcal{S} \subset \{1,\ldots,n\}:\, |\mathcal{S}|=p \\ \mu:\{1,\ldots,n\} \mapsto S}} \sum_{i=1}^{n} \|x_i - x_{\mu(i)}\|, \tag{5}$$

where

$$p = \left\lceil \frac{n}{d+1} \right\rceil,$$

$x_{\mu(i)}$ is the median assigned to point x_i, and for all j such that $j \in \mathcal{S}$, $x_{\mu(j)} = x_j$ and $x_{\mu(i)} = x_j$ for at most d of the $i \neq j$.

The following result is due to [83]:

Theorem 8 (Asymptotic convergence of the balanced medians problem)**.** *The balanced medians problem satisfies the same convergence as in Theorem 5; that is, if $X_1, X_2, \ldots$ is a sequence of random points i.i.d. according to an absolutely continuous probability density function f defined on a compact planar region $\mathcal{R}$ and $d \geq 2$ is fixed, then with probability one, the cost* BMed$(X_1, \ldots, X_n; d)$ *satisfies*

$$\lim_{n \to \infty} \frac{\text{BMed}(X_1, \ldots, X_n; d)}{\sqrt{n}} = \beta_{\text{BMed}}(d) \iint_{\mathcal{R}} \sqrt{f(x)}\, \mathrm{d}x$$

where $\beta_{\text{BMed}}(d)$ depends only on d.

4.2 Further notes on Theorem 8

This section describes a lower bound on the function $\beta_{\text{BMed}}(d)$ from Theorem 8.

Theorem 9. *The function $\beta_{\text{BMed}}(d)$ satisfies*

$$\beta_{\text{BMed}}(d) \geq \frac{\sqrt{2}d^{(3/2)}}{e\sqrt{\pi}(d+1)}.$$

That is, with probability one,

$$\lim_{n \to \infty} \frac{\text{BMed}(X_1, \ldots, X_n; d)}{\sqrt{n}} \geq \frac{\sqrt{2}d^{(3/2)}}{e\sqrt{\pi}(d+1)} \iint_{\mathcal{R}} \sqrt{f(x)}\, \mathrm{d}x.$$

Proof. See Section A of the appendix. ☐

4.3 Remark on the unconstrained p-medians problem

In addition to serving as a useful model element for many applications, the restrictive assumption of bounded degree of a median is necessary in order to arrive at a characterization of the limiting behavior of the medians problem for a general continuous demand distribution. However, for *uniformly* distributed demand, Hochbaum and Steele [31] provide a convergence result for the unbounded p-medians problem of Definition 6. They need only assume that the number of medians, p, grows linearly in n. That is, Hochbaum and Steele show the following.

Theorem 10 (Asymptotic convergence of the p-medians problem)**.** *If $X_1, X_2, \ldots$ is a sequence of points i.i.d. uniform on $[0,1]^2$ then for any $\alpha \in (0,1)$, with probability one the cost* $\mathsf{PMed}(X_1, \ldots, X_n; \lfloor \alpha n \rfloor)$ *satisfies*

$$\lim_{n\to\infty} \frac{\mathsf{PMed}(X_1, \ldots, X_n; \lfloor \alpha n \rfloor)}{\sqrt{n}} = \beta_{\mathsf{PMed}}(\alpha),$$

where $\beta_{\mathsf{PMed}}(\alpha)$ depends only on α.

This result does not provide information concerning the value of β_{PMed}, and the authors note that the question of determining exact values of such a constant is “usually hopeless.” Though it will not be necessary for our analysis of the sidekick problem, we note that by applying the same argument used in proving Theorem 9 we can obtain a lower bound for β_{PMed} as a side consequence.

Theorem 11. *The function $\beta_{\mathsf{PMed}}(\alpha)$ satisfies*

$$\beta_{\mathsf{PMed}}(\alpha) \geq \frac{\sqrt{2}(1-\alpha)^{3/2}}{e\sqrt{\pi\alpha}}.$$

That is, if $X_1, X_2, \ldots$ is a sequence of points i.i.d. uniformly on $[0,1]^2$ then for any $\alpha \in (0,1)$ then with probability one,

$$\lim_{n\to\infty} \frac{\mathsf{PMed}(X_1, \ldots, X_n; \lfloor \alpha n \rfloor)}{\sqrt{n}} \geq \frac{\sqrt{2}(1-\alpha)^{3/2}}{e\sqrt{\pi\alpha}}.$$

5 A continuous approximation analysis

This section describes a continuous approximation analysis of the sidekick routing problems (SK1) and (SK2).

5.1 Naive asymptotic analysis

Relying solely on Theorem 5, we can obtain the following partial characterization of the asymptotic behavior of both problems (SK1) and (SK2).

Claim 12. For fixed values of k, ϕ_0, and ϕ_1, let $T(p_1, \ldots, p_n)$ denote the optimal objective value of problem (SK1). Then if the customer points p_i consist of random samples P_i independently drawn from a uniform distribution on the unit square, then there exists a non-negative constant $c_{\mathsf{SK1}} = c_{\mathsf{SK1}}(k, \phi_0, \phi_1)$ such that

$$\frac{T(P_1, \ldots, P_n)}{\sqrt{n}} \to c_{\mathsf{SK1}}$$

with probability one as $n \to \infty$. The same statement holds when $T(\cdot)$ is the optimal objective value of problem (SK2), with a different constant $c_{\mathsf{SK2}} \leq c_{\mathsf{SK1}}$.

Proof. This follows immediately from Theorem 5 because it is entirely straightforward to verify that $T(\cdot)$ is a monotone subadditive Euclidean functional as defined in Definition 4. ☐

Claim 12 describes the scaling behavior of our problem as $n \to \infty$, namely that the cost scales proportionally to $\sqrt{n}$, but it tells us nothing about c_{SK1} (or c_{SK2}). For example, it is obvious that both are decreasing with respect to the three fixed parameters ϕ_0, ϕ_1, and k (since making things faster or increasing the number of helpers can only improve efficiency), and routine scaling arguments establish that $c_{\mathsf{SK1}}(k, \phi_0, \phi_1) = \phi_0 c_{\mathsf{SK1}}(k, 1, \phi_1/\phi_0)$ for all k, ϕ_0, ϕ_1 (and similarly for c_{SK2}). We devote the remainder of this section to a more precise analysis of c_{SK1} and c_{SK2}.

5.2 A lower bound for (SK2)

Of course, problem (SK2) is itself a lower bound of (SK1) by construction, so it will suffice to consider (SK2) only.

We derive a lower bound for (SK2) in terms of the Traveling Salesman tour and solution to a Bounded Medians Problem on the p_i.

Lemma 13. *Let T_n denote the optimal objective value for Problem (SK2). We have*

1. $\mathsf{TSP}(p_1, \ldots, p_n) \leq (\phi_0 + k\phi_1)T_n$
2. $\mathsf{BMed}(p_1, \ldots, p_n; d) \leq (d\phi_0 + k\phi_1)T_n$ for all $d \geq 2$.

Proof. For the first claim, we can construct a TSP solution from the (SK2) solution as follows. Consider an optimal solution $(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t}, \sigma, \pi, \theta, \mathcal{S}, \mathcal{T})$ to (SK2). For each $i \in \mathcal{S}$, the set of customers visited by the sidekicks, let

$$u_i := \operatorname*{argmin}_{u \in \{x_i, y_i\}} ||u - p_i||,$$

that is u_i is the closer to the customer of its sidekick launch and pickup points. For each $i \in \mathcal{T}$ let

$$u_i := p_i.$$

We then construct a TSP tour of the points as follows. Let the tour follow the path of the truck, visiting the customers in $\mathcal{T}$ along the tour. Whenever we reach one of the u_i for $i \in \mathcal{S}$, let the tour travel from u_i to p_i and back, then continue along the truck path. It is clear that the length added to our TSP tour coming from the truck's route is less than or equal to $\phi_0 T_n$, the truck's speed times to total time for our sidekick tour.

To bound the length from visiting points in $\mathcal{S}$ we let P_j be the set of points visited by sidekick j. Then

$$T_n \geq \frac{1}{\phi_1} \sum_{i: p_i \in P_j} ||x_i - p_i|| + ||p_i - y_i|| \quad \forall j \in \{1, \ldots, k\}.$$

That is the total sidekick tour time exceeds the time any given sidekick travels. Dividing by the above by k and summing over all j yields

$$\begin{aligned} T_n &\geq \frac{1}{k\phi_1}\sum_{i\in\mathcal{S}}||x_i - p_i|| + ||p_i - y_i|| \\ &\geq \frac{2}{k\phi_1}\sum_{i\in\mathcal{S}}||u_i - p_i||. \end{aligned}$$

Twice the sum of the $||u_i - p_i||$ is precisely what we add to our TSP tour to visit $\mathcal{S}$. Thus the contribution of this part of our TSP tour is bounded by $k\phi_1 T_n$. Adding together our truck and sidekick pieces of the TSP tour and applying the triangle inequality gives the result.

For the second bound we can construct a balanced median solution from the (SK2) solution as follows. Think of the truck as completing a tour on our u_i defined as above. Group every $d+1$ of the customer points associated with the u_i along this tour and choose as their median the point which is closest to the tour. This construction is pictured in Figure 5. By the triangle inequality, the distance from a point to its assigned median is less than or equal to the distance of traveling from that point to its corresponding u_i, then traveling along the truck tour to the median's corresponding u_i, then traveling out to the assigned median. The cost of this balanced medians solution, i.e. sum of these distances, is then less than or equal to the sum of the distances from the non-median points to their corresponding u_i, plus d times the length of the tour of the u_i, plus the sum, over all medians, of d times the distance from the median's corresponding u to the median. By our selection of the medians it is clear that this last sum is less than or equal to the sum of all of the distances from non-median points to their corresponding u. Then, noting once again that

$$\phi_0 T_n \geq \text{truck tour of the } u_i,$$

and

$$\begin{aligned} \frac{k\phi_1}{2}\cdot T_n &\geq \sum_{i\in\mathcal{S}}||u_i - p_i|| \\ &= \sum_{i=1}^{n}||u_i - p_i||, \qquad (u_i = p_i \text{ for } i \in \mathcal{T}) \end{aligned}$$

the length of the bounded medians solution is less than or equal to

$$\left(\frac{k\phi_1}{2} + d\phi_0 + \frac{k\phi_1}{2}\right)T_n.$$

□

5.3 An upper bound for (SK1)

To bound the objective value of (SK1), we describe a simple “zig-zagging” heuristic in the unit square:

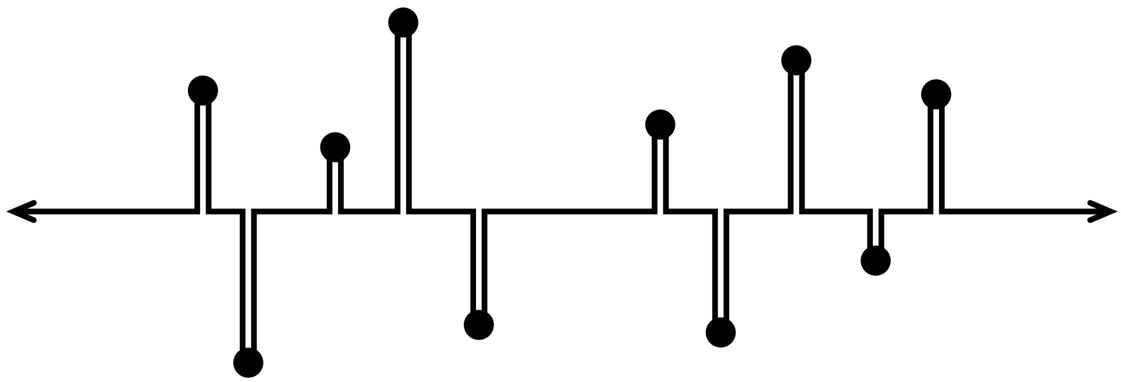


Figure 4: Constructing a TSP solution from a solution to problem (SK2). The horizontal line represents the tour of the u_i in the (SK2) solution. We follow the tour, traveling from u_i to p_i and back for each $i \in \mathcal{S}$. If T_n is the cost of the problem (SK2) solution then the total cost of the resulting TSP solution is less than or equal to $(\phi_0 + k\phi_1)T_n$.

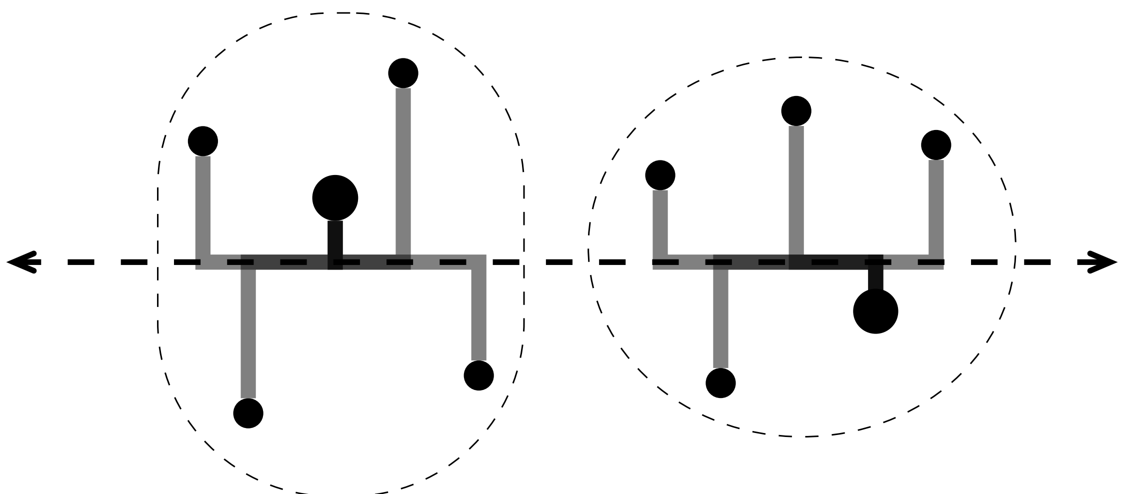
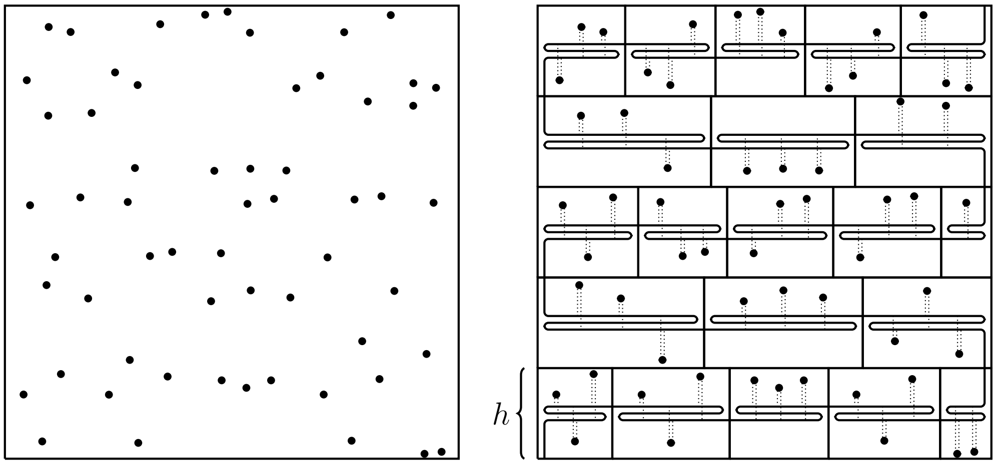


Figure 5: Constructing a balanced medians solution from a solution to problem (SK2). The horizontal line represents the tour of the u_i in the (SK2) solution. Here we choose $d = 4$ and group every 5 points along the tour. We choose as the median for these 5 points the point which is closest to the tour. Using the paths pictured, it is clear that to connect all points to their medians we need travel at most d times the length of the truck tour plus twice the total distance from the points to their u_i. If T_n is the cost of the problem (SK2) solution then the total cost of the resulting Bounded Medians solution is less than or equal to $(d\phi_0 + k\phi_1)\, T_n$.



(a) Input (b) Output

Figure 6: The tour described in Lemma 14, assuming $k = 3$.

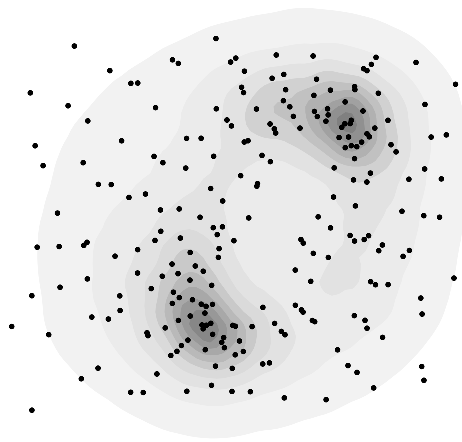
Lemma 14. *For fixed ϕ_0, ϕ_1, and k and points $p_1, \ldots, p_n$ lying in the unit square, there exists a routing strategy for problem (SK1) whose time to completion $T(p_1, \ldots, p_n)$ satisfies*

$$T(p_1, \ldots, p_n) \leq \frac{2\sqrt{3}}{\sqrt{\phi_0 \phi_1 k}} \cdot \sqrt{n} + C$$

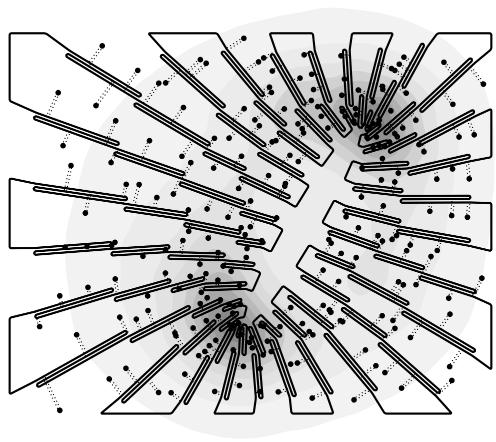
where C is a constant that depends only on ϕ_0, ϕ_1, and k.

Proof. Assume without loss of generality that $\phi_0 = 1$ and divide the unit square into strips of height $h = \sqrt{3\phi_1 k/n}$ (there may be one strip whose height is less than this due to rounding). There are $m = \left\lceil \sqrt{n/(3\phi_1 k)} \right\rceil \leq \sqrt{n/(3\phi_1 k)} + 1$ such strips. Further subdivide each strip into rectangles so that each rectangle (except possibly the rightmost in each strip) contains k points. There are at most $m + n/k$ rectangles in total. Finally, construct a tour for the truck and all helpers by traversing each rectangle three times, releasing the helpers on the first traversal and retrieving the helpers on the third traversal, as illustrated in Figure 6.

It is easy to see that for a rectangle having width w (and height h), it is possible to perform three horizontal traversals and release and retrieve the helpers in at most $3w + h/\phi_1$ time units. It is also easy to see that the only remaining time needed is for the truck to perform vertical moves to move from one strip to the next, which is a constant amount of 2 time units, plus whatever time is needed for the truck to return to its point of origin, which is also at most $\sqrt{2}$ time units. Hence, if we let w_i denote the width of rectangle i, then the total amount of time to complete this tour



(a) yo



(b) yo

Figure 7: The tour described in Lemma 15; it consists of the same kind of tour as in Figure 6, but “scaled” with respect to the probability distribution that is indicated by shading.

is at most

$$
\begin{aligned}
(2+\sqrt{2}) + \sum_i (3w_i + h/\phi_1) &\leq (2+\sqrt{2}) + 3\underbrace{\sum_i w_i}_{=m} + (m+n/k)h/\phi_1 \\
&\leq (2+\sqrt{2}) + 3\left(\sqrt{\frac{n}{3\phi_1 k}}+1\right) + \frac{1}{\phi_1}\left(\sqrt{\frac{n}{3\phi_1 k}}+1+n/k\right)\sqrt{\frac{3\phi_1 k}{n}} \\
&= \frac{2\sqrt{3}}{\sqrt{\phi_1 k}}\cdot\sqrt{n} + \sqrt{\frac{3k}{\phi_1 n}} + \frac{1}{\phi_1} + (5+\sqrt{2})
\end{aligned}
$$

as desired. □

Lemma 14 is deterministic, but also implies the following:

Lemma 15. *Let ϕ_0, ϕ_1, and k be fixed and let $P_1, \dots, P_n$ be independent samples from an absolutely continuous probability density f with compact support $\mathcal{R}$. The optimal time to completion $T(P_1, \dots, P_n)$ for problem (SK1) satisfies*

$$
\limsup_{n\to\infty} \frac{T(P_1,\dots,P_n)}{\sqrt{n}} \leq \frac{3.47}{\sqrt{\phi_0\phi_1 k}} \cdot \iint_{\mathcal{R}} \sqrt{f(x)}\,\mathrm{d}x
$$

with probability one.

Proof. This is a routine scaling argument, together with the law of large numbers and the fact $T(\cdot)$ is a subadditive Euclidean functional (see Claim 12); see Section B of the appendix for details. □

5.4 Convergence analysis for (SK1) and (SK2)

We have now collected enough supporting evidence for our main claim:

Theorem 16. *Let ϕ_0, ϕ_1, and k be fixed. Let T_n denote the optimal objective value to problem (SK1), where input points $P_1, \dots, P_n$ are independent uniform samples in the unit square. Then there exists a constant β_{SK1} satisfying $0.1389 < \beta_{\mathsf{SK1}} < 3.47$ such that*

$$\frac{T_n}{\sqrt{n}} \to \frac{\beta_{\mathsf{SK1}}}{\sqrt{\phi_0 \max\{\phi_0, \phi_1 k\}}} \tag{6}$$

with probability one as $n \to \infty$. Moreover, the same statement holds for a different constant $\beta_{\mathsf{SK2}} \le \beta_{\mathsf{SK1}}$ when T_n is the optimal objective value to problem (SK2), which also satisfies $0.1389 < \beta_{\mathsf{SK2}} < 3.47$. Finally, when the points $P_1, \dots, P_n$ are independent samples from an absolutely continuous probability density f with compact support $\mathcal{R}$, we have

$$0.1389c \le \liminf \frac{T_n}{\sqrt{n}} \le \limsup \frac{T_n}{\sqrt{n}} \le 3.47c$$

with probability one as $n \to \infty$, where T_n is the optimal objective value to either problem (SK1) or (SK2), and

$$c = \frac{\iint_{\mathcal{R}} \sqrt{f(x)}\,\mathrm{d}x}{\sqrt{\phi_0 \max\{\phi_0, \phi_1 k\}}}\,.$$

Proof. To simplify notation, assume without loss of generality that $\phi_0 = 1$ and set $t = \phi_1 k$, and rewrite the desired result (6) equivalently as

$$\sqrt{\max\{1, t\}} \cdot \frac{T_n}{\sqrt{n}} \to \beta_{\mathsf{SK1}}\,.$$

The existence of β_{SK1} and β_{SK2} was already established in Claim 12 (set $\beta_{\mathsf{SK1}} = c_{\mathsf{SK1}}\sqrt{\phi_0 \max\{\phi_0, \phi_1 k\}}$ and so forth); the real work lies in computing the bounds on these constants. Since $\beta_{\mathsf{SK2}} \le \beta_{\mathsf{SK1}}$, it will suffice to show that $0.1389 < \beta_{\mathsf{SK2}}$ and that $\beta_{\mathsf{SK1}} < 3.47$. To show that $0.1389 < \beta_{\mathsf{SK2}}$, Lemma 13 says that

$$\begin{aligned}
T_n &\ge \frac{\mathsf{TSP}(P_1, \dots, P_n)}{1+t} && (7)\\
\implies \lim_{n\to\infty} \frac{T_n}{\sqrt{n}} &\ge \lim_{n\to\infty} \frac{\mathsf{TSP}(P_1, \dots, P_n)}{(1+t)\sqrt{n}} = \frac{\beta_{\mathsf{TSP}}}{1+t} && (8)\\
\implies \sqrt{\max\{1,t\}} \cdot \lim_{n\to\infty} \frac{T_n}{\sqrt{n}} &\ge \sqrt{\max\{1,t\}} \cdot \frac{\beta_{\mathsf{TSP}}}{1+t} \ge \frac{0.625\sqrt{\max\{1,t\}}}{1+t} \\
\implies \beta_{\mathsf{SK2}} &\ge \frac{0.625\sqrt{\max\{1,t\}}}{1+t} > 0.1389 \text{ whenever } t < 18.166, && (9)
\end{aligned}$$

where in (8) we are justified in taking limits as we have seen such limits exists for problem (SK2) (Claim 12) and for the TSP (Theorem 3).

In addition Lemma 13 tells us that, provided $t \geq 2$,

$$T_n \geq \frac{\text{BMed}(P_1, \dots P_n; \lfloor t \rfloor)}{\lfloor t \rfloor + t} \tag{10}$$

$$\implies \lim_{n\to\infty} \frac{T_n}{\sqrt{n}} \geq \lim_{n\to\infty} \frac{\text{BMed}(P_1, \dots P_n; \lfloor t \rfloor)}{(\lfloor t \rfloor + t)\sqrt{n}} = \frac{\beta_{\text{BMed}}(\lfloor t \rfloor)}{\lfloor t \rfloor + t} \tag{11}$$

$$\implies \sqrt{\max\{1,t\}} \cdot \lim_{n\to\infty} \frac{T_n}{\sqrt{n}} \geq \frac{\beta_{\text{BMed}}(\lfloor t \rfloor)\sqrt{\max\{1,t\}}}{\lfloor t \rfloor + t} \geq \frac{\sqrt{2}\lfloor t \rfloor^{(3/2)}\sqrt{\max\{1,t\}}}{e\sqrt{\pi}(\lfloor t \rfloor + 1)(\lfloor t \rfloor + t)} \tag{12}$$

$$\implies \beta_{\text{SK2}} \geq \frac{\sqrt{2}\lfloor t \rfloor^{(3/2)}\sqrt{\max\{1,t\}}}{e\sqrt{\pi}(\lfloor t \rfloor + 1)(\lfloor t \rfloor + t)} > 0.1389 \text{ whenever } t \geq 18.166, \tag{13}$$

where in (11) we are justified in taking limits as we have seen such limits exists for problem (SK2) (Claim 12) and for the balanced medians problem (Theorem 8), and in (12) we have applied Theorem 9.

The upper bound $\beta_{\text{SK1}} < 3.47$ is very simple. From Lemma 14, we have

$$T_n \leq \frac{2\sqrt{3}}{\sqrt{t}} \cdot \sqrt{n} + C \tag{14}$$

$$\implies \sqrt{\max\{1,t\}} \cdot \lim_{n\to\infty} \frac{T_n}{\sqrt{n}} \leq \sqrt{\max\{1,t\}} \cdot \lim_{n\to\infty} \left(\frac{2\sqrt{3}}{\sqrt{t}} + \frac{C}{\sqrt{n}} \right)$$

$$\implies \beta_{\text{SK1}} \leq 2\sqrt{3} < 3.47 \text{ for } t \geq 1,$$

and for $t < 1$, we simply eschew the helpers altogether and visit all of the P_i's with the truck (to be precise, since the truck is not allowed to visit any points in (SK1), we bring the truck within arbitrarily small distance ϵ from each P_i and release and retrieve one of the helpers):

$$\lim_{n\to\infty} \frac{T_n}{\sqrt{n}} \leq \beta_{\text{TSP}}$$

$$\implies \sqrt{\max\{1,t\}} \cdot \lim_{n\to\infty} \frac{T_n}{\sqrt{n}} \leq \sqrt{\max\{1,t\}} \cdot \beta_{\text{TSP}} = \beta_{\text{TSP}}$$

$$\implies \beta_{\text{SK1}} \leq \beta_{\text{TSP}} \leq 0.9204 \text{ for } t < 1$$

as desired. This completes the proof of the uniform case of Theorem 16.

The non-uniform case of Theorem 16 follows the exact same logic; the only distinction is that we are no longer guaranteed that $T_n/\sqrt{n}$ has a limit, so we merely replace all instances of “$\lim_{n\to\infty} T_n/\sqrt{n}$” with either a “$\liminf_{n\to\infty}$” or a “$\limsup_{n\to\infty}$” depending on whether we are bounding from above or below. For example, the lower bound (7) becomes

$$T_n \geq \frac{\text{TSP}(P_1, \dots, P_n)}{1+t}$$

$$\implies \liminf_{n\to\infty} \frac{T_n}{\sqrt{n}} \geq \lim_{n\to\infty} \frac{\text{TSP}(P_1, \dots, P_n)}{(1+t)\sqrt{n}} = \frac{\beta_{\text{TSP}} \iint_{\mathcal{R}} \sqrt{f(x)}\,\mathrm{d}x}{1+t} \geq \frac{0.625 \iint_{\mathcal{R}} \sqrt{f(x)}\,\mathrm{d}x}{1+t}$$

$$\implies \sqrt{\max\{1,t\}} \cdot \liminf_{n\to\infty} \frac{T_n}{\sqrt{n}} \geq \frac{0.625\sqrt{\max\{1,t\}}}{1+t} \iint_{\mathcal{R}} \sqrt{f(x)}\,\mathrm{d}x > 0.1389 \iint_{\mathcal{R}} \sqrt{f(x)}\,\mathrm{d}x \text{ whenever } t < 18.1$$

$$\implies \liminf_{n\to\infty} \frac{T_n}{\sqrt{n}} \geq 0.1389c \text{ whenever } t < 18.166.$$

The same reasoning is applied for the bounded-medians-derived lower bound for $t \geq 18.166$.

The upper bound that $\limsup T_n/\sqrt{n} \leq 3.47c$ is also immediate; we already proved this for $t \geq 1$ in Lemma 15, and when $t < 1$, we again eschew the helpers altogether and use the truck:

$$\limsup_{n\to\infty} \frac{T_n}{\sqrt{n}} \leq \beta_{\mathsf{TSP}} \iint_{\mathcal{R}} \sqrt{f(x)}\,\mathrm{d}x \leq 0.9204 \iint_{\mathcal{R}} \sqrt{f(x)}\,\mathrm{d}x < 3.47c\,,$$

which completes the proof. □

6 Remarks

Informally, Theorem 16 says that the time to completion of a sidekick routing problem satisfies

$$\text{Time With Sidekicks} \propto \frac{\sqrt{n}}{\sqrt{\phi_0 \max\{\phi_0, \phi_1 k\}}}.$$

We see immediately that for $(k\phi_1)/\phi_0 < 1$, both the lower and upper bounds on the Sidekick problems that produce our result are constant multiples of the optimal TSP objective. We conclude that the asymptotic behavior of the sidekick problem looks more or less like that of the TSP in this case. This yields the managerial insight that there is no real benefit to introducing sidekicks if we are not guaranteed $(k\phi_1)/\phi_0 \geq 1$. If we will have sufficiently fast or sufficiently numerous sidekicks to guarantee this then Theorem 16 tells us as that as we add more and more customer points we can essentially say that

$$\text{Time With Sidekicks} \propto \frac{\sqrt{n}}{\sqrt{\phi_0\phi_1 k}}.$$

On the other hand in the limit the tour with just the truck has time

$$\text{Time Without Sidekicks} = \frac{\text{TSP}}{\phi_0} \propto \frac{\sqrt{n}}{\phi_0}.$$

So there is certainly a boost in efficiency to be had by introducing sidekicks. The amount of improvement due to using sidekicks is captured by

$$\frac{\text{Time With Sidekicks}}{\text{Time Without Sidekicks}} \propto \sqrt{\frac{\phi_0}{\phi_1 k}}.$$

We note that all of the above remarks hold in both the uniform and non-uniform cases because, as we have also seen in Theorem 16, the difference between these two cases merely amounts to multiplication by a factor of $\iint_{\mathcal{R}} \sqrt{f(x)}\,\mathrm{d}x$.

7 Computational results

We run simulations with points drawn from a uniform distribution in the unit square to see how the sidekick problem tours compare to our asymptotic expectations. In our computations the number of points, number of sidekicks and the ratio of the truck and sidekick speeds vary. We assume that the truck speed is always 1, letting the sidekick speed capture the ratio of the speeds. We solve Problem (SK1), in which the truck does not make deliveries. Each result is an average over 5 draws of customer points.

In order to approximately solve Problem (SK1) we use a heuristic to obtain σ and π specifying the ordering in which we launch and pickup the sidekicks for each customer, and then solve the resulting problem in the variables x, y, z, and t. To obtain the ordering, we first compute the optimal TSP tour of the customers and break this tour into consecutive chunks of size k, the number of sidekicks. To determine the launch and pickup ordering on each chunk we begin by assuming as in [56] and [13] that in the case of fast sidekicks the improvement due to multiple sidekicks is captured well by a sequence in which all sidekicks are launched and then all sidekicks are picked up. Within each chunk the launches for all the customer points occur in the order specified by the TSP tour, then the pickups occur in the order specified by the TSP tour. However, we recognize that in the case of slow sidekicks, this strategy breaks down and we expect an optimal tour to often use fewer than the full number of sidekicks at a given time, allowing the truck to cover large distances with no sidekicks out. For this reason we also consider the ordering in which we follow the TSP tour, launching for a customer and then picking up for that customer one at a time. We take the minimum result of these two approaches to be our service time. It will be clear from our plots that neither of these strategies sufficiently captures the benefits of having more and more sidekicks when the sidekicks are slower than the truck. In fact the one launch one pickup option reduces us to the one sidekick case. However, it is unreasonable to think that such a simplistic heuristic solution to this difficult problem would come close to optimality in all cases.

Letting T denote the optimal service time for the Problem (SK1), Theorem 16 tells us that we should have

$$T \approx C \frac{\sqrt{n}}{\sqrt{\phi_0 \max(\phi_0, \phi_1 k)}},$$

for some constant C.

So we plot

$$F := T \cdot \frac{\sqrt{\phi_0 \max(\phi_0, \phi_1 k)}}{\sqrt{n}},$$

for ranges of parameter values, expecting this value to be constant.

- In Figure 8 we fix the number of sidekicks and plot F over a range of values of ϕ_1/ϕ_0 and n.
- In Figure 9 we fix the ratio of the speeds and plot F for a range of values of k and n.
- In Figure 10 we once again fix the ratio of the speeds and plot F for a range of values of k and n, now for higher values of the ratio.
- In Figure 11 we fix n and plot F over a range of values of k and ϕ_1/ϕ_0.

In addition, since the TSP can be approximated by a constant times $\sqrt{n}$ for large n we would expect that, letting TSP denote the time for the truck to complete a TSP of the same points we should have

$$T \approx C' \frac{\text{TSP}}{\sqrt{\phi_0 \max(\phi_0, \phi_1 k)}},$$

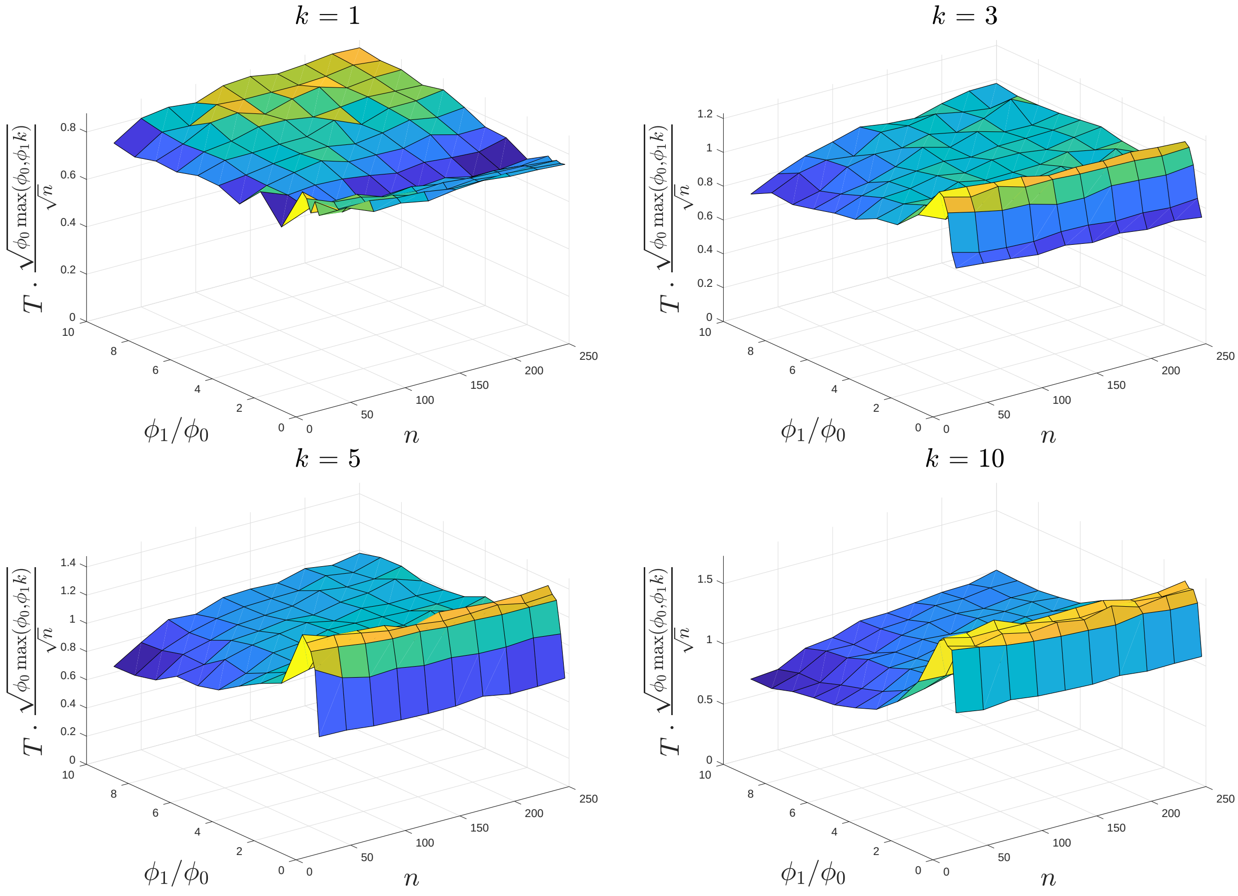


Figure 8: Value of the sidekick problem completion time times $\sqrt{\phi_0 \max(\phi_0, \phi_1 k)}/\sqrt{n}$ plotted over a range of values of ϕ_1/ϕ_0 and n for fixed k. We expect this value to be constant for large n. We can see that our heuristic approach does not sufficiently capture the benefit of introducing more and more sidekicks in the slow sidekick case, but otherwise these plots are near constant, particularly as n becomes large.

for some constant C'.

So we plot

$$F' := T \cdot \frac{\sqrt{\phi_0 \max(\phi_0, \phi_1 k)}}{\mathrm{TSP}},$$

for ranges of parameter values, expecting this value to be constant.

- In Figure 12 we fix n and plot F' over a range of values of k and ϕ_1/ϕ_0.

8 Conclusions

We have studied the limiting behavior of sidekick-assisted routing problems in the Euclidean plane and found that the improvements introduced by adding sidekicks depend on $\sqrt{\phi_0}$ and $\sqrt{\phi_1 k}$. There remain many open questions: for example, what happens when sidekicks are able to visit

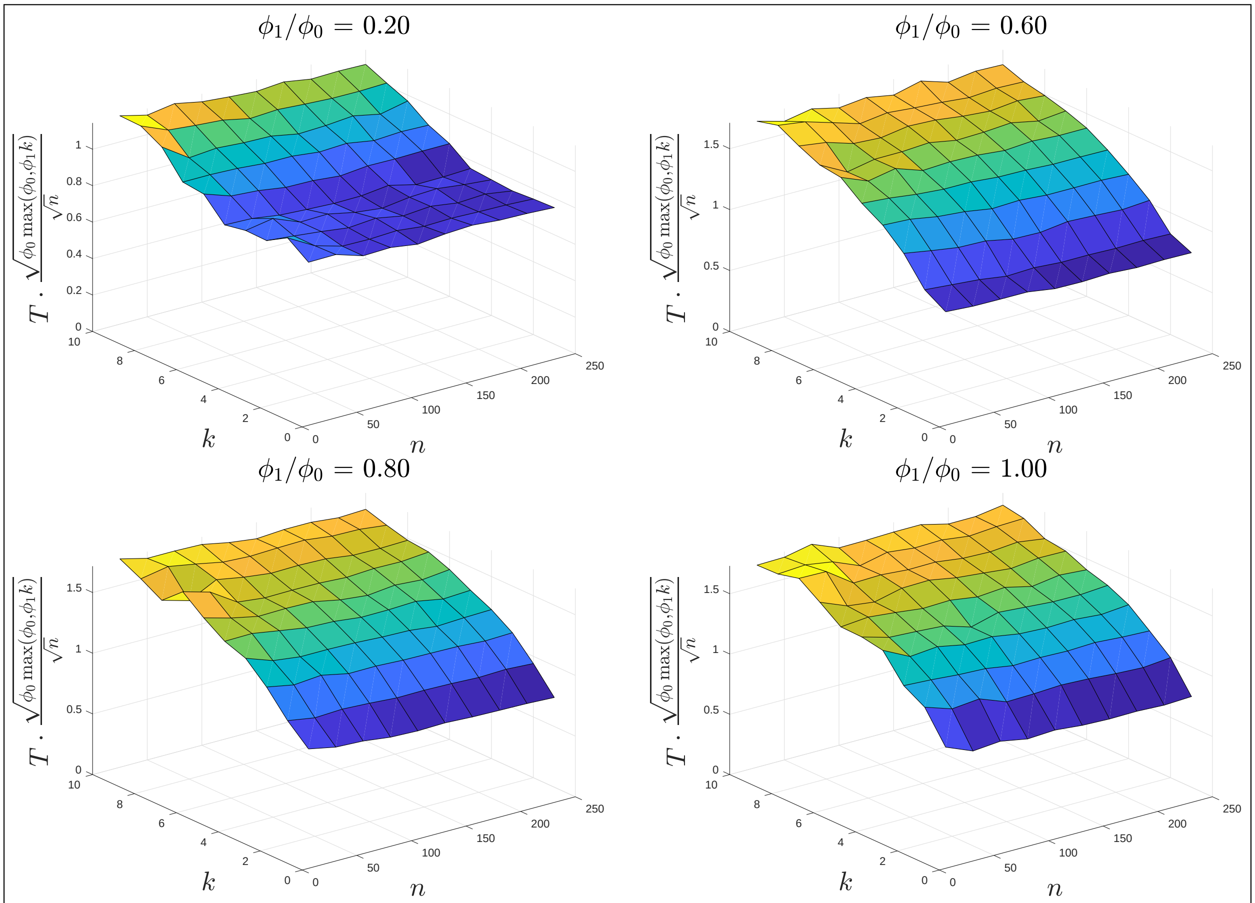


Figure 9: Value of the sidekick problem completion time times $\sqrt{\phi_0 \max(\phi_0, \phi_1 k)}/\sqrt{n}$ plotted over a range of values of k and n for fixed ϕ_1/ϕ_0. We expect this value to be constant for large n. As these values of ϕ_1/ϕ_0 are small our heuristic approach is not capturing the benefit due to introducing more and more sidekicks, and thus these plots are not constant though they fall within a range of about 1.

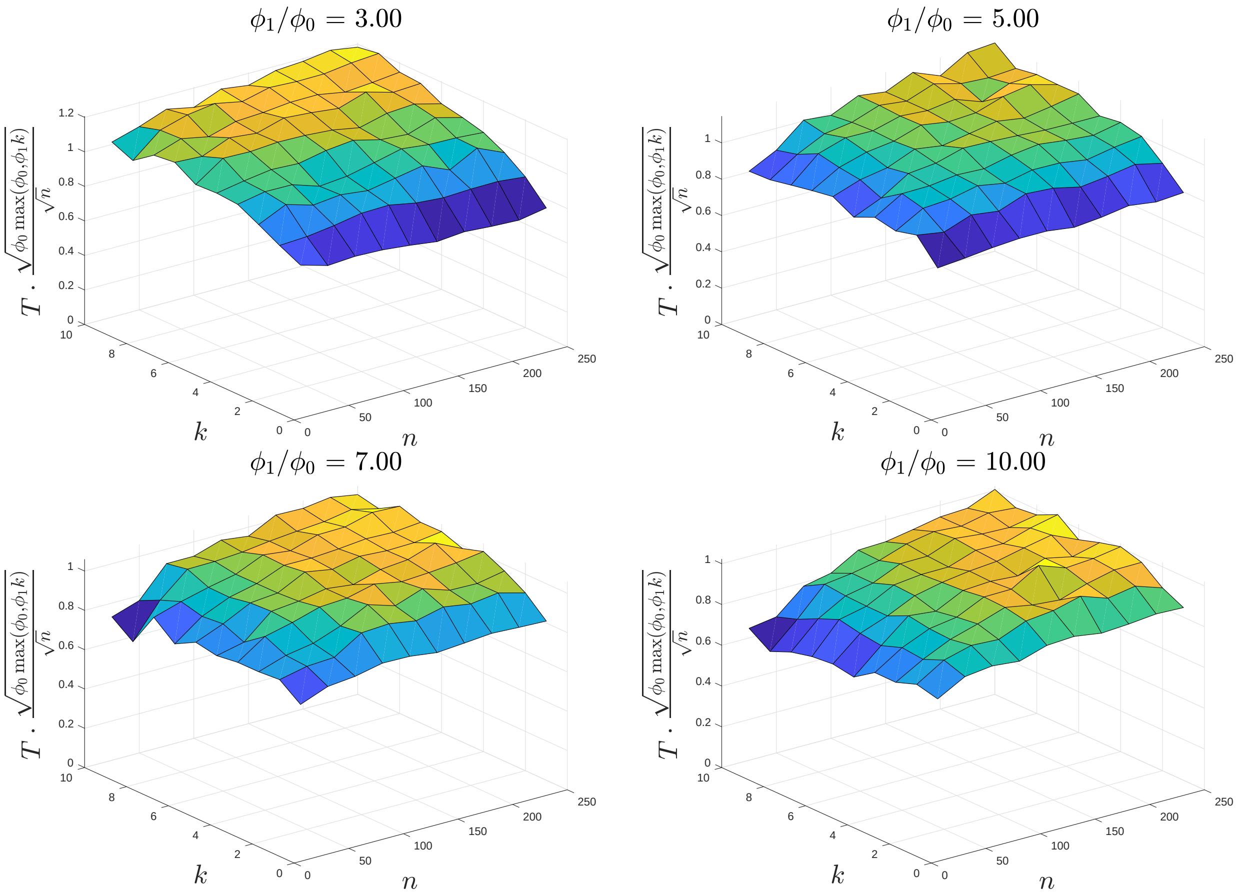


Figure 10: Value of the sidekick problem completion time times $\sqrt{\phi_0 \max(\phi_0, \phi_1 k)}/\sqrt{n}$ plotted over a range of values of k and n for fixed ϕ_1/ϕ_0. We expect this value to be constant for large n. We see the result is near constant as n becomes large.

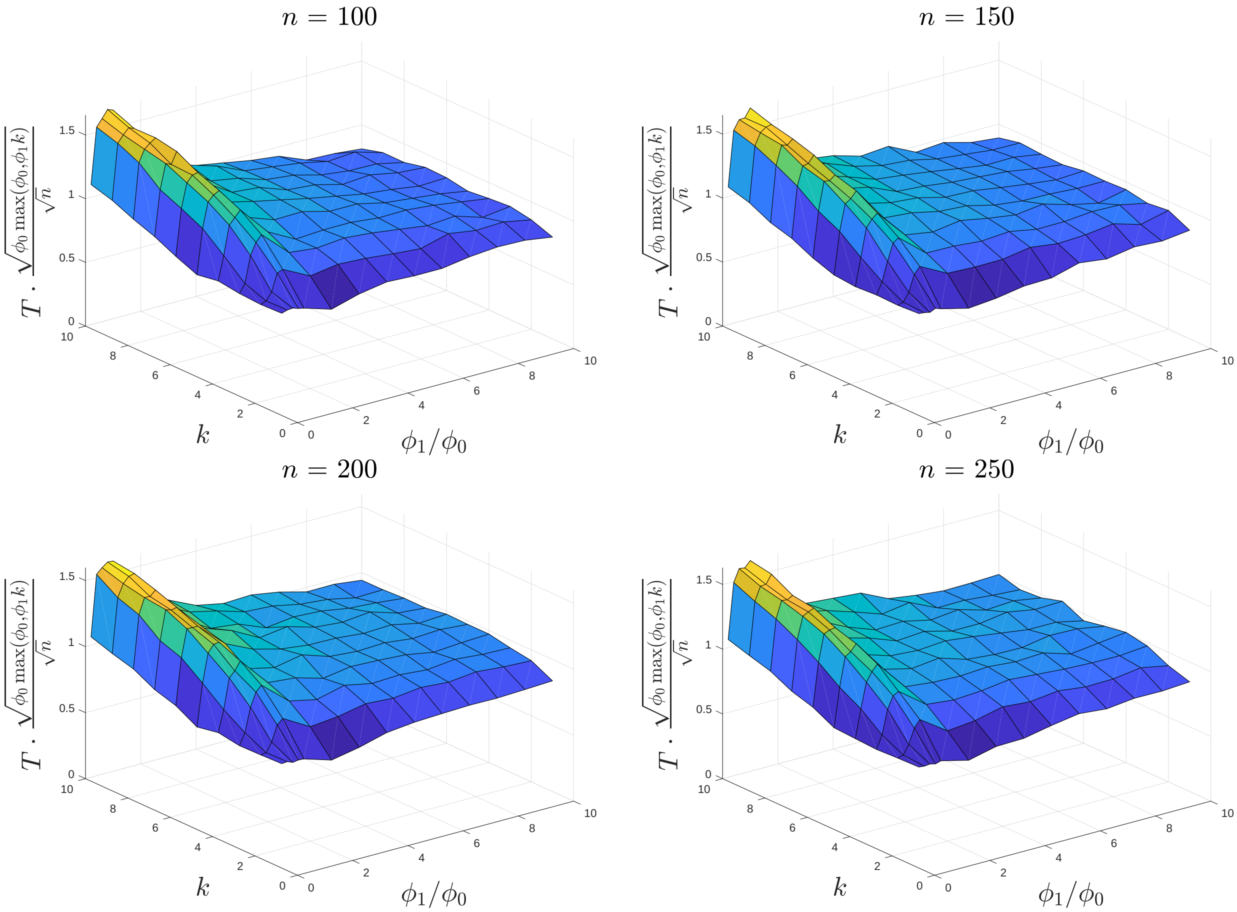


Figure 11: Value of the sidekick problem completion time, T, times $\sqrt{\phi_0 \max(\phi_0, \phi_1 k)}/\sqrt{n}$ plotted over a range of values of k and ϕ_1/ϕ_0 for fixed n. We expect this value to be constant for large n. When ϕ_1/ϕ_0 is small our heuristic approach does not sufficiently capture the benefit of introducing more and more sidekicks, but otherwise these plots are near constant.

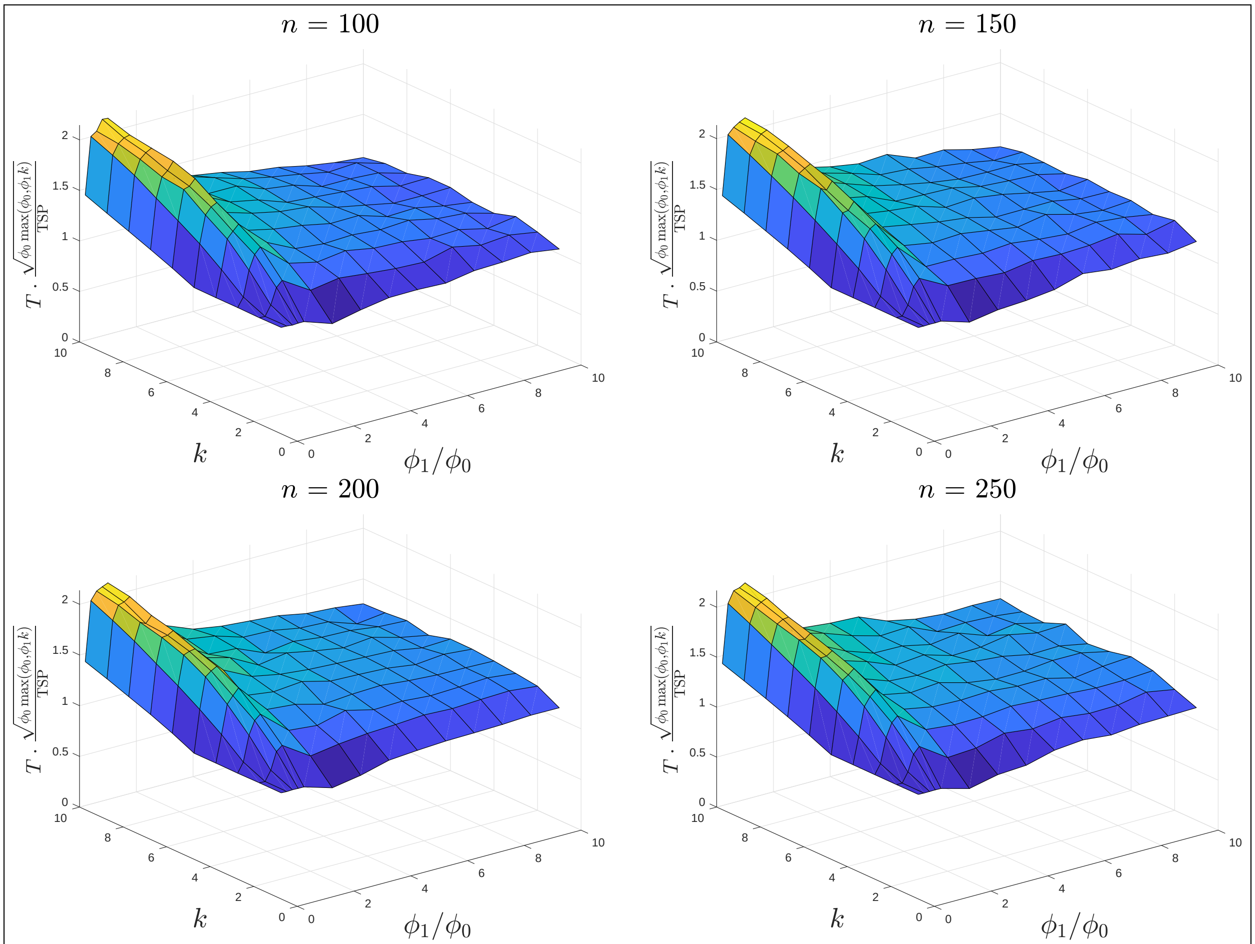


Figure 12: Value of the sidekick problem completion time times $\sqrt{\phi_0 \max(\phi_0, \phi_1 k)}$ divided by the optimal time for a TSP with only the truck, plotted over a range of values of k and ϕ_1/ϕ_0 for fixed n. We expect this value to be constant for large n. When ϕ_1/ϕ_0 is small our heuristic approach does not sufficiently capture the benefit of introducing more and more sidekicks, but otherwise these plots are near constant.

more than one customer node before returning to the truck? What happens when the truck is itself capacitated and must make returns to the depot? What happens when sidekick battery life considerations come into play? We hope to resolve these questions in future work.

References

[1] Niels Agatz, Paul Bouman, and Marie Schmidt. Optimization approaches for the traveling salesman problem with drone. *Transportation Science*, 52(4):965–981, 2018.

[2] D. Applegate, W. Cook, D. S. Johnson, and N. J. A. Sloane. Using large-scale computation to estimate the Beardwood-Halton-Hammersley TSP constant. Presentation at 42 SBPO, 2010.

[3] J. Beardwood, J. H. Halton, and J. M. Hammersley. The shortest path through many points. *Mathematical Proceedings of the Cambridge Philosophical Society*, 55(4):299–327, 1959.

[4] Mohd Shahrizan bin Othman, Aleksandar Shurbevski, Yoshiyuki Karuno, and Hiroshi Nagamochi. Routing of carrier-vehicle systems with dedicated last-stretch delivery vehicle and fixed carrier route. *Journal of Information Processing*, 25:655–666, 2017.

[5] Paul Bouman, Niels Agatz, and Marie Schmidt. Dynamic programming approaches for the traveling salesman problem with drone. *Networks*, 72(4):528–542, 2018.

[6] Nils Boysen, Dirk Briskorn, Stefan Fedtke, and Stefan Schwerdfeger. Drone delivery from trucks: Drone scheduling for given truck routes. *Networks*, 72(4):506–527, 2018.

[7] Nils Boysen, Stefan Schwerdfeger, and Felix Weidinger. Scheduling last-mile deliveries with truck-based autonomous robots. *European Journal of Operational Research*, 271(3):1085 – 1099, 2018.

[8] L. D. Burns, R. W. Hall, D. E. Blumenfeld, and C. F. Daganzo. Distribution strategies that minimize transportation and inventory costs. *Operations Research*, 33(3):469–490, 1985.

[9] Jacqueline Bühlmann. Swiss post delivery robots in use by jelmoli. `https://www.post.ch/en/about-us/company/media/press-releases/2017/swiss-post-delivery-robots-in-use-by-jelmoli`, 2017.

[10] G. P. Cachon. Retail store density and the cost of greenhouse gas emissions. *Management Science*, 60(8):1907–1925, 2014.

[11] J. F. Campbell, A. T. Ernst, and M. Krishnamoorthy. Hub arc location problems: part I – introduction and results. *Management Science*, 51(10):1540–1555, 2005.

[12] J. F. Campbell and M. E. O'Kelly. Twenty-five years of hub location research. *Transportation Science*, 46(2):153–169, 2012.

[13] James Campbell, Donald Sweeney, and Juan Zhang. Strategic design for delivery with trucks and drones, 04 2017. Supply Chain & Analytics Report SCMA-2017-0201.

[14] J. G. Carlsson. Dividing a territory among several vehicles. *INFORMS Journal on Computing*, 24(4):565–577, 2012.

[15] John Gunnar Carlsson and Siyuan Song. Coordinated logistics with a truck and a drone. *Management Science*, 64(9):4052–4069, 2017.

[16] Bruno N Coelho, Vitor N Coelho, Igor M Coelho, Luiz S Ochi, Demetrius Zuidema, Milton SF Lima, Adilson R da Costa, et al. A multi-objective green uav routing problem. *Computers & Operations Research*, 88:306–315, 2017.

[17] Rami Daknama and Elisabeth Kraus. Vehicle routing with drones. *arXiv preprint arXiv:1705.06431*, 2017.

[18] Mauro Dell’Amico, Roberto Montemanni, and Stefano Novellani. Drone-assisted deliveries: new formulations for the flying sidekick traveling salesman problem. *arXiv preprint arXiv:1905.13463*, 2019.

[19] Okan Dukkanci, Bahar Y Kara, and Tolga Bektas. The drone delivery problem. *Tolga, The Drone Delivery Problem (January 10, 2019)*, 2019.

[20] Sergio Mourelo Ferrandez, Timothy Harbison, Troy Weber, Robert Sturges, and Robert Rich. Optimization of a truck-drone in tandem delivery network using k-means and genetic algorithm. *Journal of Industrial Engineering and Management*, 9(2):374–388, 2016.

[21] L. Few. The shortest path and the shortest road through n points. *Mathematika*, 2:141–144, 1955.

[22] Júlia Cária de Freitas and Puca Huachi Vaz Penna. A variable neighborhood search for flying sidekick traveling salesman problem. *International Transactions in Operational Research*, 0(0), 2019.

[23] Emanuele Garone, Roberto Naldi, A Casavola, and Emilio Frazzoli. Cooperative mission planning for a class of carrier-vehicle systems. pages 1354 – 1359, 01 2011.

[24] J. C. Goodson, J. W. Ohlmann, and B. W. Thomas. Rollout policies for dynamic solutions to the multivehicle routing problem with stochastic demand and duration limits. *Operations Research*, 61(1):138–154, 2013.

[25] Quang Minh Ha, Yves Deville, Pham Quang Dung, and Minh Hoàng Hà. Heuristic methods for the traveling salesman problem with drone. *ArXiv*, abs/1509.08764, 2015.

[26] Quang Minh Ha, Yves Deville, Pham Quang Dung, and Minh Hoàng Hà. A hybrid genetic algorithm for the traveling salesman problem with drone. *ArXiv*, abs/1812.09351, 2018.

[27] Quang Minh Ha, Yves Deville, Quang Dung Pham, and Minh Hoàng Hà. On the min-cost traveling salesman problem with drone. *ArXiv*, abs/1512.01503, 2015.

[28] M. Haimovich and Thomas L. Magnanti. Extremum properties of hexagonal partitioning and the uniform distribution in Euclidean location. *SIAM J. Discrete Math.*, 1:50–64, 1988.

[29] Mark Harris. Amazon quietly acquired robotics company Dispatch to build Scout. `https://techcrunch.com/2019/02/07/meet-the-tiny-startup-that-helped-build-amazons-scout-robot/`, 2019. TechCrunch.

[30] D. S. Hochbaum. When are NP-hard location problems easy? *Annals of Operations Research*, 1:201–214, 1984.

[31] Dorit Hochbaum and J.Michael Steele. Steinhaus’s geometric location problem for random samples in the plane. 14, 03 1982.

[32] M. Huang, K. R. Smilowitz, and B. Balcik. A continuous approximation approach for assessment routing in disaster relief. *Transportation Research Part B: Methodological*, 50:20–41, 2013.

[33] O. Jabali, M. Gendreau, and G. Laporte. A continuous approximation model for the fleet composition problem. *Transportation Research Part B: Methodological*, 46(10):1591–1606, 2012.

[34] Ho Young Jeong, Byung Duk Song, and Seokcheon Lee. Truck-drone hybrid delivery routing: Payload-energy dependency and no-fly zones. *International Journal of Production Economics*, 214:220–233, 2019.

[35] Shucong Jia and Lin Zhang. Modelling unmanned aerial vehicles base station in ground-to-air cooperative networks. *IET Communications*, 11, 01 2017.

[36] Aline Karak and Khaled Abdelghany. The hybrid vehicle-drone routing problem for pick-up and delivery services. *Transportation Research Part C Emerging Technologies*, 102:427–449, 03 2019.

[37] S. Kim and I. Moon. Traveling salesman problem with a drone station. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(1):42–52, Jan 2019.

[38] Seon Jin Kim, Gino J. Lim, and Jaeyoung Cho. Drone flight scheduling under uncertainty on battery duration and air temperature. *Computers & Industrial Engineering*, 117:291 – 302, 2018.

[39] Seon Jin Kim, Gino J. Lim, Jaeyoung Cho, and Murray J. Côté. Drone-aided healthcare services for patients with chronic diseases in rural areas. *Journal of Intelligent & Robotic Systems*, 88:163–180, 2017.

[40] Patchara Kitjacharoenchai, Mario Ventresca, Mohammad Moshref-Javadi, Seokcheon Lee, Jose M.A. Tanchoco, and Patrick A. Brunese. Multiple traveling salesman problem with drones: Mathematical model and heuristic approach. *Computers & Industrial Engineering*, 129:14 – 30, 2019.

[41] J.K. Korpela. *Mathematical Expressions*. Suomen E-painos Oy, 2014.

[42] Yushan Li, Guangzhi Zhang, Zhibo Pang, and Lefei Li. Continuum approximation models for joint delivery systems using trucks and drones. *Enterprise Information Systems*, pages 1–30, 2018.

[43] G. J. Lim, S. Kim, J. Cho, Y. Gong, and A. Khodaei. Multi-uav pre-positioning and routing for power network damage assessment. *IEEE Transactions on Smart Grid*, 9(4):3643–3651, July 2018.

[44] Yao Liu, Zhihao Luo, Zhong Liu, Jianmai Shi, and Guangquan Cheng. Cooperative routing problem for ground vehicle and unmanned aerial vehicle: The application on intelligence, surveillance, and reconnaissance missions. *IEEE Access*, 7:63504–63518, 2019.

[45] Yao Liu, Jianmai Shi, Guohua Wu, Zhong Liu, and Witold Pedrycz. Two-echelon routing problem for parcel delivery by cooperated truck and drone, 02 2019. Preprint.

[46] Zhihao Luo, Zhong Liu, and Jianmai Shi. A two-echelon cooperated routing problem for a ground vehicle and its carried unmanned aerial vehicle. *Sensors*, 17(5):1144, 2017.

[47] Satyanarayana G Manyam, David W Casbeer, and Kaarthik Sundar. Path planning for cooperative routing of air-ground vehicles. In *2016 American Control Conference (ACC)*, pages 4630–4635. IEEE, 2016.

[48] Mario Marinelli, Leonardo Caggiani, Michele Ottomanelli, and Mauro Dell’Orco. En route truck-drone parcel delivery for optimal vehicle routing strategies. *IET Intelligent Transport Systems*, 12:253–261(8), May 2018.

[49] N. Mathew, S. L. Smith, and S. L. Waslander. Planning paths for package delivery in heterogeneous multirobot teams. *IEEE Transactions on Automation Science and Engineering*, 12(4):1298–1308, Oct 2015.

[50] Raïssa G. Mbiadou Saleu, Laurent Deroussi, Dominique Feillet, Nathalie Grangeon, and Alain Quilliot. An iterative two-step heuristic for the parallel drone scheduling traveling salesman problem. *Networks*, 72(4):459–474, 2018.

[51] Chase C. Murray and Amanda G. Chu. The flying sidekick traveling salesman problem: Optimization of drone-assisted parcel delivery. *Transportation Research Part C: Emerging Technologies*, 54:86 – 109, 2015.

[52] Chase C. Murray and R. Manu Raj. The multiple flying sidekicks traveling salesman problem: Parcel delivery with multiple drones. 2019.

[53] Phuong Nguyen, Christian Prins, and Caroline Prodhon. Solving the two-echelon location routing problem by a grasp reinforced by a learning process and path relinking. *European Journal of Operational Research*, 216:113–126, 01 2012.

[54] Alena Otto, Niels Agatz, James Campbell, Bruce Golden, and Erwin Pesch. Optimization approaches for civil applications of unmanned aerial vehicles (uavs) or aerial drones: A survey. *Networks*, 72(4):411–458, 2018.

[55] C. H. Papadimitriou. Worst-case and probabilistic analysis of a geometric location problem. *SIAM Journal on Computing*, 10:542, 1981.

[56] Stefan Poikonen. The multi-visit drone routing problem. submitted, 2018.

[57] Stefan Poikonen and Bruce Golden. The mothership and drone routing problem. to appear.

[58] Stefan Poikonen, Bruce Golden, and Edward A. Wasil. A branch-and-bound approach to the traveling salesman problem with a drone. *INFORMS Journal on Computing*, 31(2):335–346, 2019.

[59] Stefan Poikonen, Xingyin Wang, and Bruce Golden. The vehicle routing problem with drones: Extended models and connections. *Networks*, 70(1):34–43, 2017.

[60] Andrea Ponza. Optimization of drone-assisted parcel delivery. 2016.

[61] Dražen Popović, Milovan Kovač, and Nenad Bjelić. A miqp model for solving the vehicle routing problem with drones. In *4th Logistics International Conference*, 2019.

[62] Luigi Di Puglia Pugliese and Francesca Guerriero. Last-mile deliveries by using drones and classical vehicles. In *International Conference on Optimization and Decision Science*, pages 557–565. Springer, 2017.

[63] C. Redmond and J. E. Yukich. Limit theorems and rates of convergence for euclidean functionals. *The Annals of Applied Probability*, 4(4):pp. 1057–1073, 1994.

[64] Halil Savuran and Murat Karakaya. Route optimization method for unmanned air vehicle launched from a carrier. 2015.

[65] Daniel Schermer, Mahdi Moeini, and Oliver Wendt. A hybrid vns/tabu search algorithm for solving the vehicle routing problem with drones and en route operations. *Computers & Operations Research*, 109:134–158, 05 2019.

[66] Judy Scott and Carlton Scott. Drone delivery models for healthcare. In *Proceedings of the 50th Hawaii international conference on system sciences*, 2017.

[67] Marc-Oliver Sonneberg, Max Leyerer, Agathe Kleinschmidt, Florian Knigge, and Michael H Breitner. Autonomous unmanned ground vehicles for urban logistics: Optimization of last mile delivery operations. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019.

[68] J. M. Steele. Subadditive Euclidean functionals and nonlinear growth in geometric probability. *The Annals of Probability*, 9(3):pp. 365–376, 1981.

[69] J.M. Steele. *Probability Theory and Combinatorial Optimization*. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, 1987.

[70] Ziye Tang, Willem-Jan van Hoeve, and Paul Shaw. A study on the traveling salesman problem with a drone. In *International Conference on Integration of Constraint Programming, Artificial Intelligence, and Operations Research*, pages 557–564. Springer, 2019.

[71] Madjid Tavana, Kaveh Khalili-Damghani, Francisco Javier Santos Arteaga, and Mohammad-Hossein Zandi. Drone shipping versus truck delivery in a cross-docking system with multiple fleets and products. *Expert Systems with Applications*, 72:93–107, 12 2017.

[72] Pratap Tokekar, Joshua Vander Hook, David Mulla, and Volkan Isler. Sensor planning for a symbiotic uav and ugv system for precision agriculture. *IEEE Transactions on Robotics*, PP:1–1, 10 2016.

[73] M. Torabbeigi, G. J. Lim, and S. J. Kim. Drone delivery schedule optimization considering the reliability of drones. In *2018 International Conference on Unmanned Aircraft Systems (ICUAS)*, pages 1048–1053, June 2018.

[74] Maryam Torabbeigi, Gino Lim, and Seon Jin Kim. Drone delivery scheduling optimization considering payload-induced battery consumption rates. *Journal of Intelligent & Robotic Systems*, 05 2019.

[75] Marlin W. Ulmer and Barrett W. Thomas. Same-day delivery with heterogeneous fleets of drones and vehicles. *Networks*, 72(4):475–505, 2018.

[76] Antidio Viguria, Iván Maza, and Aníbal Ollero. Distributed service-based cooperation in aerial/ground robot teams applied to fire detection and extinguishing missions. *Advanced Robotics*, 24:1–23, 2010.

[77] Kangzhou Wang, Biao Yuan, Mengting Zhao, and Yuwei Lu. Cooperative route planning for the drone and truck in delivery services: A bi-objective optimisation approach. *Journal of the Operational Research Society*, 0(0):1–18, 2019.

[78] Xingyin Wang, Stefan Poikonen, and Bruce Golden. The vehicle routing problem with drones: Several worst-case results. *Optimization Letters*, 11, 04 2016.

[79] Zheng Wang and Jiuh-Biing Sheu. Vehicle routing problem with drones. *Transportation research part B: methodological*, 122:350–364, 2019.

[80] Guohua Wu, Witold Pedrycz, Haifeng Li, Manhao Ma, and Jin Liu. Coordinated planning of heterogeneous earth observation resources. *IEEE Transactions on Systems Man and Cybernetics - Part A Systems and Humans*, 10.1109/TSMC.2015.2431643, 01 2015.

[81] E. Wygonik and A. Goodchild. Evaluating the efficacy of shared-use vehicles for reducing greenhouse gas emissions: a US case study of grocery delivery. In *Journal of the Transportation Research Forum*, volume 51, 2012.

[82] W. Xie and Y. Ouyang. Optimal layout of transshipment facility locations on an infinite homogeneous plane. *Transportation Research Part B: Methodological*, 75:74–88, 2015.

[83] J.E. Yukich. *Probability Theory of Classical Euclidean Optimization Problems*. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2006.

Data Management Plan (Heading 1 style)

Products of Research

The data that were collected consist of uniformly sampled points in a geographic region as well as lat/lng pairs sampled from Southern California. All origin-destination distances can be computed using OpenStreetMaps, Google Maps, or HERE Maps.

Data Format and Content

There are no files to share; all experiments can be reproduced using only the contents of this paper.

Data Access and Sharing

The general public can access the data from this paper by repeating the experiments that we conducted, which merely require a random number generator.

Reuse and Redistribution

No restrictions to report.

Appendix

A Proof of Theorem 9

The following three lemmas are textbook-level results that we state without proof:

Lemma 17 (Stirling's approximation)**.** *The gamma function* $\Gamma(x)$ *satisfies*

$$\log \Gamma(x+1) = x \log x - x + \frac{1}{2}\log x + \frac{1}{2}\log 2 + \frac{1}{2}\log \pi + \mathcal{O}(1/x).$$

Lemma 18. *Let* $f : \mathbb{R} \mapsto \mathbb{R}$ *be a real-valued function and let* $\mathcal{B}_d(r) \subset \mathbb{R}^d$ *be the ball of radius* r *centered about the origin. We have*

$$\int_{\mathcal{B}_d(r)} f(\|x\|)dx = \int_0^r S_{d-1}(t) f(t)dt,$$

where $S_{d-1}(t)$ *is the surface area of a* $(d-1)$*-sphere of radius* t*, which is given by*

$$S_{d-1}(t) = \frac{2\pi^{d/2}}{\Gamma(d/2)} t^{d-1}.$$

Lemma 19. *The volume of a* d*-dimensional ball of radius* r *is* $\pi^{d/2} r^d / \Gamma(d/2+1)$*.*

Our final lemma to be employed in the lower bound argument is a consequence of Lemmas 18 and 19.

Lemma 20. *Let* $l > 0$ *and let* $\mathcal{D} \subset \mathbb{R}^{2n}$ *denote the set of all* n*-tuples* $(u_1, \dots, u_n)$ *of points in* $\mathbb{R}^2$ *such that*

$$\sum_{i=1}^{n} \|u_i\| < l.$$

The volume of $\mathcal{D}$*,* $\mathrm{Vol}(\mathcal{D})$*, satisfies*

$$\mathrm{Vol}(\mathcal{D}) = \frac{(2\pi)^n}{\Gamma(2n+1)} \cdot l^{2n}.$$

Proof. This volume can be expressed as the integral

$$\int_{\mathcal{B}_2(l)} \int_{\mathcal{B}_2(l-\|u_n\|)} \cdots \int_{\mathcal{B}_2(l-\sum_{i=3}^{n}\|u_n\|)} \int_{\mathcal{B}_2(l-\sum_{i=2}^{n}\|u_n\|)} 1 du_1 du_2 \cdots du_{n-1} du_n,$$

which we then compute by induction and application of Lemmas 18 and 19.

□

We are now ready to prove Theorem 9:

Proof of theorem 9. As $\beta_{\text{BMed}}(d)$ is independent of the demand distribution, we can arrive at a lower bound by first assuming we are in the case that the X_i are i.i.d. Unif$([0,1]^2)$. We employ the union bound.

$$\begin{aligned}
&\mathbb{P}(\mathrm{BMed}(X_1,\dots,X_n;d) < l)\\
&= \mathbb{P}(\text{some selection of, and assignment to, medians is of cost} < l)\\
&\leq \text{sum over all selections and assignments of } \mathbb{P}(\text{cost of selection, assignment} < l)\\
&= (\text{\# ways select medians})(\text{\# ways assign points})\mathbb{P}(\text{cost of arbitrary choice} < l)\\
&= \binom{n}{p} p^{n-p}\mathbb{P}(\text{cost of arbitrary selection and assignment} < l).
\end{aligned}$$

To obtain an upper bound on the probability that an arbitrary selection and assignment has cost less than l we fix our median indices, $\mathcal{S}$, and our assignment map μ and first recall that $X_{\mu(i)}$ is the median assigned to point X_i. Because we can reorder and adjust μ and $\mathcal{S}$ accordingly, we may assume without loss of generality that the X_i are ordered such that the medians we have selected are the last p points, $X_{n-p+1},\dots,X_n$. We then have the cost of a particular selection and assignment is given by $\sum_{i=1}^{n-p}\|X_i - X_{\mu(i)}\|$. We have

$$\mathbb{P}\left(\sum_{i=1}^{n-p}\|X_i - X_{\mu(i)}\| < l\right) =$$
$$\int_{(x_{n-p+1},\dots,x_n)\in[0,1]^{2p}} \mathbb{P}\left(\sum_{i=1}^{n-p}\|X_i - X_{\mu(i)}\| < l \,\middle|\, (X_{n-p+1},\dots,X_n) = (x_{n-p+1},\dots x_n)\right) dF_{(X_{n-p+1},\dots,X_n)}(x_{n-p+1},\dots x_n).$$

Let

$$\mathcal{E}(l,(x_{n-p+1},\dots,x_n)) := \left\{x_1,\dots,x_{n-p}\in\mathbb{R}^2 : \sum_{i=1}^{n-p}\|x_i - x_{\mu(i)}\| < l\right\}.$$

Then recalling that the X_i are drawn uniformly from the unit square we have for all $(x_{n-p+1},\dots,x_n)$,

$$\begin{aligned}
\mathbb{P}\left(\sum_{i=1}^{n-p}\|X_i - X_{\mu(i)}\| < l \,\middle|\, (X_{n-p+1},\dots,X_n) = (x_{n-p+1},\dots x_n)\right) &= \mathbb{P}((X_1,\dots,X_{n-p})\in\mathcal{E}(l,(x_{n-p+1},\dots,x_n)))\\
&= \mathrm{Vol}(\mathcal{E}(l,(x_{n-p+1},\dots,x_n))\cap[0,1]^2)\\
&\leq \mathrm{Vol}(\mathcal{E}(l,(x_{n-p+1},\dots,x_n))).
\end{aligned}$$

To compute this volume we make the volume preserving transformation $u_i := x_i - x_{\mu(i)}$ (that is, translate each median point to the origin and move its assigned set commensurately) and consider

$$\mathcal{E}'(l) := \left\{u_1,\dots,u_{n-p}\in\mathbb{R}^2 : \sum_{i=1}^{n-p}\|u_i\| < l\right\}.$$

Clearly for all possible median locations $(x_{n-p+1},\dots,x_n)$ and all of our choices of μ, we have that $\mathrm{Vol}(\mathcal{E}(l,\{x_{n-p+1},\dots,x_n\})) = \mathrm{Vol}(\mathcal{E}'(l))$ which is only dependent on l. By Lemma 19 we have

$$\mathrm{Vol}(\mathcal{E}'(l)) = \frac{(2\pi)^{n-p}}{\Gamma(2(n-p)+1)}\cdot l^{2(n-p)}.$$

Thus for all selections of $\mathcal{S}$ and μ,

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^{n-p}\|X_i - X_{\mu(i)}\| < l\right) &\leq \int \mathrm{Vol}(\mathcal{E}'(l))dF_{(X_{n-p+1},\dots,X_n)}\\ &= \mathrm{Vol}(\mathcal{E}'(l))\\ &= \frac{(2\pi)^{n-p}}{\Gamma(2(n-p)+1)}\cdot l^{2(n-p)}.\end{aligned}$$

Combining the above

$$\mathbb{P}(\mathrm{BMed}(X_1,\dots,X_n;d) < l) \leq \binom{n}{p} p^{n-p}\cdot\frac{(2\pi)^{n-p}}{\Gamma(2(n-p)+1)}\cdot l^{2(n-p)}.$$

Taking logarithms then yields

$$\begin{aligned}\log\mathbb{P}(\mathrm{BMed}(X_1,\dots,X_n;d) < l) &= \log\left(\binom{n}{p}p^{n-p}\cdot\frac{(2\pi)^{n-p}}{\Gamma(2(n-p)+1)}\cdot(c\sqrt{n})^{2(n-p)}\right)\\ &\leq \log\left(\frac{\Gamma(n+1)}{\Gamma(p+1)\Gamma(n-p+1)}\cdot p^{n-p}\cdot\frac{(2\pi)^{n-p}}{\Gamma(2(n-p)+1)}\cdot l^{2(n-p)}\right)\\ &\leq \Big[\log(p)+\log(2)+\log(\pi)+2\log(l)+\log(n)+2-\log(n-p)-\\ &\quad 2\log(2n-2p)\Big]n - 2p\log(l) - p\log(\pi) - 2p\log(p) + \log(n-p)p+\\ &\quad 2\log(2n-2p)p - p\log(2) - \log(\pi) - \log(2) - 2p + \mathcal{O}\left(\frac{1}{n}\right),\end{aligned}$$

where we have employed Lemma 17. It is clear that this upper bound goes to negative infinity as n goes to infinity if and only if the coefficient of n is negative. For large n we are justified in ignoring the ceiling and substituting $n/(d+1)$ for p. We then have as $n\to\infty$,

$$\begin{aligned}&\mathbb{P}(\mathrm{BMed}(X_1,\dots,X_n;d) < l) \to 0\\ &\Uparrow\\ &\log\left(\frac{n}{d+1}\right)+\log(2)+\log(\pi)+2\log(l)+\log(n)+2-\log\left(n-\frac{n}{d+1}\right)-2\log\left(2n-2\frac{n}{d+1}\right) < 0\\ &\Updownarrow\\ &l < \frac{\sqrt{2}d^{(3/2)}}{e\sqrt{\pi}(d+1)}\cdot\sqrt{n}.\end{aligned}$$

The result then follows easily from the almost sure convergence to β_{BMed}. □

B Proof of Lemma 15

If f is absolutely continuous, then it can be approximated arbitrarily well with finitely many step functions on $\mathcal{R}$, and we therefore assume without loss of generality that f takes precisely this form. To be more specific, we assume that $f(x) = \sum_{i=1}^{m} f_i \delta_i(x)$, where $\delta_i(x)$ is an indicator function representing membership in a square grid cell i. Let ϵ denote the area of each grid cell, so that $\iint_{\mathcal{R}} f(x)\,\mathrm{d}x = \sum_{i=1}^{m} \epsilon f_i = 1$, and let N_i denote the number of samples of $\{P_1, \ldots, P_n\}$ that belong to cell i (so that $\sum_{i=1}^{m} N_i = n$).

It is clear that we can construct a feasible tour by applying Lemma 14 to each grid cell and then “stitching” the tours within each grid cell together. Certainly, the amount of time needed to visit all N_i points in grid cell i is at most

$$\sqrt{\epsilon}\left(\frac{2\sqrt{3}}{\sqrt{\phi_0\phi_1 k}} \cdot \sqrt{N_i} + C_i\right)$$

for some constant C_i, and the amount of additional time needed to “stitch” all of the tours together is a constant C_0 that does not depend on n. Summing all of these together and letting $C = \sum_{i=0}^{m} C_i$, we have

$$\begin{aligned} T(P_1, \ldots, P_n) &\leq \sqrt{\epsilon}\sum_{i=1}^{m}\left(\frac{2\sqrt{3}}{\sqrt{\phi_0\phi_1 k}} \cdot \sqrt{N_i} + C_i\right) + C_0 \\ \implies \frac{T(P_1, \ldots, P_n)}{\sqrt{n}} &\leq \sqrt{\epsilon}\sum_{i=1}^{m}\left(\frac{2\sqrt{3}}{\sqrt{\phi_0\phi_1 k}} \cdot \sqrt{\frac{N_i}{n}}\right) + \frac{C}{\sqrt{n}} \end{aligned}$$

and since $N_i/n \to \epsilon f_i$ with probability one, we see that

$$\limsup_{n\to\infty} \frac{T(P_1, \ldots, P_n)}{\sqrt{n}} \leq \sqrt{\epsilon}\sum_{i=1}^{m} \frac{2\sqrt{3}}{\sqrt{\phi_0\phi_1 k}} \cdot \sqrt{\epsilon f_i} = \frac{2\sqrt{3}}{\sqrt{\phi_0\phi_1 k}} \sum_{i=1}^{m} \epsilon\sqrt{f_i} = \frac{2\sqrt{3}}{\sqrt{\phi_0\phi_1 k}} \iint_{\mathcal{R}} \sqrt{f(x)}\,\mathrm{d}x$$

as desired.