Title
City-wide traffic control: modeling impacts of cordon queues

Authors
Ni, Wei
Cassidy, Michael J

Publication Date
2018-03-01

INSTITUTE OF TRANSPORTATION STUDIES
UNIVERSITY OF CALITFORNIA, BERKELEY

City-wide traffic control: modeling impacts of cordon queues

Wei Ni
Michael J. Cassidy

Pre-print Series
UCB-ITS-PP-2018-01

The Institute of Transportation Studies at the University of California, Berkeley, is one of the world's leading centers for transportation research, education, and scholarship.

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. UCB-ITS-PP-2018-01	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle City-wide traffic control: modeling impacts of cordon queues		**5. Report Date** March 2018 **6. Performing Organization Code**
7. Author(s) Wei Ni Michael J. Cassidy		**8. Performing Organization Report No.** Click to enter text. Enter any/all unique alphanumeric report numbers assigned by the performing organization, if applicable.
9. Performing Organization Name and Address Institute of Transportation Studies University of California, Berkeley 109 McLaughlin Hall, MC1720 Berkeley, CA 94720		**10. Work Unit No.** **11. Contract or Grant No.**
12. Sponsoring Agency Name and Address UC Institute of Transportation Studies		**13. Type of Report and Period Covered** Pre-print **14. Sponsoring Agency Code**
15. Supplementary Notes		

16. Abstract

Optimal cordon-metering rates are obtained using Macroscopic Fundamental Diagrams in combination with flow conservation laws. A model-predictive control algorithm is also used so that time-varying metering rates are generated based on their forecasted impacts. Our scalable algorithm can do this for an arbitrary number of cordoned neighborhoods within a city. Unlike its predecessors, the proposed model accounts for the constraining effects that cordon queues impose on a neighborhood's circulating traffic. It does so at every time step by approximating a neighborhood's street space occupied by cordon queues, and re-scaling the MFD downward to describe the state of circulating traffic that results. The model is also unique in that it differentiates between saturated and under-saturated cordon-metering operations. Computer simulations show that these enhancements can substantially improve the predictions of both, the trip completion rates in a neighborhood and the rates that vehicles cross metered cordons. Optimal metering policies generated as a result are similarly shown to do a better job in reducing the Vehicle Hours Traveled in a city. The VHT reductions stemming from the proposed model and from its predecessors differed by as much as 18%.

17. Key Words Traffic control, Traffic models, Algorithms, Urban transportation	**18. Distribution Statement** No restrictions.

19. Security Classif. (of this report) Unclassified.	**20. Security Classif. (of this page)** Unclassified.	**21. No. of Pages** 17	**22. Price** 0

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

City-wide traffic control: modeling impacts of cordon queues

Wei Ni*[1] and Michael J. Cassidy†[1]

[1] *Department of Civil and Environmental Engineering, University of California, Berkeley*

Abstract

Optimal cordon-metering rates are obtained using Macroscopic Fundamental Diagrams in combination with flow conservation laws. A model-predictive control algorithm is also used so that time-varying metering rates are generated based on their forecasted impacts. Our scalable algorithm can do this for an arbitrary number of cordoned neighborhoods within a city. Unlike its predecessors, the proposed model accounts for the constraining effects that cordon queues impose on a neighborhood's circulating traffic. It does so at every time step by approximating a neighborhood's street space occupied by cordon queues, and re-scaling the MFD downward to describe the state of circulating traffic that results. The model is also unique in that it differentiates between saturated and under-saturated cordon-metering operations. Computer simulations show that these enhancements can substantially improve the predictions of both, the trip completion rates in a neighborhood and the rates that vehicles cross metered cordons. Optimal metering policies generated as a result are similarly shown to do a better job in reducing the Vehicle Hours Traveled in a city. The VHT reductions stemming from the proposed model and from its predecessors differed by as much as 18%.

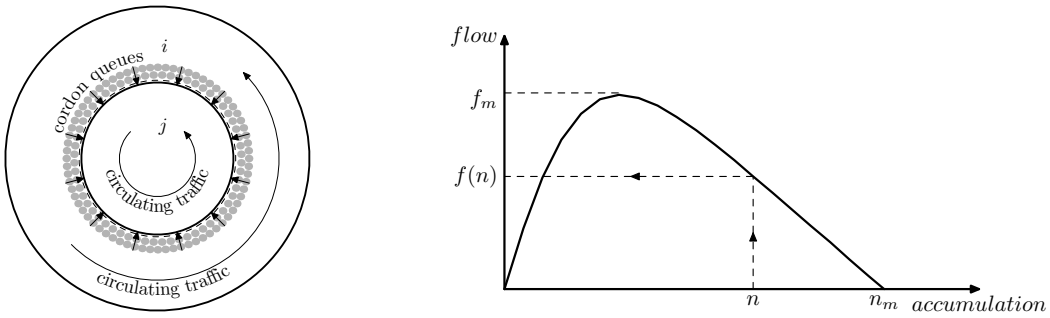
1 Introduction

A sizable literature exists on the re-timing of traffic signals to meter inflows to cordoned neighborhoods; eg, see [1–6] . In some of those efforts, metering rates were optimized using Macroscopic Fundamental Diagrams (MFDs) in combination with flow conservation laws [4–6]. The works have produced what we will call Neighborhood Transmission Models.

These NTMs separately estimate a neighborhood's time-varying accumulations of vehicles that are bound for destinations residing in that same neighborhood, and in each of the other neighborhoods within a city. The original NTM in [4, 5] has a hidden problem regarding trips of the latter (i.e inter-neighborhood) kind. When cordon metering substantially delays vehicles from crossing from one neighborhood to another, those vehicles are returned to their neighborhood's circulating traffic for another go at the boundary. They start from scratch as if previous attempts at boundary crossings had not just occurred.

*weini@berkeley.edu

†cassidy@ce.berkeley.edu



(a) Cordon queues and circulating traffic in two neighborhoods i and j

(b) Neighborhood MFD

Figure 1: Cordon queues and MFD

Recent work in [6] has addressed this deficiency by differentiating a neighborhood's accumulations of circulating vehicles from those queued at its metered cordons; see the example in Fig.1(a). Though a notable advancement, this more discriminating NTM ignores constraining effects that a neighborhood's cordon queues can impose on its circulating vehicles; i.e. the model does not recognize that the street space occupied by those (possibly lengthy) queues can diminish the neighborhood's capacity to serve the rest of its traffic. Of further concern, the model in [6] assumes that metered green times are always saturated by inflows; i.e. it ignores the possibility of under-saturated metering owing to shorter-lengthed cordon queues. The solution in [6], moreover, furnished optimal cordon-metering rates for two neighborhoods only. Finally, the discriminating NTM was tested only for internal consistency; i.e. its analytical solutions were compared against numerical ones generated by the very same (macro-level) model.

The present paper offers enhancements. The most important of these captures the impact of cordon queues on network capacity by adaptively re-scaling a neighborhood's MFD. Another enhancement enables the modeling of both saturated and under-saturated cordon-metering operations. Moreover, we use a scalable, model-predictive control algorithm that can optimize metering rates for an arbitrary number of cordoned neighborhoods. Finally, the proposed NTM and its predecessors are tested using the AIMSUN micro-level traffic simulation model [7]. It furnished the closest approximations to ground truth available to us.

The simulations show that our proposed NTM better predicts the rates that vehicles both, cross metered cordons and complete their trips as compared against predictions from the earlier models in [4–6]. These improved predictions produce, in turn, optimal cordon-metering rates that do a better job of reducing the vehicle-hours traveled in a city. Improvements can be sizable, as will be shown in due course.

Simulations also show that inputs needed by the proposed NTM can be collected from the information systems onboard connected vehicles, even when the penetration rate of those special vehicles is small. As a final aside, simulations show that optimal metering rates tend to alternate over time between one that is highly relaxed, and another that is highly restrictive. This finding is important because it lends credibility to a key assumption in our model, as will be explained.

2 Background

The logic behind the NTMs in [4–6] are now reviewed and critiqued in greater depth. For each model in that family, the vehicle accumulation in a neighborhood i at time t, $n_i(t)$, is assumed to be a measurable input. It is used with an MFD to obtain i's total flow, $f_i(n_i)$, as illustrated in Fig.1(b). Recall from sec.1 that the $n_i(t)$ is composed of i's accumulations of vehicles bound for neighborhood j, $n_{ij}(t)$, and those bound for destinations within i itself, $n_{ii}(t)$. How one might disentangle these more disaggregate accumulations from the $n_i(t)$ is not clear from the literature. We will assume that in the near future, the information available from connected vehicles in the traffic mix can be used to sample the $\frac{n_{ij}(t)}{n_i(t)}$ and $\frac{n_{ii}(t)}{n_i(t)}$, and to estimate other inputs described in due course.

2.1 Original Model

To see how the earliest NTM in [4, 5] incorporated flow conservation laws, denote L_{ij} as the average distance that vehicles in i must travel to reach i's boundary with j.[1] Thus, $\frac{f_i(n_i(t))}{L_{ij}} \cdot \frac{n_{ij}(t)}{n_i(t)}$ is the rate that i's vehicles bound for j reach that boundary. The product of that rate and the fraction that can cross the boundary when it is metered, $u_{ij}(t) \in [u_{min}, u_{max}]$, gives the rate that j-bound vehicles exit i for $j \neq i$.[2] Subtracting this boundary-crossing rate from $\lambda_{ij}(t)$, the inter-neighborhood demand generated in i, enables the updating of these n_{ij} at each time step of duration τ; i.e.

$$n_{ij}(t+\tau) - n_{ij}(t) = \tau \cdot [-\frac{f_i(n_i(t))}{L_{ij}} \cdot \frac{n_{ij}(t)}{n_i(t)} \cdot u_{ij}(t) + \lambda_{ij}(t)]. \tag{2.1}$$

The n_{ii} are similarly updated by considering the intra-neighborhood demand generated in i, $\lambda_{ii}(t)$, the trip completion rate in i, and the inflows that were allowed to cross into i from j; i.e.

$$n_{ii}(t+\tau) - n_{ii}(t) = \tau \cdot [-\frac{f_i(n_i(t))}{L_{ii}} \cdot \frac{n_{ii}(t)}{n_i(t)} + \lambda_{ii}(t) + \sum_{j \neq i} \frac{f_i(n_j)}{L_{ji}} \cdot \frac{n_{ji}(t)}{n_j(t)} \cdot u_{ji}(t)], \tag{2.2}$$

where L_{ii} is the average distance that intra-neighborhood vehicles travel in i.

The problem arises when metering delays vehicles that are queued and ready to cross the cordon beyond duration τ. Note that (2.1) and (2.2) have no mechanism to carry-over those queued vehicles into the next time step. They return instead to the n_{ij}, and are then assumed to be confronted (again!) by the distance L_{ij} that separates them from the cordon that they had just failed to cross.

These artificial additions to a neighborhood's circulating traffic can constrain its trip completions. And sending back queued vehicles can artificially starve a cordon line of flow if the metering there is suddenly relaxed.

2.2 More Discriminating Model

As noted in sec.1, the problem was addressed in [6] by decomposing the $n_{ij}(t)$ into its accumulations of circulating vehicles, $n_{c,ij}(t)$, and of vehicles queued at its cordon, $n_{q,ij}(t)$, such that $n_{ij}(t) =$

[1] The L_{ij} can be estimated knowing the physical size of neighborhood i.

[2] One might reasonably designate u_{max} to be 1, to describe the flow that can enter the cordon when the traffic signals are not re-timed to function as meters. One might impose a lower bound u_{min} to be greater than zero, to avoid the ire of j-bound drivers.

$n_{c,ij}(t) + n_{q,ij}(t)$. For convenience, denote i's total accumulations of circulating and cordon-queued vehicles as $n_{c,i}(t) = \sum_j n_{c,ij}(t)$ and $n_{q,i}(t) = \sum_j n_{q,ij}(t)$, respectively. This more discriminating NTM can therefore be formulated as

$$n_{c,ii}(t+\tau) - n_{c,ii}(t) = \tau \cdot [-\frac{f_i(n_{c,i}(t))}{L_{ii}} \cdot \frac{n_{c,ii}(t)}{n_{c,i}(t)} + \lambda_{ii}(t) + \sum_j C_{ji} \cdot u_{ji}(t)] \quad (2.3)$$

$$n_{c,ij}(t+\tau) - n_{c,ij}(t) = \tau \cdot [-\frac{f_i(n_{c,i}(t))}{L_{ij}} \cdot \frac{n_{c,ij}(t)}{n_{c,i}(t)} + \lambda_{ij}(t)] \quad (2.4)$$

$$n_{q,ij}(t+\tau) - n_{q,ij}(t) = \tau \cdot [-C_{ij} \cdot u_{ij}(t) + \frac{f_i(n_{c,i}(t))}{L_{ij}} \cdot \frac{n_{c,ij}(t)}{n_{c,i}(t)}], \quad (2.5)$$

where C denotes the unmetered capacity at the cordon separating one neighborhood from another.

Equation (2.5) means that a cordon's residual queues are carried-over from on time step to the next. Moreover, the inclusion of C in (2.3) and (2.5) remedies any underestimates of cordon-crossing rates that might otherwise occur should metering abruptly relax.

New concerns emerge, however. Note from (2.3) and (2.4) how vehicle flow in i, $f_i(n_{c,i}(t))$, is solely a function of its circulating accumulation. The constraining effect that cordon queues can impose on circulating flows is not considered. Further note how the inclusion of C in (2.3) and (2.5) assumes that metered green times are always saturated by inflows.

Of further interest, the solution furnished in [6] was an analytical one. Though an impressive achievement, the approach limited the NTM's application to scenarios involving only two neighborhoods in a city. As noted in sec.1, moreover, the analytical solutions were tested only against numerical ones from that same (more discriminating) model. Those tests showed the analytical solutions to be consistent with numerical ones, but say nothing about the model's physical realism.

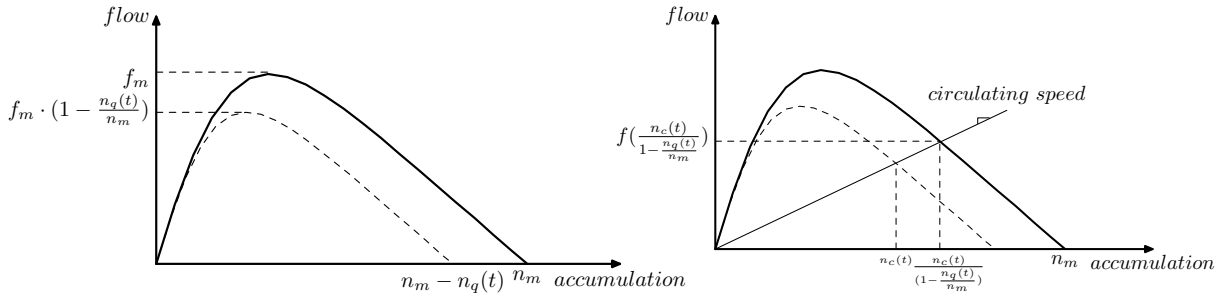
3 Proposed NTM

The model now proposed recognizes that a neighborhood's circulating flows are functions of its accumulations of both, the circulating and the cordon-queued vehicles, n_c and n_q. We model the influence of n_q in a simple way as will be explained momentarily. To support our simple approach, we will show that an optimal cordon-control scheme is of bang-singular type, but entails bang-bang control almost everywhere in the system state space. This means that optimal metering along a cordon tends to alternate over time between fully relaxed and highly restrictive rates.[3]

The implication is that cordon queues that grow long enough to constrain circulating traffic will tend to be dense. This should be the case whether bang-bang control switches frequently or infrequently between extreme metering rates.

Consider, for example, the cordon queues that persist during prolonged periods of very restrictive metering. Those queues will be dense in their own right, and can only grow denser by expanding and coalescing with queues created by other traffic signals nearby. If the queues that formed during extended periods of fully-relaxed metering were to grow long, their coalescence with other queues could make them dense as well. Or, if control switches frequently between relaxed and restrictive rates, recovery waves can be halted if and when they arrive at upstream signals during their red phases. Queues on the upstream links would remain jammed for the remainders of those reds, and possibly beyond if turning vehicles create queues downstream; see [8].

[3]Theoretical derivations are presented in Appendix A, and experimental support is furnished in sec.5.



(a) Re-scaling an MFD (b) Circulating speed and flow

Figure 2: Re-scaled MFD and implication on circulating speed

We exaggerate the above considerations and assume that cordon queues exhibit jam densities at all times. The assumption enables a rescaling downward of a neighborhood's MFD. The re-scaled version approximates the neighborhood's states of circulating traffic (only) that result from the street space lost in storing cordon queues. The approximations are clearly based on lower-bound estimates of the lost street space: vehicles in less-than-jammed traffic exhibit larger-than-jammed spacings, and thus collectively occupy greater space. Still, our lower-bound estimates will capture at least some of the constraining effects that cordon queues impose on neighborhood circulation, and are improvements over ignoring the efforts entirely as in [4–6]. The point will be illustrated in sec.5. Details of the MFD re-scaling are furnished below.

3.1 Modeling Cordon-Queue Impacts

With the assumption of jammed cordon queues, the maximum accumulation of circulating vehicles that a neighborhood can store becomes $n_m - n_q$; i.e. the neighborhood's vehicle-storage capacity, n_m, is diminished by the factor $(1 - \frac{n_q}{n_m})$, as shown in Fig.2(a). That same multiplicative factor applied to f_m, the neighborhood's maximum possible flow, gives the capacity available for circulating flows. The assumption here is only that the neighborhood's capacity to circulate traffic diminishes in proportion to its street space given to cordon queues.

Capturing the constraining effect that those queues exert on circulating speeds would require adjustments using the same factor. Hence, the neighborhood's average circulating speed is given by $f(\frac{n_c}{1-\frac{n_q}{n_m}}) \cdot (\frac{n_c}{1-\frac{n_q}{n_m}})^{-1}$, as exemplified by the slope of the chord in Fig.2(b). The product of this speed and n_c is the neighborhood's circulating flow. That flow for neighborhood i is therefore

$$\bar{f}_i(n_c, n_q) = f(\frac{n_c}{1 - \frac{n_q}{n_m}}) \cdot (1 - \frac{n_q}{n_m}), \tag{3.1}$$

as given by a re-scaled MFD, like the dashed ones in Figs.2(a) and 2(b).

Recall that (3.1) represents a departure from the logic behind the original NTM [4, 5], which assumes that circulating flows $f(n) = f(n_q + n_c)$, as if vehicles in the n_q were themselves circulating. The proposed logic is likewise different from that of the discriminating NTM [6], which ignores the influence of n_q by assuming $f(n) = f(n_c)$. The proposed re-scaling can, of course, occur in every time step.

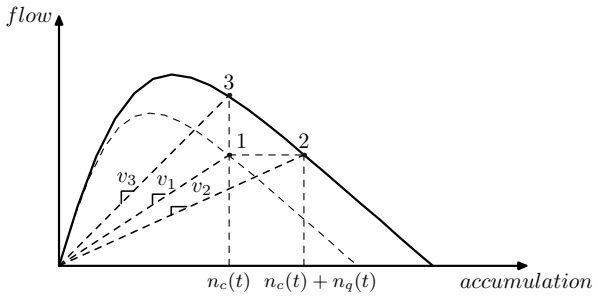


Figure 3: Circulating speeds predicted by proposed NTM and predecessors

3.2 Saturated vs Under-Saturated Metering and Model Formulation

Denote as $d_{ij}(t)$ the number of vehicles to cross a cordon from i to j during time window $[t, t+\tau]$. Unlike [6], we do not assume that $d_{ij}(t)$ always equals $C_{ij} \cdot u_{ij}(t) \cdot \tau$, as if the meters were always saturated. When cordon queues are sufficiently short, $d_{ij}(t)$ is instead the sum of the $n_{q,ij}(t)$, the cordon's queued vehicles present at the start of the time window, and the inter-neighborhood vehicles to arrive at the cordon during duration τ. Hence,

$$d_{ij}(t) \quad = \quad min\{C_{ij} \cdot u_{ij}(t) \cdot \tau, n_{q,ij}(t) + \tau \cdot \frac{\bar{f}_i(n_{c,i}(t), n_{q,i}(t))}{L_{ij}} \cdot \frac{n_{c,ij}(t)}{n_{c,i}(t)}\}. \tag{3.2}$$

The formulation for the proposed NTM is thus

$$n_{c,ii}(t+\tau) - n_{c,ii}(t) = \tau \cdot [-\frac{\bar{f}_i(n_{c,i}(t), n_{q,i}(t))}{L_{ii}} \cdot \frac{n_{c,ii}(t)}{n_{c,i}(t)} + \lambda_{ii}(t)] + \sum_j d_{ji}(t) \tag{3.3}$$

$$n_{c,ij}(t+\tau) - n_{c,ij}(t) = \tau \cdot [-\frac{\bar{f}_i(n_{c,i}(t), n_{q,i}(t))}{L_{ij}} \cdot \frac{n_{c,ij}(t)}{n_{c,i}(t)} + \lambda_{ij}(t)] \tag{3.4}$$

$$n_{q,ij}(t+\tau) - n_{q,ij}(t) = -d_{ij}(t) + \tau \cdot \frac{\bar{f}_i(n_{c,i}(t), n_{q,i}(t))}{L_{ij}} \cdot \frac{n_{c,ij}(t)}{n_{c,i}(t)}. \tag{3.5}$$

3.3 Theoretical Insights

Comparisons between the proposed NTM and its predecessors support earlier remarks regarding model predictions of trip-completion rate. We start by comparing the proposed NTM with the original in [4, 5]. To that end, we note that

$$\frac{\bar{f}_i(n_{c,i}(t), n_{q,i}(t))}{n_{c,i}(t)} \geq \frac{f_i(n_{c,i}(t) + n_{q,i}(t))}{n_{c,i}(t) + n_{q,i}(t)}, \tag{3.6}$$

which indicates that an average circulating speed predicted for a certain scenario by the proposed model tends to exceed that predicted by the original one. Example speeds labeled v_1 and v_2 in Fig.3 illustrate a case in point.

General verification of (3.6) comes by expanding its left-side as

$$\frac{\bar{f}_i(n_{c,i}(t), n_{q,i}(t))}{n_{c,i}(t)} = f(\frac{n_{c,i}(t)}{1 - n_{q,i}(t)/n_m}) \cdot \frac{1 - n_{q,i}(t)/n_m}{n_{c,i}(t)}, \tag{3.7}$$

so that the expansion has the form $g(x) = f(x)/x$. The function $g(x)$ is non-increasing because

$$g^{'}(x) = \frac{f^{'}(x) \cdot x - f(x)}{x^2} = \frac{1}{x^3}(f^{'}(x) - \frac{f(x) - f(0)}{x - 0}) \leq 0, \tag{3.8}$$

where the inequality in (3.8) holds because f is concave and $f(0) = 0$. With function g, (3.6) becomes

$$g(\frac{n_{c,i}(t)}{1 - n_{q,i}(t)/n_m}) \geq g(n_{c,i}(t) + n_{q,i}(t)). \tag{3.9}$$

As regards the left-side of (3.9), it can be shown that

$$\frac{n_{c,i}(t)}{1 - n_{q,i}(t)/n_m} = n_{c,i}(t) + n_{q,i}(t) \cdot \frac{n_{c,i}(t)}{n_m - n_{q,i}(t)}, \tag{3.10}$$

which cannot exceed the right-side argument of (3.9), $n_{c,i}(t) + n_{q,i}(t)$. Hence, (3.9) holds knowing that g is non-increasing. Equation (3.6) therefore holds as well.

Having thus verified (3.6), we can expand both sides of that inequality as

$$\frac{\bar{f}_i(n_{c,i}(t), n_{q,i}(t))}{L_{ii}} \cdot \frac{n_{c,ii}(t)}{n_{c,i}(t)} \geq \frac{f_i(n_{c,i}(t) + n_{q,i}(t))}{L_{ii}} \cdot \frac{n_{c,ii}(t)}{n_{c,i}(t) + n_{q,i}(t)}, \tag{3.11}$$

which indicates that trip-completion rates predicted by the proposed NTM tend to exceed those predicted by the original model. As noted, the latter's artificial inclusion of $n_{q,i}(t)$ in circulating traffic tends to damp trip completions.

Comparing predictions between the proposed NTM and the discriminating one in [6] tells a different story. To that end, we note that

$$\frac{\bar{f}_i(n_{c,i}(t), n_{q,i}(t))}{f_i(n_{c,i}(t))} \leq 1, \tag{3.12}$$

as is clear from the example data points labeled "1" and "3" in Fig.3. Thus,

$$\frac{\bar{f}_i(n_{c,i}(t), n_{q,i}(t))}{L_{ii}} \cdot \frac{n_{c,ii}(t)}{n_{c,i}(t)} \leq \frac{f_i(n_{c,i}(t))}{L_{ii}} \cdot \frac{n_{c,ii}(t)}{n_{c,i}(t)}, \tag{3.13}$$

which indicates that trip-completion rates predicted by the proposed NTM tend to be smaller than those predicted by the discriminating model in [6]. As noted, the latter's failure to account for constraining effects of cordon queues can result in overly-optimistic predictions of trip completion.

We have also noted how the original model [4, 5] and its more discriminating counterpart [6] may in some circumstances under- and over-estimate cordon-crossing rates, respectively. Experimental evidence in support of all these theoretical insights is furnished in sec.5.

4 Control

As in [4–6], we seek $u^{\star}_{ij}(t)$, the optimal metering action to manage cordon crossings from i to j so as to minimize vehicle-hours traveled (VHT) on a network. And like those previous works, we do so assuming that trip-making demand, $\lambda_{ij}(t)$, is given. We discretize the entire control period into

small time steps of duration τ, and obtain numerical solutions using model-predictive control with a rolling horizon of H steps. Denoting the start of each planning horizon as $t = t_0$, our objective function for each horizon takes the form

$$min_u \sum_{1 \le h \le H} \sum_{i,j} n_{q,ij}(t_0 + h \cdot \tau) + n_{c,ij}(t_0 + h \cdot \tau), \tag{4.1}$$

where system dynamics are constrained by (3.3)-(3.5), and accumulations at t_0 are estimated as described below.

4.1 Initial System States

Denote as α the penetration of connected vehicles in the traffic mix, since their onboard systems can provide needed inputs. At the start of every planning horizon,

$$n_{c,ij}(t_0) + n_{q,ij}(t_0) = \frac{K}{\alpha} \tag{4.2}$$

if at that t_0 there are K connected vehicles in i observed heading for j. Assume for simplicity that $n_{q,ij}(t_0)$ need travel no distance to reach i's boundary with j. It is easy to show that

$$n_{c,ij}(t_0) = \frac{\sum_k l_k(t_0)}{\alpha \cdot L_{ij}} \tag{4.3}$$

$$n_{q,ij}(t_0) = \frac{K}{\alpha} - \frac{\sum_k l_k(t_0)}{\alpha \cdot L_{ij}}, \tag{4.4}$$

where $l_k(t_0)$ is the distance that connected vehicle k must travel at time t_0 to reach the boundary, as measured by that vehicle.

4.2 Numerical Solution

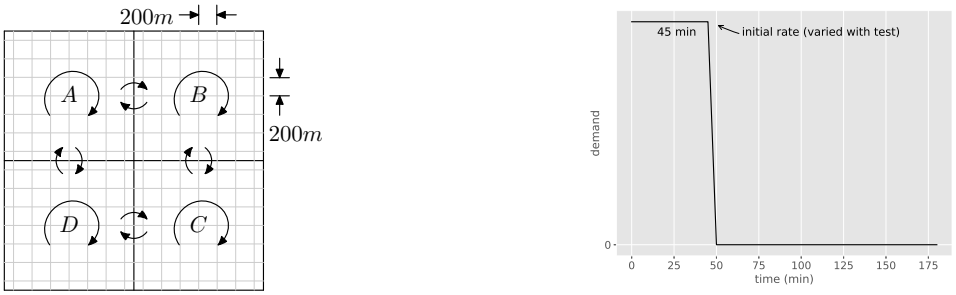
The dimensions of an NTM system increase quadratically with the number of cordoned neighborhoods. Analytical solutions for an arbitrary number of neighborhoods therefore seem out of reach.[4]

We turned to a numerical method instead. Rather than directly solving the non-linear problem using a single-shooting method as in [4] and [5], we sought a more efficient approach that is scalable with problem size.[5] Hence we chose the iterative Linear Quadratic Regulator (iLQR) method [9], along with several improvements to the regularization and line-search aspects of the algorithm developed in [10]. Details of this algorithm and brief discussion of why it suits our problem so well are furnished in Appendix B.

[4] Recall that the analytical solution furnished in [6] was for two neighborhoods (only).

[5] The single-shooting method was also used in [6] to verify the analytical solutions furnished there.

5 Numerical Analysis



(a) Network with 4 cordoned neighborhoods (b) Time-varying pattern of each O-D demand

Figure 4: Test inputs

Tests of the proposed NTM and its predecessors entailed use of the AIMSUN software [7] to simulate traffic on a signalized network of 14x14 two-way streets. Each street segment was 200m long, with 2 lanes in each travel direction. The network was cordoned into four equally-sized neighborhoods labeled A-D in Fig.4(a).

Demands for the O-Ds (shown with arrows in Fig.4(a)) each took the time-varying pattern shown in Fig.4(b).[6] Rates varied across tests as will be described in due course, but were always set so that all trips were completed within each test's 3-h period. All (simulated) drivers received traffic updates at 5-min intervals, and responded by altering their routes as per the model in [7]. The penetration of connected vehicles, α, was set at just 5%.

5.1 Predictive Strengths

The following illustrates how the proposed NTM can produce improved estimates of neighborhood traffic conditions. Initial demand for each O-D was set at 8,000 vph. Cordon queues were created in the early going by restrictively metering the collective crossings from neighborhood A to B at 1,500 vph. The metering occurred from minute 5 to 35 of the simulated test. No other cordon-control actions were taken.

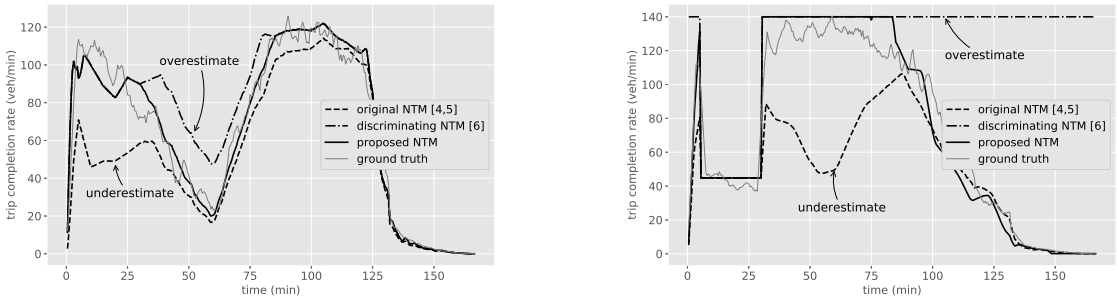
Figure 5(a) presents trip-completion rates in A as predicted by: (i) the original NTM in [4, 5]; (ii) the more discriminating NTM in [6]; and (iii) the proposed model. "Ground-truth" rates generated by AIMSUN are shown as well. Visual inspection of Fig.5(a) clearly shows that predictions from the proposed NTM fit the ground-truth data best.[7]

In contrast, the original NTM [4, 5] underestimated rates in the early going, when A's cordon queues were lengthiest. The queued vehicles that were sent back into A's circulating traffic seem indeed to have artificially constrained trip completions. The discriminating NTM [6] overestimated rates toward the middle of the test, which is when circulating traffic entered the congested regime.

[6]Each trip crossed no more than one cordon line to avoid the complications of modeling the driver route-choice behavior involved in touring three or more neighborhoods.

[7]Comparing predictions from the proposed model against ground truth produces a root-mean-square error of 7.7 vehicles/min. This is substantially smaller than the RMSEs of 20.5 vehicles/min and 14.7 vehicles/min produced by the original and the discriminating NTMs, respectively.

Overestimates occurred because that model failed to account for any of the network capacity lost in storing cordon queues.



(a) Trip completion rates in A

(b) Cordon-crossing flows from A to B

Figure 5: Predictive strengths of different NTMs

We turn now to the time-varying rates at which vehicles crossed the cordon when traveling from A to B. Figure 5(b) presents these predictions made by the proposed NTM and by its predecessors, as well as the simulated ground-truth rates. Note from the latter how cross-cordon flows: fell during the 30-mins of restrictive metering; rebounded when metering was discontinued; and gradually dropped thereafter when cordon queues (and demand) diminished. Note too that the proposed NTM matches that pattern best.[8]

The original NTM underestimated cross-cordon flows during the hour that followed metering's deactivation. Because that model sent back vehicles that were queued at the cordon, they were not on the scene when opportunity came to saturate the cordon. The more discriminating NTM overestimated cordon crossings late in the test. The model mistakenly assumed that flows from A to B still saturated the cordon after queues and demand had dissipated.

5.2 Control Outcomes

Our NTM's enhanced predictions lead to improved cordon-control policies. To demonstrate, initial demand was increased to 20,000 vph for each O-D pair bound for A, and was lowered to 5,000 vph for each remaining O-D. All inter-neighborhood trips were subject to cordon metering. As per the reasoning in footnote 2, the control variable, u_{ij}, was constrained in [0.33,1.0], which are expressed as ratios of the metering rate to the cordon's unmetered capacity. Optimal values, $u^{\star}_{ij}$, were separately obtained from the proposed NTM and from its predecessors. To facilitate fair comparisons: the $u^{\star}_{ij}(t)$ were in all cases selected using the iLQR method with a rolling horizon of $H = 20$ steps, each of duration $\tau = 5$ mins;[9] and the resulting time-varying accumulations over the network were always generated using AIMSUN. Outcomes for the entire network are shown in Fig.6, along with a fourth curve for the do-nothing (no metering) case. Each curve is the average of five simulations with distinct random seeds.

[8]Predictions from the proposed model produced an RMSE of 12.3 vehicles/min. The RMSEs for the original and the discriminating NTMs were 35.1 vehicles/min and 77.3 vehicles/min, respectively.

[9]At each control step, the iLQR method always converged within 10 iterations. Computation time for each step was about 5 seconds on an Intel i7-4770 CPU @ 3.4GHz. The algorithm was implemented in Python with Numpy and Tensorflow.

The proposed NTM wins again. It reduced VHT on the network (the area under a curve) by nearly 15% relative to the do-nothing strategy. The original NTM reduced VHT by only 3%. Network VHT increased by 3% under the more discriminating model. Hence, the proposed model saved 12% more VHT than did the original one, and 18% more than did the discriminating one.

To explain these outcomes, we examine the optimal metering actions imposed on vehicles traveling from B to A, $u^{\star}_{BA}(t)$. The values selected by the proposed NTM and its predecessors are shown in Fig.7. Note that in all cases, the controller alternated between the minimum and maximum values allowed, $u_{min} = 0.33$ and $u_{max} = 1.0$. The outcome supports the theoretical analysis in Appendix A indicating that bang-bang metering tends to be dominant over a control period.

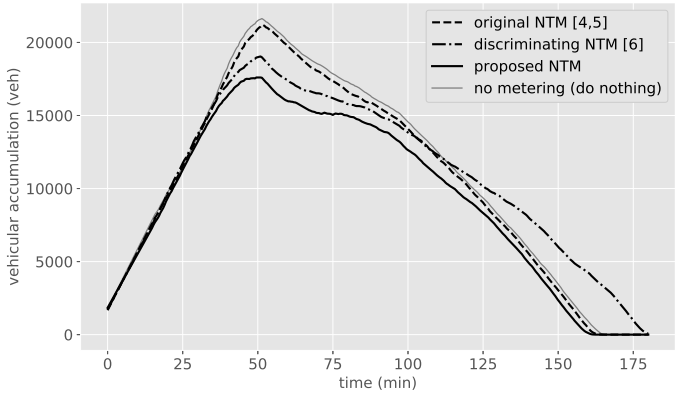
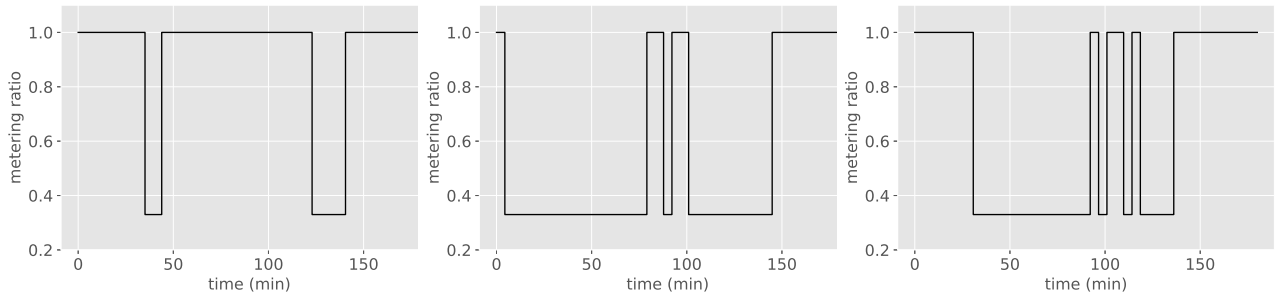


Figure 6: Network-wide vehicle accumulation over time

Note by comparing Fig.7(a) with its counterparts, how the original NTM [4, 5] selected u_{min} sparingly relative to the other models; i.e. u_{min} was deployed under the original model only twice, and only for brief periods during the 3-h test. By underestimating cordon-crossing rates (as we saw in Fig.5(b)), the original NTM overestimated metering's negative impacts. It therefore selected a metering policy that was quite relaxed. This is why its resulting network accumulations so closely resemble the do-nothing strategy, as evident in Fig.6.

The control policy generated by the more discriminating NTM [6] is once again a different story. By ignoring the constraining effects of cordon queues and assuming that meters were always saturated, that model underestimated metering's negative impacts. Its selected control policy was therefore a very restrictive one. Note from Fig.7(b) how most of the test period was metered at u_{min} and how that restrictive control was twice deployed for relatively extended durations. This aggressive control slowed the network's eventual recovery. Note how the dot-dash curve in Fig.6 lies above its thin-solid (do-nothing) counterpart during the test's final hour or so.

The proposed NTM generated the metering pattern in Fig.7(c). This more effective policy is a kind-of compromise between those of its predecessors; i.e. note how the pattern in Fig.7(c) alternates more frequently between u_{min} and u_{max}, and how the former tends to be deployed for durations that fall between those selected using the other models.



(a) Original NTM (b) Discriminating NTM (c) Proposed NTM

Figure 7: Control actions for vehicles crossing from B to A

6 Summary and Future Research

Like the others of its kind, the Neighborhood Transmission Model presently proposed addresses the cordon-control problem on large geographic scales. Like its most recent predecessor [6], our NTM treats a neighborhood's circulating vehicles differently from those queued at its cordon. Unlike other NTMs, however, ours approximately accounts for the time-varying street space occupied by cordon queues, and models the attendant effects on circulating traffic. This is done by re-scaling the neighborhood's MFD with the assumption that cordon queues always exhibit jam densities.

The assumption is supported to some degree by our theoretical finding that an optimal metering strategy entails bang-singular control, which includes very restrictive metering; see Appendix A. Experimental support for this finding is furnished from simulations. The simulations further indicate that the proposed NTM does a better job of predicting neighborhood traffic dynamics than do its predecessors, and that this leads to more effective cordon-metering policies. Policies were obtained by discritizing system dynamics into small time steps and applying the iLQR method to solve a rolling-horizon optimization. The method scales linearly with the number of variables (see Appendix B), and always converged in relatively few iterations (see again footnote 7).

Inputs needed at the start of each planning horizon may in the future come from connected vehicles. Our estimation methods that stemmed from this were a bit coarse, particularly the assumption that vehicles in a cordon queue need travel zero distance to reach the cordon line. Since the assumption will be subject to greater error as the queue lengths expand at cordons, future research might seek to refine the present estimates. Perhaps refinements will come by further coupling estimates with real-time measurements.

It is further worth noting that the proposed NTM implicitly assumes that inter-neighborhood-bound vehicles evenly distribute themselves across a cordon. This may be reasonable when drivers can readily adjust their routes in response to cordon queues. Still, our ongoing research on the subject indicates that network performance can often be improved by varying metering rates along a cordon line, in part to balance queue lengths there.

Finally, the proposed NTM assumes that the partitioning of neighborhoods occurs in static and *a priori* fashion. Our ongoing work in this realm suggests that there is merit in adapting cordon sizes and locations in real time, to accommodate a city's evolving congestion patterns.

Appendix A: "Bang-Singular" control

We show that the optimal control action, $u^\star$, is "bang-singular" type by re-writing the discrete optimization problem of (4.1) in continuous fashion.[10] Thus,

$$\begin{aligned} & \int_{t_0}^{t_0+H\tau} 1^T \mathbf{n} dt \\ s.t. \quad & \dot{\mathbf{n}} = \lambda + f(\mathbf{n}) + g(\mathbf{n}, \mathbf{u}) \\ & \mathbf{u} \in [0, 1], \end{aligned}$$

where $\mathbf{n}$ is the vector of $[...n_{c,ij}...n_{q,ij}...]$; $\mathbf{u}$ is the vector of $[...u_{c,ij}...u_{q,ij}...]$; λ is traffic demand; function f is the uncontrolled part of the system dynamics; and function g represents the cordon-crossing flows.

Thanks to the special form of $g(\mathbf{n}, \mathbf{u})$, we can rewrite the problem as follows:

$$\begin{aligned} & \int_{t_0}^{t_0+H\tau} 1^T \mathbf{n} dt \\ s.t. \quad & \dot{\mathbf{n}} = \lambda + f(\mathbf{n}) + A \cdot \mathbf{u} \\ & \mathbf{u} \in [u_{min}(\mathbf{n}), u_{max}(\mathbf{n})], \end{aligned}$$

where we transform function $g(\mathbf{n}, \mathbf{u})$ into $A \cdot \mathbf{u}$ and state-dependent control bounds $u_{max}(n)$ and $u_{min}(n)$.

The Hamiltonian of the system is:

$$\begin{aligned} H(\mathbf{n}, \mathbf{u}, \mathbf{p}, t) &= \mathbf{p}^T(\lambda + f(\mathbf{n}) + A \cdot \mathbf{u}) + 1^T \mathbf{n} \\ \nabla_{\mathbf{u}} H(\mathbf{n}, \mathbf{u}, \mathbf{p}, t) &= A^T \cdot \mathbf{p}, \end{aligned}$$

where $\mathbf{p}$ is the costate vector. Proof that the optimal control action, $u^\star$, is "bang-singular" is now straightforward because our system is linear in the control variables and the objective function does not involve those variables. If the coefficient in $A^T \cdot \mathbf{p}$ is nonzero, use of Pontryagin's maximum principle shows that the corresponding control is bang-bang. If any one of the coefficients in $A^T \cdot \mathbf{p}$ equals zero for a non-zero time interval, the corresponding control is of singular type, meaning that for a time control is not extreme. (We note for future reference that when undergoing control, the system trajectory is called a singular arc.) Now suppose $(A^T \cdot \mathbf{p})_i = 0$ for a non-zero time interval, i.e. the i-th coefficient equals zero and u_i is of singular type. Consider the costate equation

$$-\dot{\mathbf{p}}^T = \nabla_{\mathbf{n}} H(\mathbf{n}, \mathbf{u}, \mathbf{p}, t) = \mathbf{p}^T \cdot \nabla_{\mathbf{n}} f + 1^T,$$

and multiply both sides by A, such that

$$-\dot{\mathbf{p}}^T A = \mathbf{p}^T \cdot \nabla_{\mathbf{n}} f \cdot A + 1^T \cdot A.$$

Since $(\dot{\mathbf{p}}^T \cdot A)_i = 0$ and $1^T \cdot A = 0$, we have

$$(p^T \cdot \nabla_{\mathbf{n}} f \cdot A)_i = 0,$$

[10] The continuous formulation is, of course, roughly equivalent to the discrete version when time-step duration, τ, is small.

which, together with $(A^T \cdot \mathbf{p})_i = 0$, constrain the singular arc in a zero-measure subset of the system state space. Since our use of model-predictive control optimizes the system trajectory at each step but only applies the first control action, the observed initial system state at each time step has near-zero probability of lying exactly on these singular arcs. In the unlikely event that a state did lie on the singular arcs, measurement noise and prediction errors from the embedded dynamic model would soon cause the state to drift from singular control. Optimal control therefore tends almost always to be "bang-bang", which explains why we did not see singular control in our numerical results.

Appendix B: Iterative LQR algorithm

The following algorithm is adopted from [9, 10]. Consider the trajectory optimization problem

$$\begin{aligned} min_U \quad & l_f(x_N) + \sum_{1 \le i < N} l(x_i, u_i) \\ s.t. \quad & x_{i+1} = f(x_i, u_i), \forall 1 \le i < N, \end{aligned}$$

where x is the system state; u is the control variable; f is the system dynamic; l is the cost function; and l_f is the terminal state cost. By defining $U_i = \{u_i, u_{i+1}, ...u_{N-1}\}$, the cost to go function J is defined as

$$J_i(x_i, U_i) = \sum_{i \le j \le N-1} l(x_j, u_j) + l_f(x_N).$$

The value function is defined by $V(x_i, i) = min_{U_i} J_i(x_i, U_i)$, which has the recursive form

$$V(x_i, i) = min_{u_i}[l(x_i, u_i) + V(f(x_i, u_i), i + 1)].$$

The Q function is defined as

$$Q(x_i, u_i) = l(x_i, u_i) + V(f(x_i, u_i), i + 1).$$

With all the above definitions, the iLQR algorithm first initializes a random control sequence U and then alternatively performs a "backward pass" and a "forward pass" to improve U until convergence.

The backward pass is done as follows:

$$\begin{aligned} Q_x &= l_x + f_x^T V_x^{'} \\ Q_u &= l_u + f_u^T V_x^{'} \\ Q_{xx} &= l_{xx} + f_x^T (V_{xx}^{'} + \mu I_n) f_x \\ Q_{uu} &= l_{uu} + f_u^T (V_{xx}^{'} + \mu I_n) f_u \\ Q_{ux} &= l_{ux} + f_u^T (V_{xx}^{'} + \mu I_n) f_x \\ k &= -Q_{uu}^{-1} Q_u \\ K &= -Q_{uu}^{-1} Q_{ux} \\ \Delta V(i) &= \frac{1}{2} k^T Q_{uu} k + k^T Q_u \\ V_x(i) &= Q_x + K^T Q_{uu} k + K^T Q_u + Q_{ux}^T k \\ V_{xx}(i) &= Q_{xx} + K^T Q_{uu} K + K^T Q_{ux} + Q_{ux}^T K, \end{aligned}$$

where μ is a regularization factor. Note that in our application, $l_{xx}, l_{uu}, l_{ux}, l_u$ are all zeros, i.e. no direct hessian matrix computation is involved, which can save considerable computation time and makes the iLQR algorithm ideal for our application.

The forward pass to update the trajectory $(\hat{x}, \hat{y}, i)$ is:

$$\begin{aligned} \hat{u}(i) &= u(i) + \beta k(i) + K(i)(\hat{x}(i) - x(i)) \\ \hat{x}(i+1) &= f(\hat{x}(i), \hat{u}(i)) \\ \hat{x}(1) &= x(1); \end{aligned}$$

where β is the step size found by line-search. The iLQR algorithm repeats the backward and forward passes alternatively until the trajectory converges.

References

[1] Carlos F Daganzo. Urban gridlock: macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological*, 41(1):49–62, 2007.

[2] Konstantinos Aboudolas and Nikolas Geroliminis. Perimeter and boundary flow control in multi-reservoir heterogeneous networks. *Transportation Research Part B: Methodological*, 55: 265–281, 2013.

[3] Mehdi Keyvan-Ekbatani, Markos Papageorgiou, and Victor L Knoop. Controller design for gating traffic control in presence of time-delay in urban road networks. *Transportation Research Part C: Emerging Technologies*, 59:308–322, 2015.

[4] Nikolas Geroliminis, Jack Haddad, and Mohsen Ramezani. Optimal perimeter control for two urban regions with macroscopic fundamental diagrams: A model predictive approach. *IEEE Transactions on Intelligent Transportation Systems*, 14(1):348–359, 2013.

[5] Mohsen Ramezani, Jack Haddad, and Nikolas Geroliminis. Dynamics of heterogeneity in urban networks: aggregated traffic modeling and hierarchical control. *Transportation Research Part B: Methodological*, 74:1–19, 2015.

[6] Jack Haddad. Optimal perimeter control synthesis for two urban regions with aggregate boundary queue dynamics. *Transportation Research Part B: Methodological*, 96:1–25, 2017.

[7] Jordi Casas, Jaime L Ferrer, David Garcia, Josep Perarnau, and Alex Torday. Traffic simulation with aimsun. In *Fundamentals of traffic simulation*, pages 173–232. Springer, 2010.

[8] Nikolas Geroliminis and Alexander Skabardonis. Identification and analysis of queue spillovers in city street networks. *IEEE Transactions on Intelligent Transportation Systems*, 12(4):1107–1115, 2011.

[9] Weiwei Li and Emanuel Todorov. Iterative linear quadratic regulator design for nonlinear biological movement systems. In *ICINCO (1)*, pages 222–229, 2004.

[10] Yuval Tassa, Tom Erez, and Emanuel Todorov. Synthesis and stabilization of complex behaviors through online trajectory optimization. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2012)*, pages 4906–4913. IEEE, March 2017.