Optimal Policies for Platooning and Ride Sharing in Autonomy-Enabled Transportation

Aviv Adler, David Miculescu, and Sertac Karaman**

Massachusetts Institute of Technology,
77 Massachusetts Avenue, Cambridge, MA, 02139
{adlera, dmicul, sertac}@mit.edu,

Abstract. Rapid advances in autonomous-vehicle technology may soon allow vehicles to platoon on highways, leading to substantial fuel savings through reduced aerodynamic drag. While these aerodynamic effects have been widely studied, the systems aspects of platooning have received little attention. In this paper, we consider a class of problems, applicable to vehicle platooning and passenger ride-sharing, from the systems perspective. We consider a system in which vehicles arrive at a station according to a stochastic process. If they wait for each other, they can form platoons to save energy, but at the cost of incurring transportation delays. Our analysis explores this tradeoff between energy consumption and transportation delays. Among our results is the derivation of the Pareto-optimal boundary and characterization of the optimal polices in both the open-loop and feedback regimes. Surprisingly, the addition of feedback improves the energy-delay curve very little when compared to open-loop policies.

Keywords: Logistics, Autonomy-Enabled Transportation Systems, Platooning, Optimization

1 Introduction

Autonomous vehicles hold the potential to revolutionize transportation. Autonomous cars may enable effective mobility-on-demand systems [1]. Autonomous drones may allow us to deploy the next generation urban logistics infrastructure [2, 3]. Even though tremendous effort has been invested in developing autonomous vehicles themselves, it is still unclear how teams of autonomous vehicles may operate as an efficient and sustainable transportation system.

We draw inspiration from the development of technology that allows autonomous vehicles to follow each other very closely [4], beyond what human drivers and pilots can safely do. This technology may allow road/air vehicles to drive/fly in platoons/formation in order to save energy. It is known that platooning trucks [4–9] and formation flying airplanes [10, 11] can save a substantial amount of fuel, simply by mitigating the effects of aerodynamic drag. In fact, many bird species frequently utilize formation flight precisely for this purpose

** The first two authors contributed equally, and they are listed in alphabetical order.

K. Goldberg et al. (Eds.): *Algorithmic Foundations of Robotics XII*, SPAR 13, pp. 848–863, 2020.
https://doi.org/10.1007/978-3-030-43089-4_54

on long-distance flights [12, 13]. Similarly, bikers and race car drivers often pack behind a leader in long-distance competitions [14], a behavior called *drafting*.

Motivated by this potential, we study the systems aspects of this technology from the perspective of sustainability-efficiency tradeoffs, in a mathematically rigorous way. Even though aerodynamics of platooning and control policies for best platooning results have been studied, to our knowledge this paper is the first to study its systems aspects from a queueing perspective as in [15, 16].

Specifically, we consider vehicles arriving at an initial location, from here on called the station, and all headed to the same destination. If the vehicles wait at the station for each other, then they can travel in platoons to reduce energy consumption per vehicle. However, clearly, forming long platoons will require some vehicles to wait for others, which will impact the average delay. In this paper, we explore this fundamental "energy-delay tradeoff."

Our contributions include the following: formalizing the problem, including formal definitions for station control policies and the energy-delay tradeoff; several general results, greatly narrowing down the range of policies we need to find optimal policies; a full characterization of the optimal delay-energy combinations, as well as a description of a class of policies which achieve them; and a description and analysis of a class of open-loop policies, which are potentially much simpler to implement. Surprisingly, in many cases, feedback policies show little improvement compared to open-loop policies, which we show rigorously.

Our results apply directly to truck platooning, air vehicle formation flight, and bicycle and race car drafting. These results also apply to ride-sharing (or pooling) systems [17, 18]. Consider passengers arriving at a bus station. When should each bus head out? What policies guarantee minimize delays and energy consumption (and the negative environmental impact of transportation) at the same time? Even though we utilize terminology from platooning, drafting, and formation flight throughout the paper, our results also apply to ride sharing systesm, which may become prominent with the introduction of autonomous vehicles [18].

Furthermore, while most recent rigorous work in transportation is on efficiency metrics, for instance focusing on capacity (throughput) and delay tradeoffs [19], our study involves the tradeoff between efficiency and sustainability, in this case, energy and delay. The study of the tradeoff between efficiency and sustainability in autonomy-enabled transportation systems may become more prominent in the future, as autonomous vehicles are deployed broadly.

This paper is organized as follows. In Section 2, we introduce our model and formalize the metrics of evaluation. In Section 3, we provide a set of general results characterizing optimal policies. Then, we focus on Poisson arrivals and restrict our attention to certain classes of energy usage models. In Section 4, we analyze optimal feedback policies. In Section 5, we analyze optimal open-loop policies. In Section 6, we compare these two policies, and conclude with remarks.

Due to space constraints some of our proofs are omitted and replaced either with sketches or the underlying intuition. The complete proofs are available in our full online paper [20].

2 Problem Setup

Even though the problem setup is relevant for both platooning and pooling, we motivate the rest of the paper with the truck platooning problem and use terminology from platooning. We remind the reader that all results reported in this paper can be applied to similar pooling problems as well.

We are interested in identifying the *energy-delay tradeoff* for platooning vehicles running between two locations connected by a highway. The vehicles arrive at an initial location, called the station, at random time intervals. All vehicles are headed to the same final location. In this case, if the vehicles wait for each other, then they can platoon to reduce energy consumption. Platoons with many vehicles are clearly better in terms of energy consumption; however, to form large platoons, the vehicles need to wait long periods of time for each other, which increases delay. Our aim is to analyze this fundamental energy-delay tradeoff.

The model includes two essential components: the vehicle arrival model and the energy consumption of a platoon; there are two metrics of performance: average delay and average energy per vehicle. We consider two types of policies: time-table policies and feedback optimal policies. In what follows, we formalize these six notions, leading to a formal problem definition.

Vehicle Arrival Model: Generically, the vehicles arrive at the station according to the stochastic process $\{t_i : i \in \mathbb{Z}_{>0}\}$, where t_i is the (random) arrival time for the ith vehicle (so $t_1 \leq t_2 \leq \dots$). In this paper, we assume that the arrival process $\{t_i : i \in \mathbb{Z}_{>0}\}$ is a Poisson process with intensity λ. That is, the inter-arrivals $a_{i+1} = t_{i+1} - t_i$ are independent random variables whose distributions are exponential with parameter λ.

Energy Consumption Model: The energy usage of a platoon of vehicles is a function of the number of vehicles in the platoon. This function is called the *total energy function*, and denoted by $\mathrm{TE}\colon \mathbb{Z}_{\geq 0} \to \mathbb{R}_{\geq 0}$, where $\mathrm{TE}(k)$ denotes the combined energy usage of a platoon consisting of k vehicles, which is called a k-platoon. By convention, we define $\mathrm{TE}(0) = 0$.

In the case of vehicle platooning, these functions are determined by the aerodynamics of drafting for the particular vehicles and speeds involved. In the case of pooling, these relate to the energy characteristics of the engine that powers the vehicles carrying the passengers.

We also define a particular class of total energy functions, which are simple but capture the intuition that while the first and last vehicles in a platoon pay higher energy costs, vehicles in the middle of a platoon spand roughly the same energy as each other. Thus, each vehicle will have an energy cost c, with an additional cost of c^* which is not influenced significantly by the number of vehicles in the platoon. We call this the class of *affine energy functions*:

Definition 1. *The* affine energy function *with parameters $c \geq 0$ and $c^* > 0$ is* $\mathrm{TE}^{\mathrm{aff}}(k) = ck + c^*$ *for $k > 0$, and* $\mathrm{TE}^{\mathrm{aff}}(0) = 0$.

This has been previously used in studying platooning optimization problems [5].

Control policies: We formally define the state of the system at time t as $x(t) := \big(t, (t_i^{\mathrm{arr}}, t_i^{\mathrm{dep}})_{i=1}^n\big)$, where t denotes the current time in the system, and $(t_i^{\mathrm{arr}}, t_i^{\mathrm{dep}})$ is the arrival time and departure time of the ith vehicle to enter the system. If a vehicle has arrived but not departed, then we formally define the departure time of such a vehicle as ∞. Note that the state of the system incorporates the entire history of arrivals and departure in the system, but only up until the current time t. We denote the collection of all states as the state space $\mathcal{X}$:

$$\mathcal{X} := \Big\{\big(t, (t_i^{\mathrm{arr}}, t_i^{\mathrm{dep}})_{i=1}^n\big) \colon\ 0 \le t_i^{\mathrm{arr}} \le t\,,\ t_i^{\mathrm{arr}} \le t_i^{\mathrm{dep}}\,,\ t_i^{\mathrm{arr}} \le t_{i+1}^{\mathrm{arr}}\,, n \in \mathbb{N}\Big\}.$$

Let $\mathcal{Y}$ be a predetermined sequence of random variables. We define a *control policy* $\pi\colon (\mathcal{X}, \mathcal{Y}) \to 2^{\mathbb{N}}$ (the $\mathcal{Y}$ is there to accommodate control policies with internal randomness), identifying which vehicles must leave the station as a platoon at the current time t. Note that a control policy can control a vehicle at the instant of its arrival in the station, but has no knowledge of future arrivals.

We also define the notion of *regularity* of a policy. This is to capture the intuition that whatever is optimal now should be optimal also in the future, and to avoid the possibility of a policy which changes permanently over time and prevents average delay and/or energy consumption from converging. In order to properly define it, we first assume that any reasonable policy will *sometimes* send all vehicles out of the station; otherwise, there will be always be at least one vehicle at the station, and we can immediately improve the policy by always having one fewer vehicle at the station (without altering when platoons leave).

Definition 2 (Reset Points and Regular Policies). *Suppose we have a policy π and it has been running up until time t^*. Then, t^* is a* reset point *if:*

- *no vehicles are present at the station at time t^**
- *the distribution of the future behavior of the system is identical to the distribution of future behavior at $t = 0$*

A policy π is regular *if the probability of getting a reset point at some time in the future is* 1*, and the expected time between reset points is finite.*

The fundamental problem considered here is the tradeoff between *energy consumption* and *delay* on the vehicles. The station seeks to minimize energy usage by combining vehicles into platoons. However, in order to do this, it will have to hold some vehicles at the station to wait for additional vehicles to arrive. The amount of time a vehicle spends at the station is its *delay*. Both energy consumption and delay for a policy are calculated as *averages per vehicle* as the policy runs over long time horizons.

Definition 3 (Average Energy Consumption Metric). *Suppose we run policy π on a Poisson-distributed sequence of vehicles (indexed by order of arrival). Let k_i be the size of the platoon in which vehicle i leaves. We define $E(i) = \frac{\mathrm{TE}(k(i))}{k(i)}$. If there is some E such that*

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n E(i)}{n} = E \textit{ with probability } 1$$

then we refer to this as the average energy *of policy π and denote it by* $\mathbb{E}[E(\pi)]$.

Note that in this definition, we are using $E(i) = \frac{\mathrm{TE}(k(i))}{k(i)}$ – the average energy consumption of a vehicle in vehicle i's platoon – as a stand-in for the amount of energy actually consumed by vehicle i. Individually, vehicle i may consume a different amount of energy, but this averages out correctly when considering all the vehicles in vehicle i's platoon.

Definition 4 (Average Delay Metric). *Suppose we run policy π on a Poisson-distributed sequence of vehicles (indexed by order of arrival). If there is some D such that*

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n}(t_i^{\mathrm{dep}} - t_i^{\mathrm{arr}})}{n} = D \textit{ with probability } 1,$$

then we refer to this as the average delay *of policy π and denote it by* $\mathbb{E}[D(\pi)]$.

Note that these definitions require that the average energy or delay per vehicle converges to a predictable value with probability 1 as the policy is run over an infinite time. Although this may seem strict, it is necessary since otherwise there is no clear way to measure a policy's performance in terms of saving energy and delay. We restricted ourselves to regular policies because, in addition to being intuitively the only reasonable kind of policy, the regularity condition guarantees that the above metrics are well-defined:

Lemma 1. *If π is regular, then $\mathbb{E}[E(\pi)]$ and $\mathbb{E}[D(\pi)]$ exist.*

The proof is straightforward, but due to space constraints we omit it.

Now that we have a rigorous notion of the energy consumption and delay incurred by different policies, we can discuss what it means for a policy to be 'efficient' with respect to these metrics. We use the notion of *Pareto optimality*, in which a policy is optimal for a class Π if no other policy in Π can improve one metric without performing worse on the other.

Definition 5 (Pareto-optimality). *A policy π is* Pareto-optimal *among a class of policies Π if there is no other policy $\pi' \in \Pi$ such that $\mathbb{E}[E(\pi')] \leq \mathbb{E}[E(\pi)]$ and $\mathbb{E}[D(\pi')] \leq \mathbb{E}[D(\pi)]$ and at least one of the two inequalities is strict.*

When implementing a policy, one should always choose from the set of Pareto-optimal policies, if possible. This then leads to the notions of the *profile* of policy π and the *energy/delay tradeoff curve*:

Definition 6 (Profile and energy-delay curve). *The* profile *of policy π is the point $(\mathbb{E}[D(\pi)], \mathbb{E}[E(\pi)]) \in \mathbb{R}^2$. The* energy-delay curve *is the set of profiles of Pareto-optimal policies.*

Equivalent ways of viewing the energy-delay curve are: (1) as the lower convex hull of the set of profile points of policies; (2) the curve representing solutions of the optimization problem, "minimize average energy among all policies with average delay at most D".

All-or-Nothing Policies: We also note that, intuitively, keeping vehicles at the station when a platoon leaves is wasteful – if we wanted to send a k-platoon, we should do it when there are k vehicles rather than wait for additional vehicles to arrive. This leads to the notion of an *All-or-Nothing policy*, which sends every vehicle present with every platoon.

Definition 7. *An* all-or-nothing policy *is a mapping* $\pi : (\mathcal{X}, \mathcal{Y}) \to \{\emptyset, \mathbb{N}\}$.

In fact, we will show later that *only* All-or-Nothing policies (and others that behave similarly) are optimal (see Theorem 2). Hence, we can restrict ourselves to considering only these policies.

We now introduce the two specific kinds of policies we will analyze: *time-table* and *mixed-threshold* policies. We are interested in time-table policies because they require almost no information on the state of the system, and hence can be much easier and quicker to implement. Mixed-threshold policies are also relatively simple (though they require more information on the state of the system) and simple to analyze, and, as it turns out, there is always a Pareto-optimal policy of this form. Both are by definition All-or-Nothing policies.

Time-Table Policies: General policies are allowed to use a significant range of information and internal randomization to make their decisions, while time-table policies are restricted to have no randomization and make decisions based only on the elapsed time since the beginning of the process. Formally:

Definition 8. *A* time-table policy *is a policy* $\pi : \mathbb{R}_+ \to \{0, \mathbb{N}\}$.

Thus, a time-table policy is really just a list of predetermined times for which it sends all vehicles out of the station – hence the name. Because of the regularity condition, time-table policies must have reset points after which their behavior repeats. Thus, a time-table policy can be thought of in terms of a repeating cycle of intervals during which vehicles collect at the station; at the end of each interval, the vehicles exit.

We define the (ordered) set of *waiting times* $(T_1, T_2, \ldots, T_M)$ so that $T_m > 0$ for all m, and denote $C := \sum_{m=1}^{M} T_m$ as the cycle time, *i.e.*, the waiting times repeat every C time units in the same order. In other words, let $\mathcal{T}$ be the set of times when vehicles exit the station. Then, $\mathcal{T}$ has the following structure:

$$\mathcal{T} := \left\{ kC + \sum_{m=1}^{n} T_m : \quad k \in \mathbb{Z}_{\geq 0},\, n \in \{1, 2, \ldots, M\} \right\},$$

where $k \geq 0$ counts the number of cycles and $n \in \{1, \ldots, M\}$ tracks the current interval T_n. Thus, for the remainder of the paper, we will use the simpler but equivalent representation of time-table policies:

$$\Pi := \left\{ (T_1, T_2, \ldots, T_M) \in \mathbb{R}^M_{>0} : \; M \in \mathbb{Z}_{>0} \right\}$$

We note that a time-table policy has a reset point every time the cycle ends.

Mixed-Threshold Policies: Another simple type of policy is one which determines what distribution of platoon sizes it wants to send and directly produces them.

A *mixed-threshold* policy π is determined by a probability distribution $p(\pi) = (p_1, p_2, \dots)$ over platoon sizes; p_k is the probability that any given platoon will have k vehicles. Every time it sends out a platoon, some k is drawn according to $p(\pi)$ (independently of everything that has happened so far) and the next platoon leaves when the kth vehicle arrives at the station.

3 General Results

We present some general results pertaining to this optimization problem.

Theorem 1 (Mixing Regular Policies). *Given two regular policies π, π' and some $\alpha \in [0,1]$, it is possible to create a 'mixed' policy π^* such that*

$$\mathbb{E}[D(\pi^*)] = \alpha\mathbb{E}[D(\pi)] + (1-\alpha)\mathbb{E}[D(\pi')] \text{ and } \mathbb{E}[E(\pi^*)] = \alpha\mathbb{E}[E(\pi)] + (1-\alpha)\mathbb{E}[E(\pi')]$$

Proof. π^* is constructed by using the "cycles" of π and π'. A *cycle* of a policy is an interval between consecutive reset points. By the definition of regularity and because vehicle arrivals are Poisson, a policy can be viewed as a sequence of independent cycles. Furthermore, each cycle has finite expected length (by definition of regularity) and therefore a finite expected number of vehicle arrivals. Let $L(\pi)$ be the expected number of vehicle arrivals in a cycle of π, and $L(\pi')$

We then construct π^* as follows: randomly decide between π and π', and run one cycle; then repeat. The random decisions are i.i.d. (so every time a cycle ends, we have a reset point of π^*). The probability of picking π is:

$$\mathbb{P}[\text{pick } \pi \text{ to control next cycle}] = \frac{\alpha L(\pi')}{\alpha L(\pi') + (1-\alpha)L(\pi)}$$

This is because we want each *vehicle* to be controlled under π with probability α, and under π' with probability $(1-\alpha)$ – and thus, we pick more often from the policy with shorter cycles to address this imbalance. Thus, by linearity of expectations, the average delay and energy per vehicle under π^* are $\alpha\mathbb{E}[D(\pi)] + (1-\alpha)\mathbb{E}[D(\pi')]$ and $\alpha\mathbb{E}[E(\pi)] + (1-\alpha)\mathbb{E}[E(\pi')]$ respectively. ■

Corollary 1. *The energy-delay curve for regular policies is convex.*

Proof. This is a simple consequence of Theorem 1, since it shows that any point on the line segment between the profiles of π and π' is the profile of some π^*. Thus, the set of profile points is a convex set and so its lower boundary (which is the energy-delay curve) is a convex function. ■

We now define the *platoon distribution* of a given policy π, which is a table of what fraction of platoons have k vehicles and leave behind j at the station.

Definition 9 (Platoon Distribution). *Consider running a policy π. Let $\rho_{j,k}^{(m)}$ denote the number of platoons of the first m which leave with k vehicles while j stay behind at the station. Then, if there is some set of values $\{\rho_{j,k}\}$ (for $j \in \mathbb{Z}_{\geq 0}$ and $k \in \mathbb{Z}_{>0}$) such that $\lim_{m\to\infty} \frac{\rho_{j,k}^{(m)}}{m} = \rho_{j,k}$ with probability 1 for all j, k, we call $\{\rho_{j,k}\}$ the* platoon distribution *of π, and denote it $\rho(\pi)$.*

Note that $\rho(\pi)$ is really a probability distribution over $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{>0}$. For this notion to be useful, we want to make sure that it exists for regular policies.

Lemma 2. *For any regular policy π, $\rho(\pi)$ exists.*

This result is similar to Lemma 1.

Note that for an all-or-nothing policy, $\rho_{j,k} = 0$ for all $j > 0$ since no vehicles are ever left behind. The platoon distribution is important because it turns out that the platoon distribution of a policy π can be used to calculate the average energy and average delay.

Theorem 2 (Policy Equivalence from Platoon Dist.). *Let π, π' be policies such that $\rho(\pi) = \rho(\pi')$. Then $\mathbb{E}[D(\pi)] = \mathbb{E}[D(\pi')]$ and $\mathbb{E}[E(\pi)] = \mathbb{E}[E(\pi')]$.*

Proof (sketch). Since $\rho(\pi) = \rho(\pi')$, we can use $\rho_{j,k}$ to refer to elements of both; we also refer to the time between consecutive platoons as a *platoon interval.*

First, we note that by definition, the probability of getting a platoon of size k is $p_k = \sum_{j=0}^{\infty} \rho_{j,k}$ (this is of course the same for both policies). Since energy consumption of a platoon depends only on its size, $\mathbb{E}[E(\pi)] = \mathbb{E}[E(\pi')]$.

Now we turn to the delay. The basic intuition for why the average delay is the same in both policies is that waiting with m vehicles at the station has the same effect no matter what happened in the past, because future vehicle arrivals are Poisson-distributed and hence not dependent on past events. Formally, the proof works by computing $\mathbb{E}[D(\pi)]$ using only $\rho(\pi)$ – which automatically shows that $\mathbb{E}[D(\pi')]$ can be computed the same way and is therefore the same.

The basic method is to consider running π, and look at the set S_m of times where exactly m vehicles are at the station. If we allow empty segments, each platoon interval has exactly one segment of time with m vehicles. Furthermore, each segment can end either because (i) a new vehicle arrived or (ii) a platoon leaves. We will compute the probability of (i) from $\rho(\pi)$.

Note that if we piece together the segments of S_m, including where new vehicles arrived, we get a Poisson distribution of new vehicle arrivals (as Poisson arrivals are independent of the past, and π is only dependent on the past).

But we also note that a segment of S_m ends with a new vehicle if and only if the platoon interval starts with m or fewer vehicles and ends with more than m vehicles, which can be calculated from $\rho(\pi)$. This, and the insight that new vehicle arrivals on S_m are Poisson, gives us the average length (empty segments included) of a segment of S_m, which gives us the average amount of time in a platoon interval spent with exactly m vehicles at the station. Since we can compute this for all m, we also get the average length of a platoon interval.

This then gives us the fraction of total time with m vehicles at the station (when m vehicles are accruing delay) – call this fraction α_m. Thus, the average amount of delay being accrued at any moment as $\sum_{m=0}^{\infty} m\alpha_m$. Since the average number of vehicles which show up in a unit of time is λ, this means if we divide the total amount of delay accrued over a long period of time by the number of vehicles which arrived during that time, we get (with cancellation of T):

$$\mathbb{E}[D(\pi)] = \lim_{T\to\infty} \left(\frac{\text{total delay in } T \text{ time}}{\#\text{ vehicles in } T \text{ time}} \right) = \frac{\sum_{m=0}^{\infty} m\alpha_m}{\lambda}$$

with the approximation converging to the right-hand side as $T \to \infty$.

Thus, we have shown that we can compute $\mathbb{E}[D(\pi)]$ using only $\rho(\pi)$. ■

This proof has a very useful consequence: it can be used to show that for every policy π, there is an all-or-nothing policy π^* which performs at least as well; and if it has a positive probability of leaving behind vehicles when sending a (nonempty) platoon, then π^* has strictly lower average delay.

Theorem 3 (Optimality of All-or-Nothing). *Let π be a regular policy, and π^* be the mixed-threshold policy whose probability of sending a k-platoon is $p_k^* = \sum_{j=0}^{\infty} \rho_{j,k}$ for all k (that is, the mixed-threshold policy with the same proportion of k-platoons as π). Then, $\mathbb{E}[E(\pi)] = \mathbb{E}[E(\pi^*)]$ and $\mathbb{E}[D(\pi)] \geq \mathbb{E}[D(\pi^*)]$ with a strict inequality if and only if there is some (j', k') such that $j' > 0$ and $\rho_{j',k'} > 0$.*

The proof of this relies on the fact that π^* has the same distribution of platoon sizes as π. This immediately gives $\mathbb{E}[E(\pi)] = \mathbb{E}[E(\pi^*)]$.

For the delay, we use the method from the proof of Theorem 2 to compute the average delay of π and π^*. If there is no (j', k') such that $j' > 0$ and $\rho_{j',k'} > 0$, this means that π and π^* have exactly the same platoon distribution (and π is effectively all-or-nothing since it never leaves vehicles behind at the station) and so we apply Theorem 2 to get $\mathbb{E}[D(\pi)] = \mathbb{E}[D(\pi')]$. If there is such a (j', k'), then we can show algebraically (through the notion of stochastic dominance) that $\mathbb{E}[D(\pi)] > \mathbb{E}[D(\pi^*)]$.

4 On Pareto-optimal Regular Policies

We now consider the problem of finding policies which are Pareto-optimal over all regular policies. We do this by using the machinery from the previous section. In particular, we know from Theorem 3 that we do not need to consider policies which are not all-or-nothing. We then analyze the average delay and energy of these policies. First, we define the class of *hard-threshold* policies:

Definition 10. *The* hard-threshold *policy $\bar{\pi}_k$ is the policy which sends all vehicles out of the station whenever there are k vehicles (or more) waiting.*

The "or more" is just for completeness, as such a policy would prevent the queue from becoming longer than k vehicles. Furthermore, a hard-threshold policy has a reset point every time a platoon is sent out.

Theorem 4 (Mixed-Threshold is Optimal). *Any point on the energy-delay curve can be achieved by a mixed-threshold policy with at most two platoon sizes.*

Proof. Recall that a mixed-threshold policy is one which pre-selects the distribution of platoon sizes it produces with a probability distribution $(p_1, p_2, \dots)$, with p_k representing the probability of a platoon having k vehicles; after every platoon is sent, it chooses the size k of the next platoon from this distribution (independently) and sends the next platoon when the kth vehicle arrives.

If (D, E) is on the energy-delay curve, then there must be a policy π for which it is the profile. But by Theorem 3, we can use $\rho(\pi)$ to construct the mixed-threshold policy π^* where $p_k = \sum_{j=0}^{\infty} \rho_{j,k}$, and π^* will then by definition achieve (D, E) as well.

We now want to show why at most two unique platoon sizes are necessary. We note that mixed-threshold policies can be expressed as probabilistic mixtures of the set of policies $\{\bar{\pi}_k\}$, as in Theorem 1. Thus, the set of all profiles of mixed-threshold policies is just the convex hull of the set of profiles of $\{\bar{\pi}_k\}$.

But the energy-delay curve lies on the boundary of this set, and thus any point on the curve is on a line between at most two profile points – and thus is the profile of a policy that mixes at most two elements of $\{\bar{\pi}_k\}$. ■

In order to gain a complete characterization of the energy-delay curve, we compute the average energy and delay of hard-threshold policies:

Lemma 3 (Delay and Energy of Threshold Policies). *If the vehicles arrive with intensity λ, then $\mathbb{E}[D(\bar{\pi}_k)] = \frac{k-1}{2\lambda}$ and $\mathbb{E}[E(\bar{\pi}_k)] = \frac{\mathrm{TE}(k)}{k}$*

Proof. Because $\bar{\pi}_k$ has a reset point every time it sends a platoon, we can examine one platoon in isolation. By definition, every platoon it sends has k vehicles and hence $\mathbb{E}[E(\bar{\pi}_k)] = \frac{\mathrm{TE}(k)}{k}$ is immediate.

Recall that $a_{i+1} = t_{i+1}^{\mathrm{arr}} - t_i^{\mathrm{arr}}$ (the amount of time between the arrivals of vehicles i and $i+1$). Because arrivals are Poisson with intensity λ, we know that a_{i+1} follows an exponential distribution with parameter λ, which in turn means $\mathbb{E}[a_{i+1}] = 1/\lambda$ for all i. But the amount of time that vehicle i waits before the platoon leaves is $t_k^{\mathrm{arr}} - t_i^{\mathrm{arr}} = \sum_{j=i+1}^{k} a_j$, which means that $\mathbb{E}[t_k^{\mathrm{arr}} - t_i^{\mathrm{arr}}] = \frac{k-i}{\lambda}$. Averaging this over vehicles $i = 1, 2, \ldots, k$ then gives $\mathbb{E}[E(\bar{\pi}_k)] = \frac{k-1}{2\lambda}$. ■

We can now give the form of the energy-delay curve:

Corollary 2. *The energy-delay curve is piecewise-linear, with its vertices at $\left(\frac{k_i-1}{2\lambda}, \frac{\mathrm{TE}(k_i)}{k_i}\right)$ for some (increasing) sequence of integers $k_1, k_2, \ldots$.*

Proof. This follows from Lemma 3 and Theorem 4. The sequence $\{k_i\}$ just represents the profiles of vertices the lower convex hull. ■

Finally, we look at the optimal policies for the affine total energy function:

Corollary 3. *When the total energy function is $\mathrm{TE}^{\mathrm{aff}}$, the Pareto-optimal policy achieving delay at most x is the mixed-threshold policy which, letting $k^* = \lfloor 2\lambda x + 1 \rfloor$ and $p^* = (2\lambda x + 1) - k^*$, sends a platoon of size $k^* + 1$ with probability p^* and k^* with probability $1 - p^*$.*

Also, the minimum possible energy with delay x is $y = \left(\frac{c^*}{k^*} - \frac{p^*}{k^*(k^*+1)}\right) + c$, which is just the piecewise-linear interpolation of the function $f(x) = \frac{c^*}{2\lambda x+1} + c$ with vertices at $\{x = \frac{k-1}{2\lambda} : k \in \mathbb{Z}_{>0}\}$.

Proof. This follows simply from the fact that $\frac{\mathrm{TE}^{\mathrm{aff}}(k)}{k} = c + \frac{c^*}{k}$ for all $k > 0$, which is decreasing and convex – so all points in the sequence $(x_k, y_k) = \left(\frac{k-1}{2\lambda}, \frac{\mathrm{TE}(k)}{k}\right) = \left(\frac{k-1}{2\lambda}, c + \frac{c^*}{k}\right)$ are on the lower hull. The exact definition of k^* is derived from the platoon-mixing construction in Theorem 1's proof. ■

5 On Pareto-optimal Time-table Policies

We now analyze the important class of *time-table* policies. Recall that a time-table policy is a control policy that releases all currently present vehicles as a platoon at a set of predetermined and regular departure times, as described in Section 2. These policies are *open-loop* policies, in the sense that no information regarding the number or arrival times of the vehicles is required to realize them. In practice, there may be difficulties in obtaining state information and/or sharing it throughout the system, making time-table policies important to consider.

The average delay of policy π can also be expressed as the expected cumulative delay in one cycle divided by the expected number of arrivals:

$$\mathbb{E}[D(\pi)] = \frac{\frac{\lambda}{2}\sum_{m=1}^{M} T_m^2}{\lambda \sum_{m=1}^{M} T_m} = \frac{\sum_{m=1}^{M} T_m^2}{2\sum_{m=1}^{M} T_m}.$$

Note that the average delay of policy π is not a function of the arrival rate λ. Additionally, for any time-table policy $\pi \in \Pi$, there exists a time-table policy with only one interval size that yields the same average delay, namely, the policy $(T) \in \Pi$ satisfying $T = 2\mathbb{E}[D(\pi)]$. We call such a policy *fixed-interval.*

Similarly, we can express the average energy of a time-table policy by dividing the expected cumulative energy in a cycle by the expected number of arrivals:

$$\mathbb{E}[E(\pi)] = \frac{\sum_{m=1}^{M} \mathbb{E}[\mathrm{TE}(N_m)]}{\lambda \sum_{m=1}^{M} T_m},$$

where N_m is the number of arrivals in interval T_m, *i.e.*, N_m is distributed as a Poisson random variable with parameter λT_m.

We now look at the *fixed-interval* time-table policies, *i.e.*, $\mathcal{T} = \{kT_1 : k \in \mathbb{N}_{\geq 0}\}$, which releases the accumulated vehicles at the station every T_1 time units. To simplify notation, define $\epsilon(\lambda T)$ as the average accumulated energy in time interval T with arrival rate λ:

$$\epsilon(\lambda T) := \mathbb{E}[\mathrm{TE}(N)] = e^{-\lambda T} \sum_{k>0} \mathrm{TE}(k) \frac{(\lambda T)^k}{k!}, \tag{1}$$

where $N \sim Pois(\lambda T)$. Note that $\epsilon(0) = 0$, and $\epsilon(T) > 0$ for all $T > 0$. We assume that ϵ is infinitely differentiable on the open interval $(0, \infty)$. The function $\epsilon(x)$ is not necessarily infinitely differentiable in general, *e.g.*, $\mathrm{TE}(k) = k^k$; however, such TE are pathological and not meaningful to consider.[1] We can express the average energy of a fixed-interval policy $(T) \in \Pi$ as $\mathbb{E}[E(T)] = \frac{\epsilon(\lambda T)}{\lambda T}$.

Another important class of time-table policies is the class of *alternating time-table policies*, *i.e.*, $(T_1, T_2) \in \Pi$. Notice that the average energy of an alternating time-table policy $(T_1, T_2) \in \Pi$ can be expressed as $\mathbb{E}[E(T_1, T_2)] = \frac{\epsilon(\lambda T_1) + \epsilon(\lambda T_2)}{\lambda T_1 + \lambda T_2}$.

[1] We note that a sub-exponential, monotonically increasing TE guarantees that ϵ is infinitely differentiable. It is sufficient to check that $\left| k \frac{\mathrm{TE}(k-1)}{\mathrm{TE}(k)} \right|$ goes to ∞ in order to establish that ϵ is infinitely differentiable.

In Lemma 4 below, we characterize the class of TE functions for which fixed-interval policies are Pareto-optimal among all time-table policies. Consider a general time-table policy $\pi := (x_1, x_2, \ldots, x_M) \in \Pi$ and its corresponding fixed-interval policy $\pi_f := (x) \in \Pi$, where $x = 2\mathbb{E}[D(\pi)]$. Item 1 of Lemma 4 essentially states that the fixed-interval policy π_f never has larger average energy than the original time-table policy π, *i.e.*, $\mathbb{E}[E(\pi)] \geq \mathbb{E}[E(\pi_f)]$, *i.e.*, Pareto-optimality of fixed-interval policies among all *time-table* policies. Item 2 is similarly states that fixed-interval policies are Pareto-optimal among all *alternating* time-table policies, *i.e.*, $M = 2$. Item 3 gives us a key insight into the general structure of the Pareto-optimal energy-delay curve among time-table policies, since $\frac{\epsilon(x)}{x}$ is the energy-delay curve up to horizontal scaling.

Lemma 4. [2] *Given some* TE *function, suppose the corresponding* $\epsilon(x)$ *(see Equation* (1)*) is infinitely differentiable. Then, the following are equivalent:*

1. $\frac{\sum_m \epsilon(x_m)}{\sum_m x_m} \geq \frac{\epsilon\left(\frac{\sum_m x_m^2}{\sum_m x_m}\right)}{\frac{\sum_m x_m^2}{\sum_m x_m}}$, *for all* $x_m \geq 0$, *such that* $\sum_{m=1}^M x_m > 0$;
2. $\frac{\epsilon(x)+\epsilon(y)}{x+y} \geq \frac{\epsilon\left(\frac{x^2+y^2}{x+y}\right)}{\frac{x^2+y^2}{x+y}}$, *for all* $x, y \geq 0$, *such that* $x + y > 0$;
3. $\frac{\epsilon(x)}{x}$ *is convex on* $(0, \infty)$.

Note that the special case of Item 1 for $M = 2$ yields Item 2. Also, item 3 implies Item 1, via a simple application of Jensen's inequality. The main difficulty of this lemma is in showing Item 2 implies Item 3. Notice that Item 2 is similar to the notion of *midpoint convexity* of the function $\frac{\epsilon(x)}{x}$. The key insight behind the proof is the following. A particular TE function uniquely determines ϵ, from which one obtains the function $\epsilon(\lambda T)/(\lambda T)$ describing the (not necessarily optimal) energy-delay curve of all fixed-interval time-table policies. The previous lemma says that if the given TE function is such that the energy-delay curve of all fixed-interval time-table policies is convex, then fixed-interval time-table policies are Pareto-optimal among all other time-table policies.

Next, we present a more meaningful sufficient condition on the TE function to ensure the Pareto-optimality of fixed-interval time-table policies (among all time-table policies). This sufficient condition has the class of affine TE functions as a special case.

Theorem 5. *Let* TE *be given such that* ϵ *as defined in Equation* (1) *is infinitely differentiable. Furthermore, suppose* $\left\{\frac{\mathrm{TE}(k)}{k}\right\}_{k>0}$ *is a convex sequence. Then, fixed-interval time-table policies are Pareto-optimal among all time-table policies.*

Proof. We show that TE satisfies Item 3 of Lemma 4. Then, by Lemma 4, Item 3 implies Item 1; hence, Pareto-optimality of fixed-interval time-table policies follows directly. Since ϵ is infinitely differentiable, we prove convexity of ϵ by showing $(\epsilon(x)/x)'' \geq 0$. Expanding $(\epsilon(x)/x)''$ as a power series, one has

[2] For ease of presentation, we write $\sum_m$ in the place of $\sum_{m=1}^M$ in Item 1.

$$\left(\frac{\epsilon(x)}{x}\right)'' = e^{-x}\sum_{k\geq 0}\left(\frac{\mathrm{TE}(k+1)}{k+1} - 2\frac{\mathrm{TE}(k+2)}{k+2} + \frac{\mathrm{TE}(k+3)}{k+3}\right)\frac{x^k}{k!}.$$

Since $\mathrm{TE}(k)/k$ is convex, all the coefficients of the series are nonnegative, from which the series is never negative for any $x > 0$. ■

In particular, for an affine TE function, we have the following result:

Theorem 6. *Consider the affine total energy function,* $\mathrm{TE}^{\mathrm{aff}}(k) = ck + c^*\mathbf{1}(k > 0)$ *with* $c \geq 0$, $c^* > 0$. *Then, fixed-interval policies are Pareto-optimal among all time-table policies. Furthermore, the Pareto-optimal energy-delay curve is given by* $c + c^*\left(\frac{1-e^{-2\lambda D}}{2\lambda D}\right)$, *where* D *is the average delay and* λ *is the arrival rate.*

Proof. We begin by deriving the energy-delay curve for the fixed-interval time-table policies, and then verify its convexity. We compute $\frac{\epsilon(\lambda T)}{\lambda T} = c + c^*\frac{1-e^{-\lambda T}}{\lambda T}$. Note that $c \geq 0$, $c^* > 0$ guarantees convexity. By invoking Lemma 4, this energy-delay curve is Pareto-optimal among all time-table policies (equivalence of Item 3 and Item 1). Note that the energy-delay curve in terms of delay is given by tracing out the curve: $(\frac{T}{2}, \frac{\epsilon(\lambda T)}{\lambda T})$. By changing variables from time interval T to delay D, *i.e.*, $D = T/2$, we arrive at the optimal energy-delay curve.

Example 1. Consider Figure 1. The simplest nontrivial TE function is $\mathrm{TE}^{\mathrm{aff}}(k) = ck + c^*\mathbf{1}(k > 0)$ (given in Section 2) with $c \geq 0$ and $c^* > 0$. Note that the sufficient condition of Theorem 5 is satisfied; thus, fixed-interval policies are Pareto-optimal among all time-table policies.

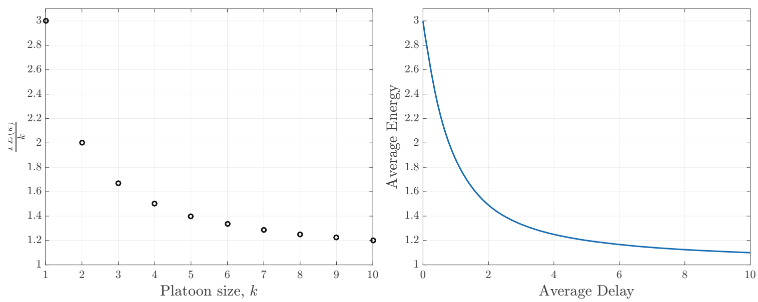

Fig. 1. Pareto-optimal (among time-table policies) energy-delay curve for affine TE function: $\mathrm{TE}(k) = 2 + k$. Note convexity of both the energy-delay curve (Lemma 4) and $\frac{\mathrm{TE}(k)}{k}$ (Theorem 5).

Example 2. Consider Figure 2. Note that Theorem 5 is not a necessary condition for optimality of fixed-interval time-table policies. Consider the following function $\mathrm{TE}(k) = 2\cdot\mathbf{1}(k > 0) + \lfloor k/2 \rfloor$. This TE function does not meet the convexity condition of Theorem 5, while its energy-delay curve is indeed Pareto-optimal among all time-table policies according to Lemma 4.

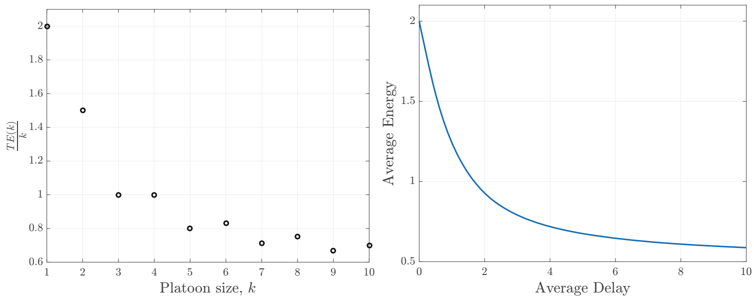

Fig. 2. Pareto-optimal (among time-table policies) energy-delay curve with non-convex $\frac{\mathrm{TE}(k)}{k}$; $\mathrm{TE}(k) = 2 + \lfloor k/2 \rfloor$. Note convexity of energy-delay curve (Lemma 4).

Example 3. Now, we show that if one violates Items 1, 2, or 3 of the Lemma 4, then a fixed-interval time-table policy can be arbitrarily worse than some alternating time-table policy. Consider the almost affine TE function: $\mathrm{TE}(k) = k+2$, for all $k \neq 4$, and $\mathrm{TE}(4) = K$. Consider the alternating policy $(1, 5)$, and its delay-equivalent fixed-interval policy $(\frac{13}{3})$. Then, the difference in average energy between these two policies grows linearly with K, with the alternating policy having lower average energy.

6 Discussion and Conclusion

In this section, we compare time-table and threshold policies. We assume that vehicles arrive at the station according to a Poisson arrival, and we assume an affine total energy function. See Figure 3. In the plot on the left, we show the

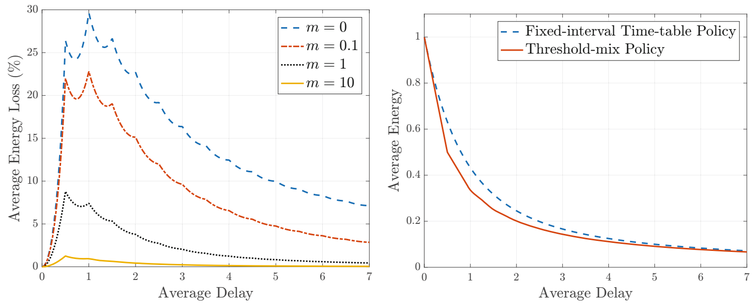

Fig. 3. On the left, average energy loss versus average delay plot for various total energy curves, $\mathrm{TE}(k) = mk + 1$. On the right, Pareto-optimal energy-delay tradeoff for the class of time-table (fixed-interval) and feedback policies (mixed-threshold) for $\mathrm{TE}(k) = 1$; Poisson arrival process, intensity $\lambda = 1$.

loss in average energy by implementing a time-table policy in lieu of a feedback policy, for various energy functions. We only consider energy functions of the form $\mathrm{TE}(k) = mk + 1$, since the plot of average energy loss is the same for any other total energy function $\mathrm{TE}(k) = cmk + c$, for any $c > 0$. We note that the energy loss is no greater than thirty percent, and corresponds to a flat total energy function, $m = 0$, *i.e.*, the total energy of a platoon is the same regardless of how many vehicles are in the platoon. Also, we fixed the intensity of Poisson arrivals λ to 1, since changing the arrival rate λ simply corresponds to a horizontal scaling of λ in the plots of Figure 3.

One interesting insight is that hard-threshold policies – the building blocks of the Pareto-optimal mixed-threshold policies – derive much of their advantage over fixed-interval time-table policies by getting (on average) one more vehicle into each platoon. Specifically, the hard-threshold policy $\bar{\pi}_k$ gets k-vehicle platoons for an average delay of $\frac{k-1}{2\lambda}$ (as shown in Lemma 3); a time-table policy with an average delay of $\frac{k-1}{2\lambda}$ would require an interval of length $\frac{k-1}{\lambda}$, which gives it an average platoon size of $k-1$.

Another interesting point is the 'spikiness' of the curves on the left of Figure 3. These curves plot the relative energy usage of timetable and mixed-threshold policies. The spikes occur at average delays which match up exactly to $\frac{k-1}{2\lambda}$ for some k – at these points, the Pareto-optimal policies are hard-threshold. This happens because mixing two hard-threshold policies is in some sense a compromise; ideally, one would like to have a hard-threshold policy for non-integer sized platoons. However, of course, that is impossible, and mixing hard-threshold policies is thus a stand-in for this. On the other hand, fixed-interval timetable policies smoothly increase their intervals, and hence don't have to make this compromise.

In this paper, we considered a class of decision-making problems which model a number of logistical scenarios – including platooning, drafting, vortex riding, ride sharing, and many others – that are relevant in transportation systems involving autonomous vehicles. In our problem, vehicles arrive at a station following a Poisson process; the station then decides when to send the vehicles currently present to their destination. Sending more vehicles at once allows for more energy-efficient transportation, but to do so consistently the station must delay some vehicles to wait for others to arrive. We explored the energy-delay tradeoff this problem presents, focusing on open loop (time-table) and feedback policies, and we characterized both the Pareto-optimal policies and the optimal energy-delay curves for each case. We also proved some general results concerning the form that the optimal policies can take, and showed that the performance of any policy is uniquely determined by the sizes of the platoons it produces.

Future work will include studies involving multiple stations, complex transportation networks, and unknown statistics, among other generalizations. In this setting, we hope to identify the tradeoffs between efficiency metrics, *e.g.*, capacity and delay, and sustainability metrics, *e.g.*, energy consumption and environmental impact.

Acknowledgements This work was supported by NSF under grant #1544413.

References

1. J. B. Greenblatt and S. Saxena. Autonomous taxis could greatly reduce greenhouse-gas emissions of US light-duty vehicles. *Nature Climate Change*, 5(9):860–863, July 2015.
2. K. B. Sandvik and K. Lohne. The Rise of the Humanitarian Drone: Giving Content to an Emerging Concept. *Millennium - Journal of International Studies*, 43(1):145–164, September 2014.
3. R. D'Andrea. Guest Editorial Can Drones Deliver? *IEEE Transactions on Automation Science and Engineering*, 11(3):647–648, July 2014.
4. B. Besselink, V. Turri, S. H. van de Hoef, K.-Y. Liang, A. Alam, J. Martensson, and K. H Johansson. Cyber–Physical Control of Road Freight Transport. *Proceedings of the IEEE*, 104(5):1128–1141, April 2016.
5. S. H. van de Hoef, K. H. Johansson, and D. V. Dimarogonas. Fuel-Optimal Centralized Coordination of Truck Platooning Based on Shortest Paths. In *American Control Conference*, July 2015.
6. U. Franke, F. Bottiger, Z. Zomotor, and D. Seeberger. Truck Platooning in Mixed Traffic. In *IEEE Intelligent Vehicles Symposium*, 1995.
7. S. Tsugawa, S. Kato, and K. Aoki. An Automated Truck Platoon for Energy Saving. In *IEEE/RSJ IROS Conference*, September 2011.
8. Overview of Platooning Systems. In *ITS World Congress*, October 2012.
9. O. Gehring and H. Fritz. Practical results of a longitudinal control concept for truck platooning with vehicle to vehicle communication. In *IEEE Conference on Intelligent Transportation Systems*, November 2003.
10. W. B. Blake, S. R. Bieniawski, and T. C. Flanzer. Surfing aircraft vortices for energy. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, 12(1):31–39, December 2014.
11. M. J. Vachon, R. J. Ray, K. R. Walsh, and K. Ennix. F/A-18 Performance Benefits Measured During Autonomous Formation Flight Project. Technical report, 2003.
12. H. Weimershirsh, J. Martin, Y. Clerquin, P. Alexandre, and S. Jiraskova. Energy saving in flight formation. *Nature*, pages 697–698, October 2001.
13. M. T. Florian and M. H. Dickinson. Fly with a little flap from your friends. *Nature*, pages 295–296, January 2014.
14. J. Brisswalter and C. Hausswirth. Consequences of Drafting on Human Locomotion: Bene ts on Sports Performance. *International Journal of Sports Physiology and Performance*, 3:3–15, 2008.
15. D. Gross and C. Harris. *Fundamentals of Queueing Theory*. Wiley, 1998.
16. R. C. Larson and A. R. Odoni. *Urban Operations Research*. MIT Press, 1999.
17. N. A. H. Agatz, A. L. Erera, M. W. P. Savelsbergh, and X. Wang. Optimization for dynamic ride-sharing: A review. *European Journal of Operational Research*, 223(2):295–303, December 2012.
18. D. J. Fagnant and K. M. Kockelman. Dynamic ride-sharing and fleet sizing for a system of shared autonomous vehicles in Austin, Texas. In *Transportation Research Board Annual Meeting*, September 2015.
19. Z. Tian, T. Urbanik, and R. Engelbrecht. Variations in capacity and delay estimates from microscopic traffic simulation models. *Transportation Research Record: Journal of the Transportation Research Board*, 1802:23–31, 2002.
20. A. Adler, D. Miculescu, and S. Karaman. Optimal policies for platooning and ride sharing in autonomy-enabled transportation. http://peris.mit.edu/AdlerMiculescuKaramanWAFR16.pdf, July 2016.