

Markov chains

Beyond basic facility dynamics

Cathy Wu

1.041/1.200 Transportation: Foundations and Methods

References

1. Optional readings: Prof Ayalvadi Ganesh. "Markov Chains." Lecture notes.
https://people.maths.bris.ac.uk/~maajg/teaching/pgt/markov_chains.pdf
2. Wu, Cathy, Eugene Vinitzky, Aboudy Kreidieh, and Alexandre Bayen. "Multi-lane reduction: A stochastic single-lane model for lane changing." In *International Conference on Intelligent Transportation Systems (ITSC)*, 2017.
3. Slides adapted from Carolina Osorio

Outline

1. A few more notes on analytical queues
2. Global balance equations
3. Markov chains

Outline

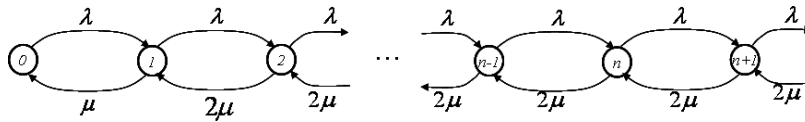
1. **A few more notes on analytical queues**
 - a. $M/M/2$
 - b. $M/M/c$
 - c. $M/M/c/K$
2. Global balance equations
3. Markov chains

Facilities in transportation

- So far, we have studied stochastic models for **a single transportation facility**:
- Examples:
 - A single bus stop
 - A single intersection
 - A single road (this lecture)
 - A single gas station (may have multiple gas pumps)
 - A single subway station (may have multiple trains)
 - A single airport (may have multiple runways)
 - A single port (may have many docking areas)
- Next lecture: **multiple facilities**, that affect one another (**queuing networks**)
- Examples
 - A road network
 - A bus system
 - A rail / subway system

$M/M/2$ queueing system

- What happens if we have two parallel, independent servers, each with service rate μ and exponentially distributed service times?
- State transition diagram:



- Balance equations:
$$\begin{cases} \lambda P_0 = \mu P_1 \\ \lambda P_1 = 2\mu P_2 \\ \lambda P_2 = 2\mu P_3, \dots \end{cases}$$

$$\begin{cases} P_1 = \frac{\lambda}{\mu} P_0 \\ P_2 = \frac{\lambda}{2\mu} P_1 = \frac{1}{2} \left(\frac{\lambda}{\mu}\right)^2 P_0 \\ P_3 = \frac{\lambda}{2\mu} P_2 = \left(\frac{1}{2}\right)^2 \left(\frac{\lambda}{\mu}\right)^3 P_0 \\ P_n = \left(\frac{1}{2}\right)^{n-1} \left(\frac{\lambda}{\mu}\right)^n P_0 \end{cases}$$

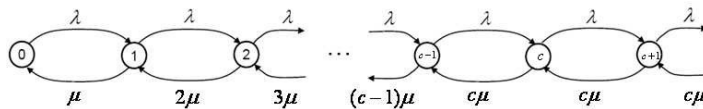
$M/M/2$ queue

$$P_0 = \frac{1 - \frac{\lambda}{2\mu}}{1 + \frac{\lambda}{2\mu}} \quad \text{for } \lambda < 2\mu$$

$$P_n = 2 \left(\frac{\lambda}{2\mu} \right)^n P_0, \quad n \geq 1$$

$$\dots \bar{N} = \frac{\frac{\lambda}{\mu}}{\left(1 + \frac{\lambda}{2\mu}\right) \left(1 - \frac{\lambda}{2\mu}\right)}$$

$M/M/c$ queue



- This model is a reasonable assumption at toll booths on turnpikes or at toll bridges where there is often more than one toll booth open.
- Traffic intensity / utilization factor: $\rho = \frac{\lambda}{c\mu}$
- Stability: $\frac{\lambda}{c\mu} < 1$
- Stationary dbn:

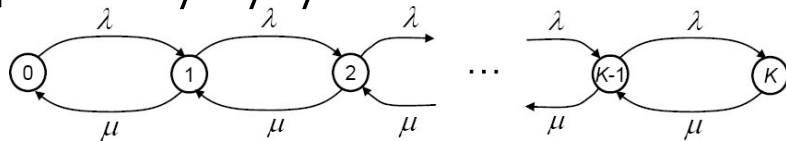
$$P_k = \begin{cases} \frac{(\lambda/\mu)^k}{k!} P_0, & k = 1, 2, \dots, c-1 \\ \frac{(\lambda/\mu)^k}{c! c^{k-c}} P_0, & k = c, c+1, \dots \end{cases}$$

$$P_0 = \left[\frac{(\lambda/\mu)^c}{c! (1 - \rho)} + \sum_{k=0}^{c-1} \frac{(\lambda/\mu)^k}{k!} \right]^{-1}$$

- Expected queue length (in the buffer)?

$M/M/c/K$

- Finite capacity queue: $M/M/1/K$



$$P_n = \frac{1 - \rho}{1 - \rho^{K+1}} \rho^n \quad n = 0, 1, \dots, K$$

- The queue is always stable ($\forall \rho$), so steady state is always reached
- When system is full: arrivals are lost
- Effective* arrival rate (i.e. rate of arrivals that actually enter the system): $\lambda(1 - P_K)$
- Careful when applying Little's Law! Count only the vehicles that actually join the system:

$$\begin{aligned} \lambda' &= \lambda(1 - P_K) \\ \bar{N} &= \lambda' \bar{T} \end{aligned}$$

Outline

1. A few more notes on analytical queues
2. **Global balance equations**
 - a. Transition rate matrix
 - b. Container terminal problem
3. Markov chains

Beyond the basic queues

- In this lecture, we discuss **a more general strategy** for modeling and analyzing more sophisticated facility dynamics.
- We will introduce ideas of: transition rate matrix, global balance equations. We will leverage fundamental results from Markov chains.
- Solutions will no longer be analytical (in general) and can be instead computed **numerically**.
- For now, we will still use properties of the **Exponential distribution**. In Lecture 11, we will relax this to allow modeling of (even) more general facility dynamics (see Computational Lab #2).
- **Why do we need fancier tools?**

Container terminal

■ Arrival

- Ships arrive at a container terminal according to a Poisson process with rate λ . Each ship brings thousands of containers.
- There are c berths, i.e. only c ships can dock or anchor at the port simultaneously.
- If all berths are occupied arriving ships must wait at the entrance of the port. A maximum of r ships can wait for entrance, further ships are redirected to another port.

■ Berth operations

- Upon arrival to the berth, a ship unloads its containers. The unloading time follows an exponential distribution with parameter μ .



Departure

- Once all containers are unloaded, the ship can leave the port, as long as the exit is clear, which occurs with probability p_f .
- Suppose each berth has its own exit. If the exit is occupied, it remains so during an exponentially distributed time with parameter β .

Container terminal

Upon arrival to the port a ship (arrival rate λ):

1. [queue]
2. is served (service rate μ)
3. [blocked] (rate at which exit is occupied β , also exponentially distributed)
4. Departs (with probability p_f)

Note: Blocking after service: blocking mechanism used to describe how congestion arises and propagates

State space of port :

$$S = \{(A, B, W) \in N^3, A + B \leq c, W \leq r\}$$

- A = ships being served
- B = ships waiting to depart
- W = ships waiting to be served

Stationary distribution:

$$P = (P((A, B, W) = (a, b, w)), (a, b, w) \in S)$$

Problem: For $c = 2, r = 2$, define the possible transitions and their corresponding rates (tabulate).

Reminder (Departure): Once all containers are unloaded, the ship can leave the port, as long as the exit is clear, which occurs with probability p_f . Suppose each berth has its own exit. If the exit is occupied, it remains so during an exponentially distributed time with parameter β .



General strategy for stationary analysis

■ Time perspective (what we've considered so far):

- The random process $X_t \in \mathbb{N}$ evolves according to transition rates.
- Represented by state transition diagram and/or transition rate matrix Q .



■ Distributional perspective: $P_n(t) := P(X_t = n)$

- Discrete-time form: $P_n(k) := P(N(\Delta t k) = n)$
- $P(k) := (P_n(k))_{n \in \mathbb{N}}$

1. Define the evolution of $P(k)$ as:

$$P(k+1)^T = P(k)^T T$$

- Where $T := I + Q\Delta t$ is the transition probability matrix



2. Invoke fundamental theorem of Markov Chains.

- Gives convergence to stationary distribution: $P(k) \rightarrow P$



Transition rate matrix

- $Q = (q_{ij})_{i \in I, j \in J}$, where q_{ij} is transition rate from state i to state j
- Possible transitions from a given state i , with their corresponding rates and conditions under which these transitions can take place
- Example: M/M/1

initial state s	new state j	rate q_{ij}	condition
n	$n + 1$	λ	always possible
n	$n - 1$	μ	$n \geq 1$

Container terminal

Upon arrival to the port a ship (arrival rate λ):

1. [queue]
2. is served (service rate μ)
3. [blocked] (rate at which exit is occupied β , also exponentially distributed)
4. Departs (with probability p_f)

Note: Blocking after service: blocking mechanism used to describe how congestion arises and propagates

State space of port :

$$S = \{(A, B, W) \in N^3, A + B \leq c, W \leq r\}$$

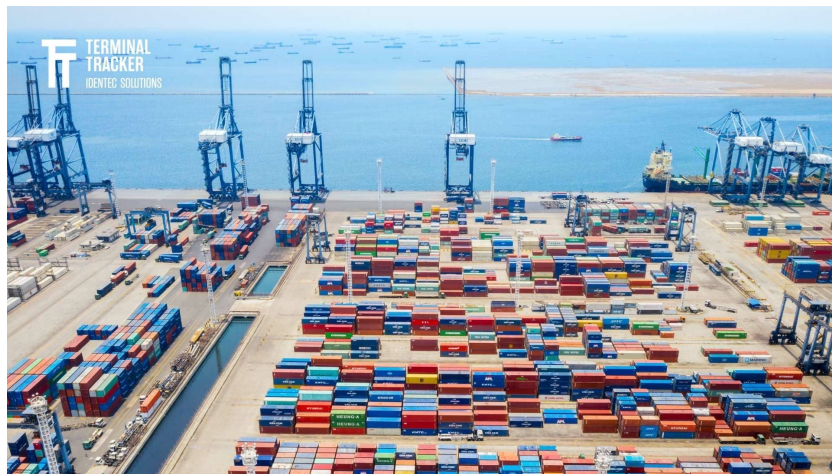
- A = ships being served
- B = ships waiting to depart
- W = ships waiting to be served

Stationary distribution:

$$P = (P((A, B, W) = (a, b, w)), (a, b, w) \in S)$$

Problem: For $c = 2, r = 2$, define the possible transitions and their corresponding rates (tabulate).

Reminder (Departure): Once all containers are unloaded, the ship can leave the port, as long as the exit is clear, which occurs with probability p_f . Suppose each berth has its own exit. If the exit is occupied, it remains so during an exponentially distributed time with parameter β .



Global balance equations

Outline

1. A few more notes on analytical queues
2. Global balance equations
3. **Markov chains**
 - a. Python demo
 - b. Application: (Simple) multi-lane modeling

Markov chain

- The sequence $\{X_t, t \geq 0\}$ that goes from state i to j with probability τ_{ij} , independently of the states visited before, is a **Markov chain**.
- τ_{ij} is also called a **transition probability**.
- **Markov property**: the current state contains all information for predicting the future of the process/chain.

Preview (Unit 4): Markov decision processes (MDP)

- Extension of Markov chains, where, in addition to the current state, the transition is also affected by **decisions (control)**, i.e. τ_{ijk} .
- Decisions = traffic signal control, autonomous vehicle steering & acceleration, speed limits, bus schedules, routing apps (e.g. Google Maps, Waze), rideshare matching algorithms

Transition probability matrix

- $T = (\tau_{ij})_{i \in I, j \in J}$
- How come $P(k+1)^T = P(k)^T T$, where $T := I + Q\Delta t$?

Fundamental theorem of Markov chains

Distribution after k steps (each Δt , i.e. $t = \Delta t k$)

$$P_n(k+1) = \sum_{i \in N} P_i(k) \tau_{in}$$

Stationary distribution

$$P_n = \sum_{i \in N} P_i \tau_{in}$$

Often written in other texts
(different notation) as: $\pi P = \pi$.

Definitions:

- Finite: finite number of possible states ($N < \infty$)
- Irreducible: can reach any state from any state, possibly after many steps
- Aperiodic: all states return to itself with some final cycle length > 1

Theorem (Fundamental theorem of Markov chains)

1. If the Markov chain is finite and irreducible, it has a unique invariant distribution P and P_n is the long-term fraction that $X(t) = n$.
2. If the Markov chain is also aperiodic, then the distribution $P(k)$ of $X(t)$ converges to P .

Outline

1. A few more notes on analytical queues
2. Global balance equations
3. Markov chains
 - a. **Python demo**
 - b. Application: (Simple) multi-lane modeling

Outline

1. A few more notes on analytical queues
2. Global balance equations
3. Markov chains
 - a. Python demo
 - b. Application: (Simple) multi-lane modeling**

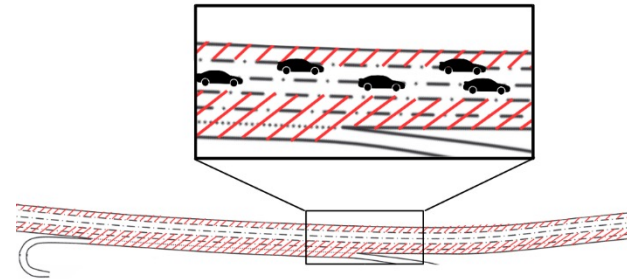


(Simple) multi-lane modeling

Goal: Enable analysis of vehicle controllers for more complex traffic, in particular for multi-lane roads.

Key challenges:

1. Driver behavior models too complex.
2. Steady-state analysis of car following models is less useful amidst frequent perturbations.

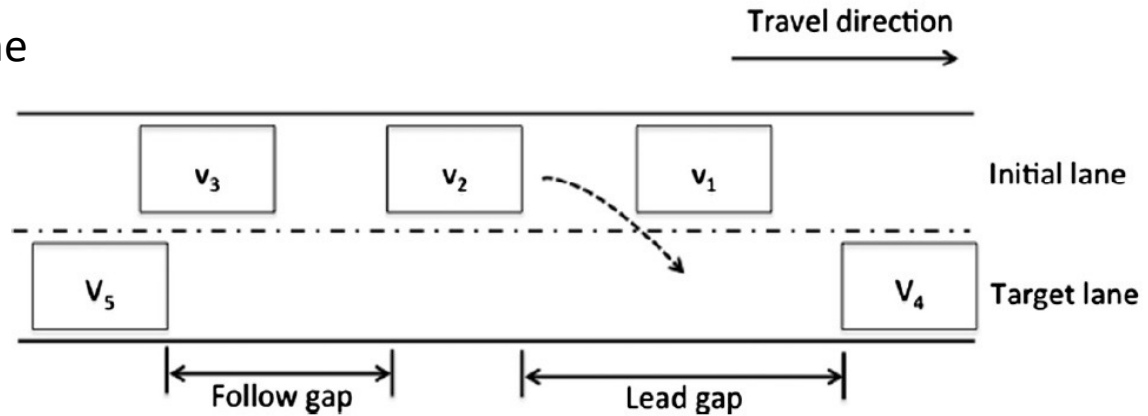


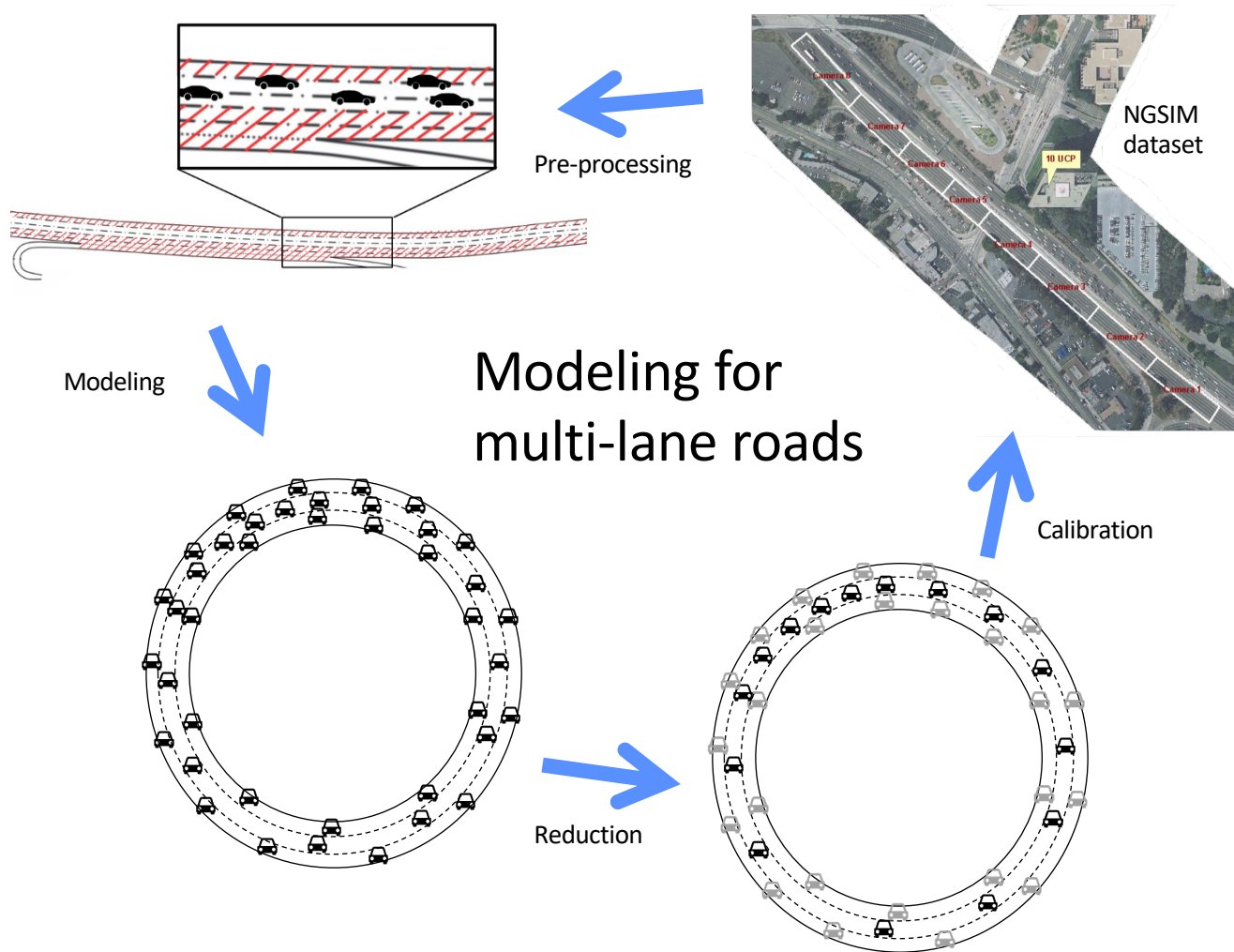
Contributions:

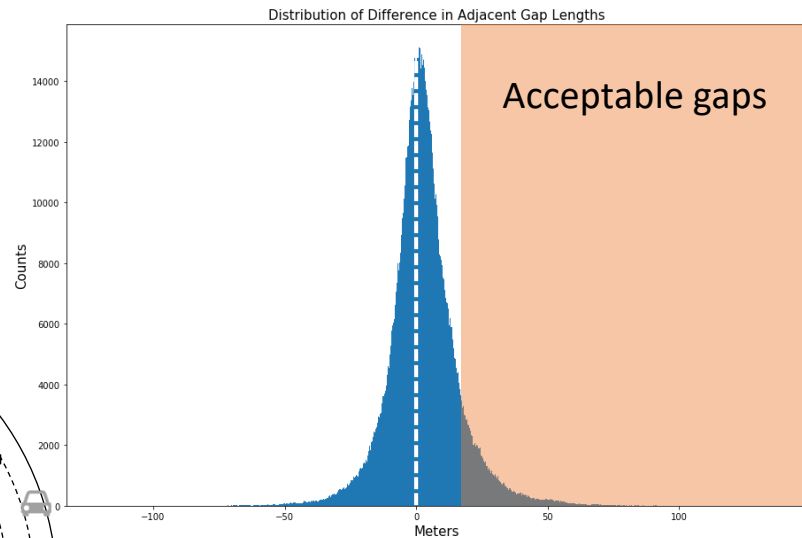
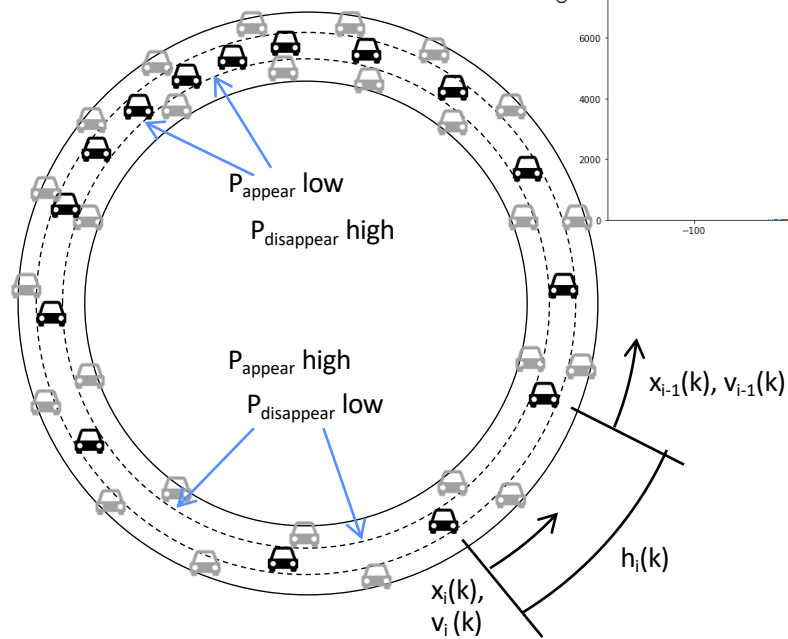
- *Multi-lane roads*: Reduction of a multi-lane model into a stochastic single-lane model.
- *Multi-lane counts*: Markov chain describing macroscopic multi-lane traffic, derived from microscopic vehicle dynamics.
- Enables analysis of variance of velocity, for understanding energy use.

Dynamical System Lane Changing

- Major source of instability
- Focus on discretionary changes
- Multifactorial
 - Gap size
 - Lead and lag speed
 - Speed of adjacent lane
 - Braking distance
 - Cooperation
 - Driver politeness
 - Traffic density



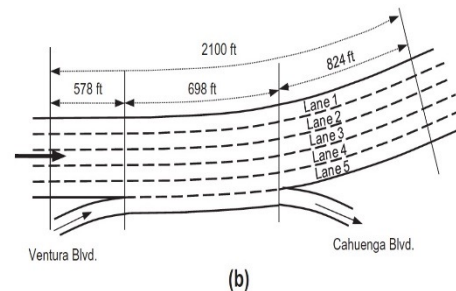
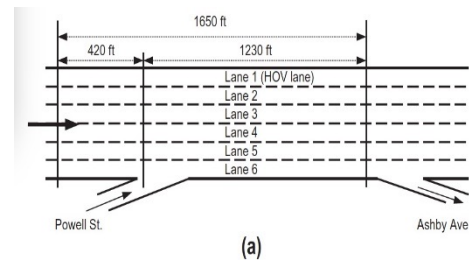
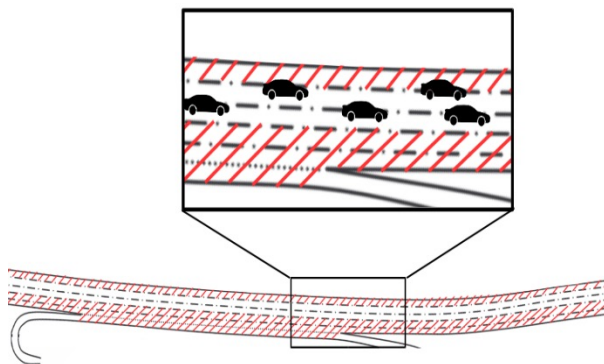




This distribution of differences in adjacent gap lengths from the NGSIM dataset indicates that there is plenty of probability mass with positive gap differences (the **right side**), providing opportunity for lane changes into the lane.

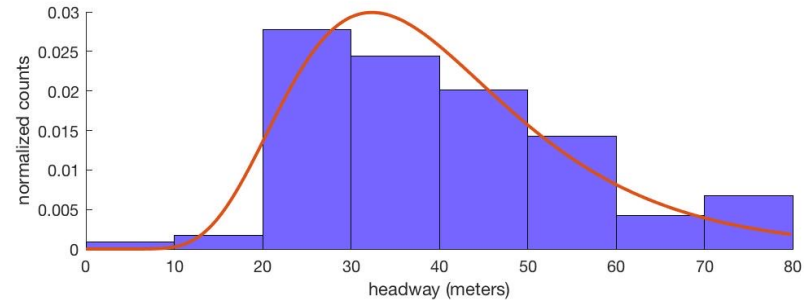
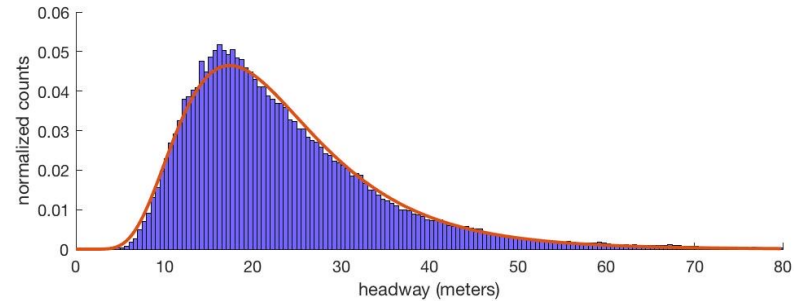
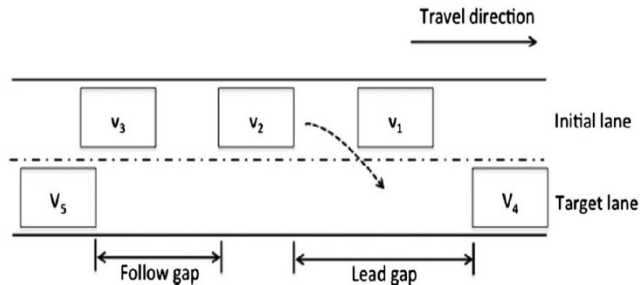
Model Calibration - NGSIM

- Two highways
- 1 1/2 hours of camera footage
- Filter out trucks
- Extract left two lanes



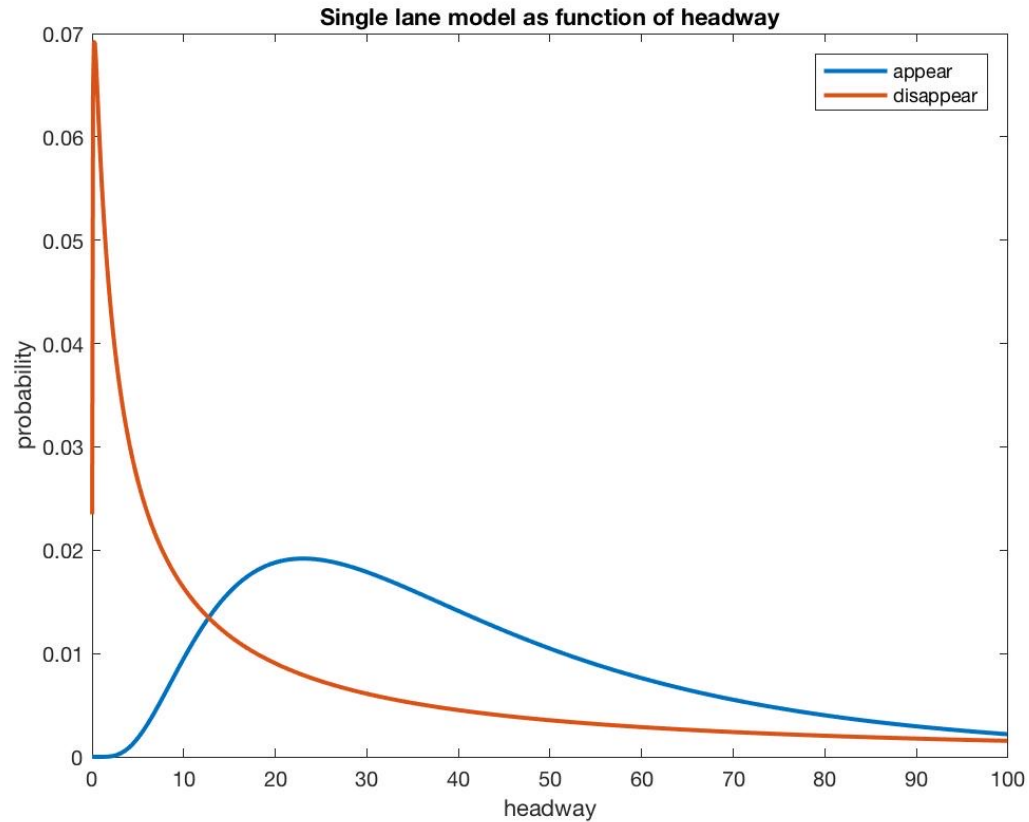
Model Calibration - Probabilities

- Extract headways when lane changes occur
- Fit to log-normal distribution
- $$p(lc|h) = \frac{p(h|lc)*p(lc)}{p(h)}$$



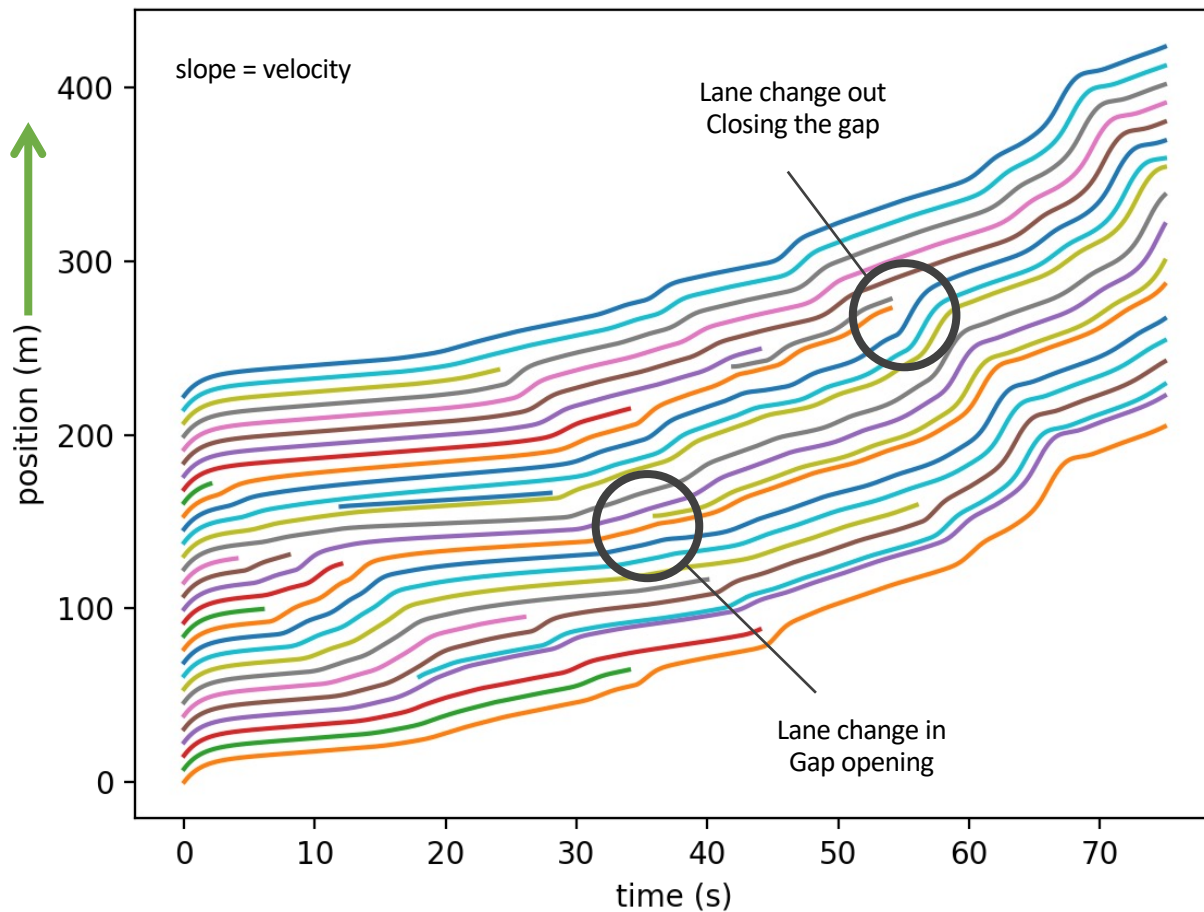
Fit of log-normal distribution for the total distribution of headways and the conditional distribution of headways when a vehicle lane changes into a lane, respectively, computed for 7:50 am on the US 101.

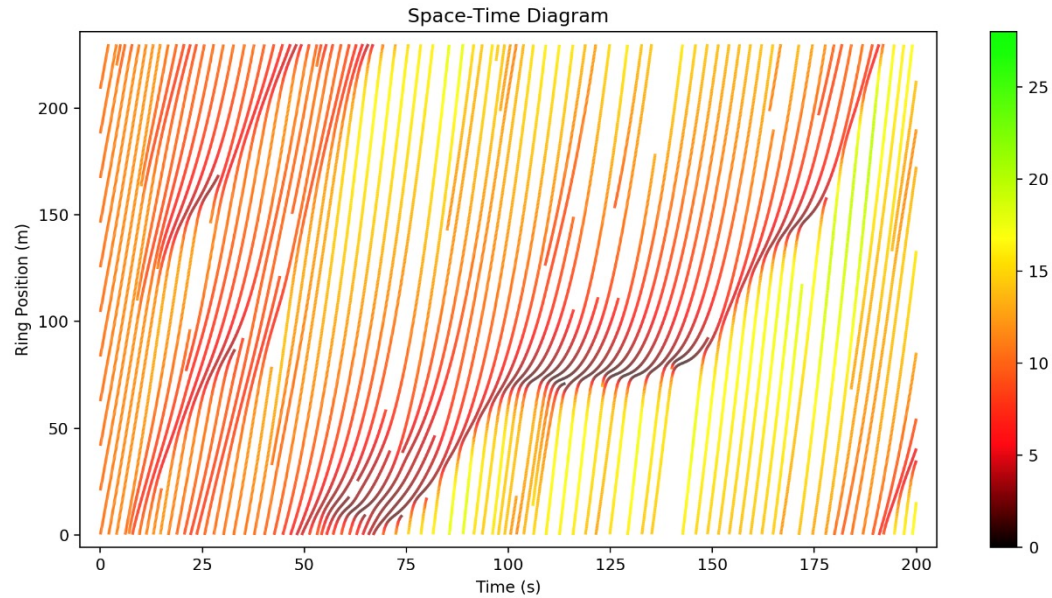
Single lane model



Time-space diagram

Position Profile of Each Car

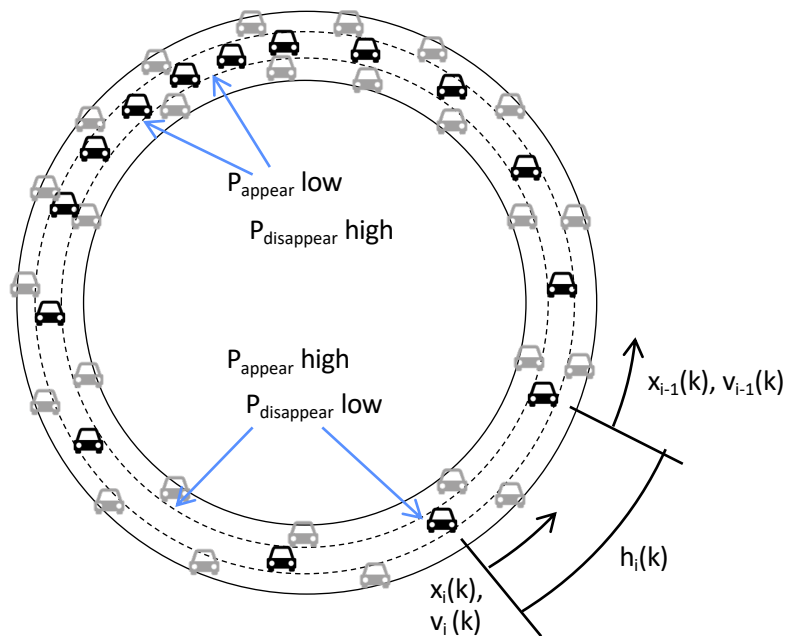




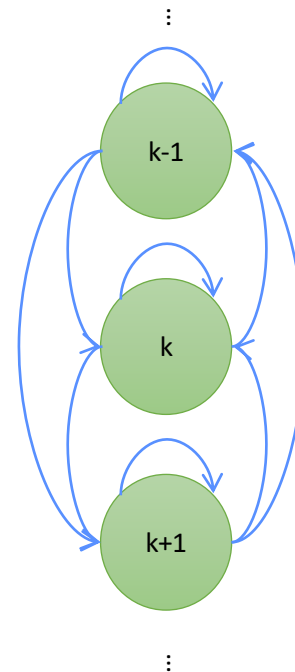
Time step (deterministic car following) = 0.025 sec
Time step (stochastic lane changing) = 6 sec

Two reductions

Stochastic single-lane model



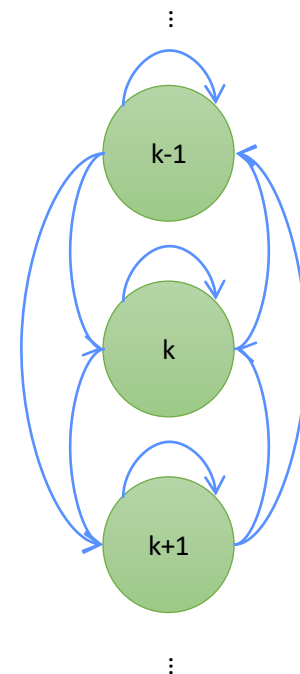
Macroscopic multi-lane traffic



Macroscopic multi-lane model

$$T_{n,n'} = \sum_{i=0}^{n-\Delta} \overbrace{\binom{n}{i+\Delta} [p_a(h_n^*)]^{\Delta+i} [1-p_a(h_n^*)]^{n-(\Delta+i)}}^{i+\Delta \text{ vehicles appear}} \times \underbrace{\binom{n}{i} [p_d(h_n^*)]^i [1-p_d(h_n^*)]^{n-i}}_{i \text{ vehicles disappear}}$$

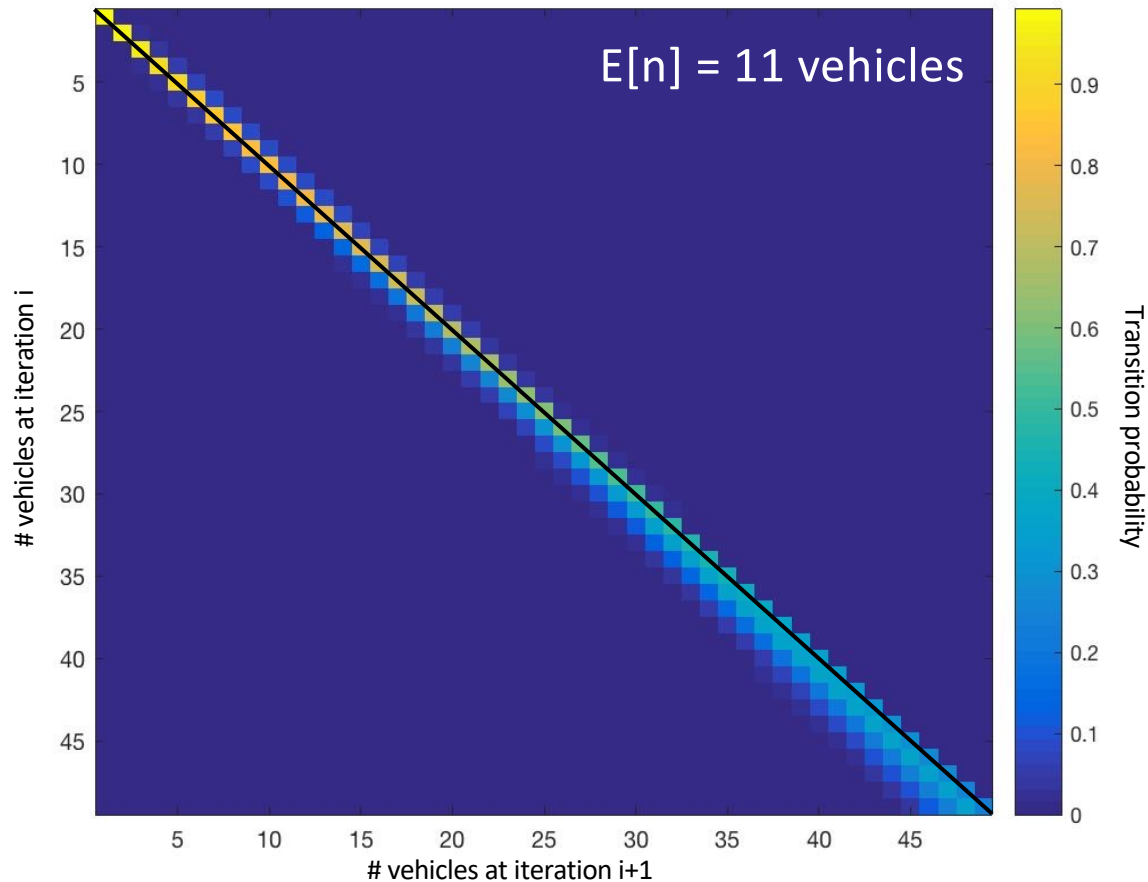
$$\Delta := n' - n$$



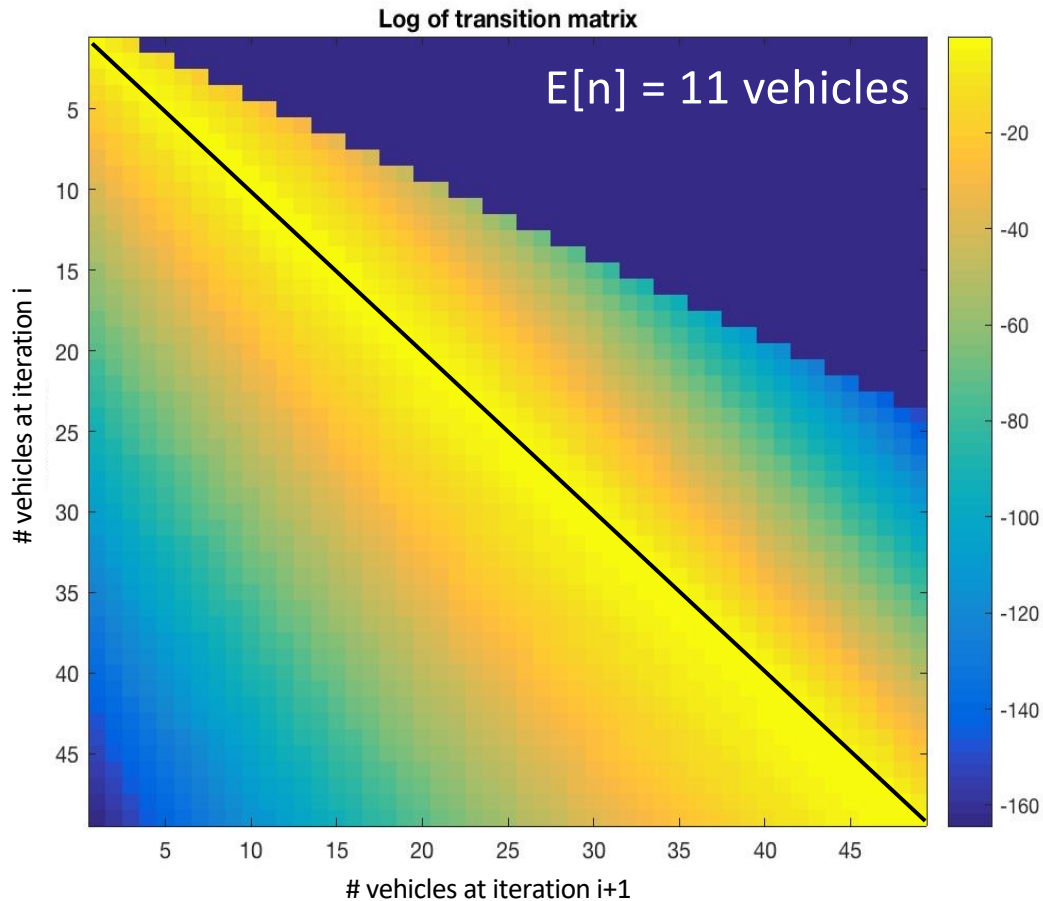
Theorem (Stationary)

This Markov chain converges to a stationary distribution.

Transition matrix for the number of vehicles in the lane



Transition matrix for the number of vehicles in the lane



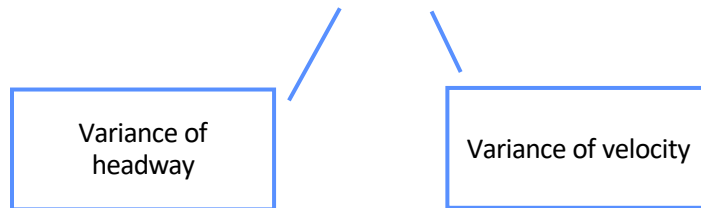
Variance analysis

Observation: Variance of velocity is related to energy consumption.

Experiment Type	σ_h^2	σ_v^2
Lane Changing Simulation	61	80
No Lane Changing Simulation	58	97
Markov Chain	61	41

(Does not model lane changing)

(Does not model transients, stop-and-go traffic)



Insight: Discretionary lane changes may aid in reducing stop-and-go traffic waves rather than increase them.