What if you don't know your probabilities?

A crash course in imprecise probabilities and their application to Markov chains

GHENT UNIVERSITY

Jasper De Bock

25 November 2019

Brighton

What if you don't know your probabilities?

A crash course in imprecise probabilities and their application to Markov chains

?

GHENT UNIVERSITY

MODELLING UNCERTAINTY

p-boxes

set theory

logic

probability measures

belief functions

random sets

imprecise probabilities

probability intervals

choice functions

sets of probability measures

lower and upper expectations

sets of desirable gambles

Markov chains → imprecise Markov chains

Bayesian networks → credal networks

And much more…

imprecise probabilities

Bounds on probabilities and expectations

Robust (set-valued) decision making

WWW.SIPTA.ORG

society for

imprecise probabilities

theories and applications

What if you don't know your probabilities?

A crash course in imprecise probabilities and their application to Markov chains

GHENT UNIVERSITY

Jasper De Bock

25 November 2019

Brighton

FLip Foundations Lab on imprecise probabilities

Jasper De Bock

GHENT UNIVERSITY

FLip

Foundations Lab on imprecise probabilities

Gert de Cooman & Jasper De Bock

GHENT UNIVERSITY




FLip

Foundations Lab on imprecise probabilities

GHENT UNIVERSITY


WWW.SIPTA.ORG

society for

imprecise probabilities

theories and applications

What if you don’t know
your probabilities?

A crash course in imprecise probabilities and their application to Markov chains

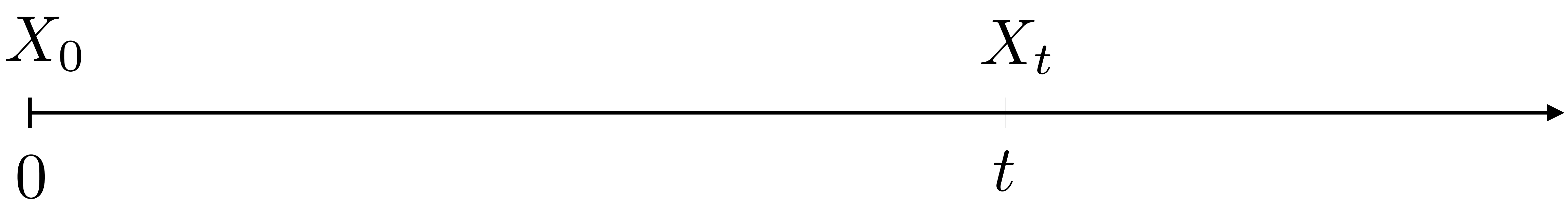
GHENT UNIVERSITY

Jasper De Bock

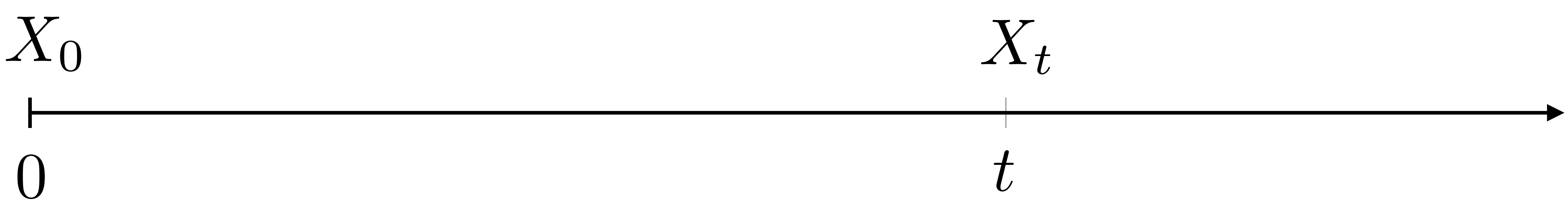
25 November 2019

Brighton

stochastic process



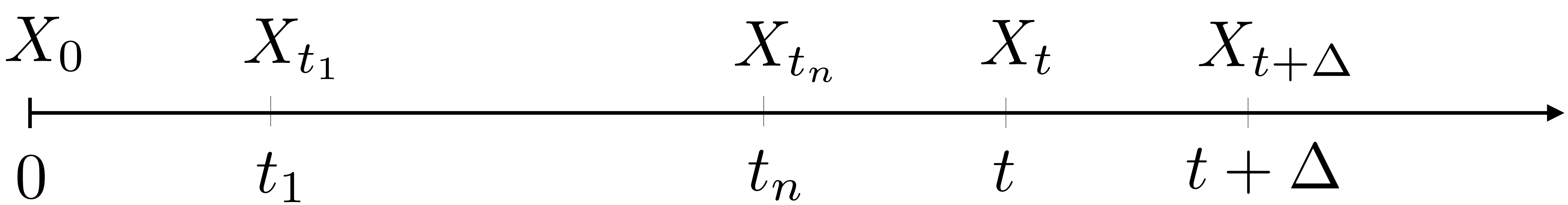
Continuous-time stochastic process



Continuous-time stochastic process

$$P(X_0 = x) = \pi_0(x)$$

$$P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x)$$

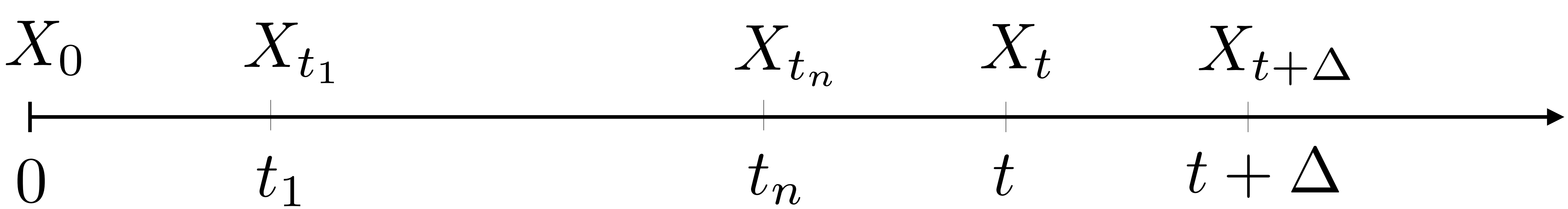


Continuous-time Markov chain

$$P(X_0 = x) = \pi_0(x)$$

$$P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x)$$
$$= P(X_{t+\Delta} = y | X_t = x)$$

Markov assumption

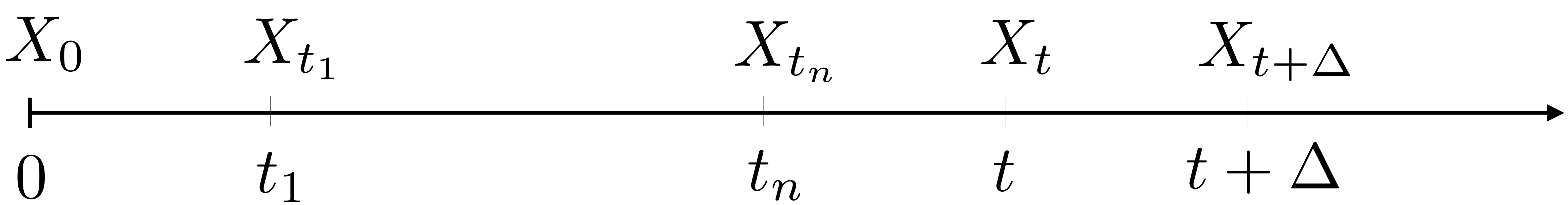


Continuous-time Markov chain

$$P(X_0 = x) = \pi_0(x)$$

$$\begin{aligned} P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \\ = P(X_{t+\Delta} = y | X_t = x) \\ \approx I(x, y) + \Delta Q_t(x, y) \end{aligned}$$

differentiability assumption

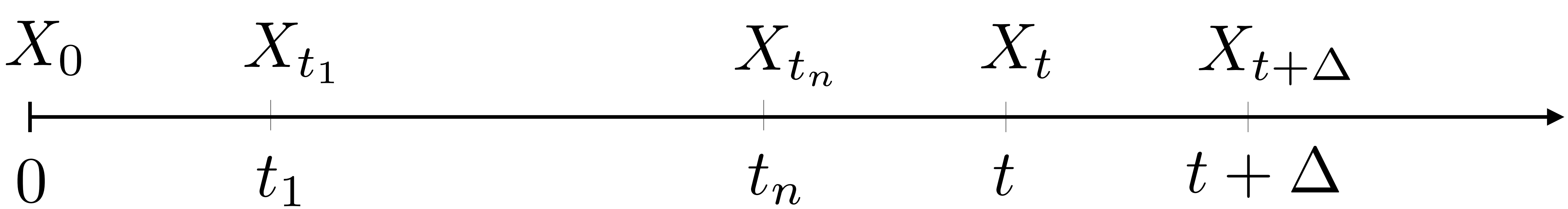


Continuous-time Markov chain

$$P(X_0 = x) = \pi_0(x)$$

$$\begin{aligned} P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \\ = P(X_{t+\Delta} = y | X_t = x) \\ \approx I(x,y) + \Delta \underset{\substack{\| \\ Q(x,y)}}{Q_t(x,y)} \end{aligned}$$

assumption of time-homogeneity



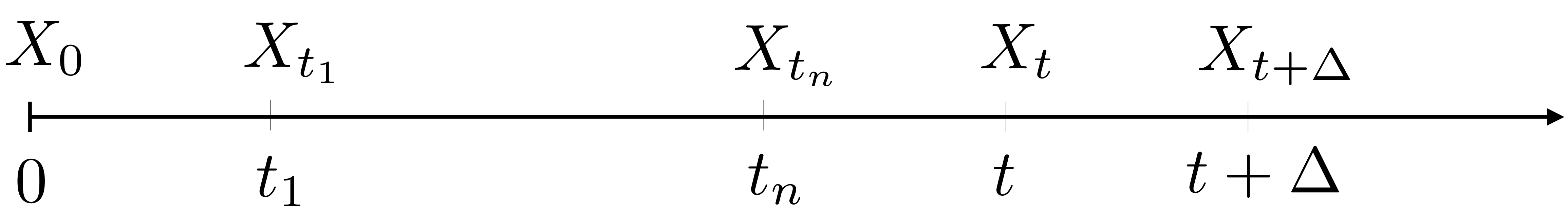
Continuous-time Markov chain

$$P(X_0 = x) = \boxed{\pi_0(x)}$$ **initial distribution**

$$P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x)$$
$$= P(X_{t+\Delta} = y | X_t = x)$$
$$\approx I(x, y) + \Delta Q_t(x, y)$$

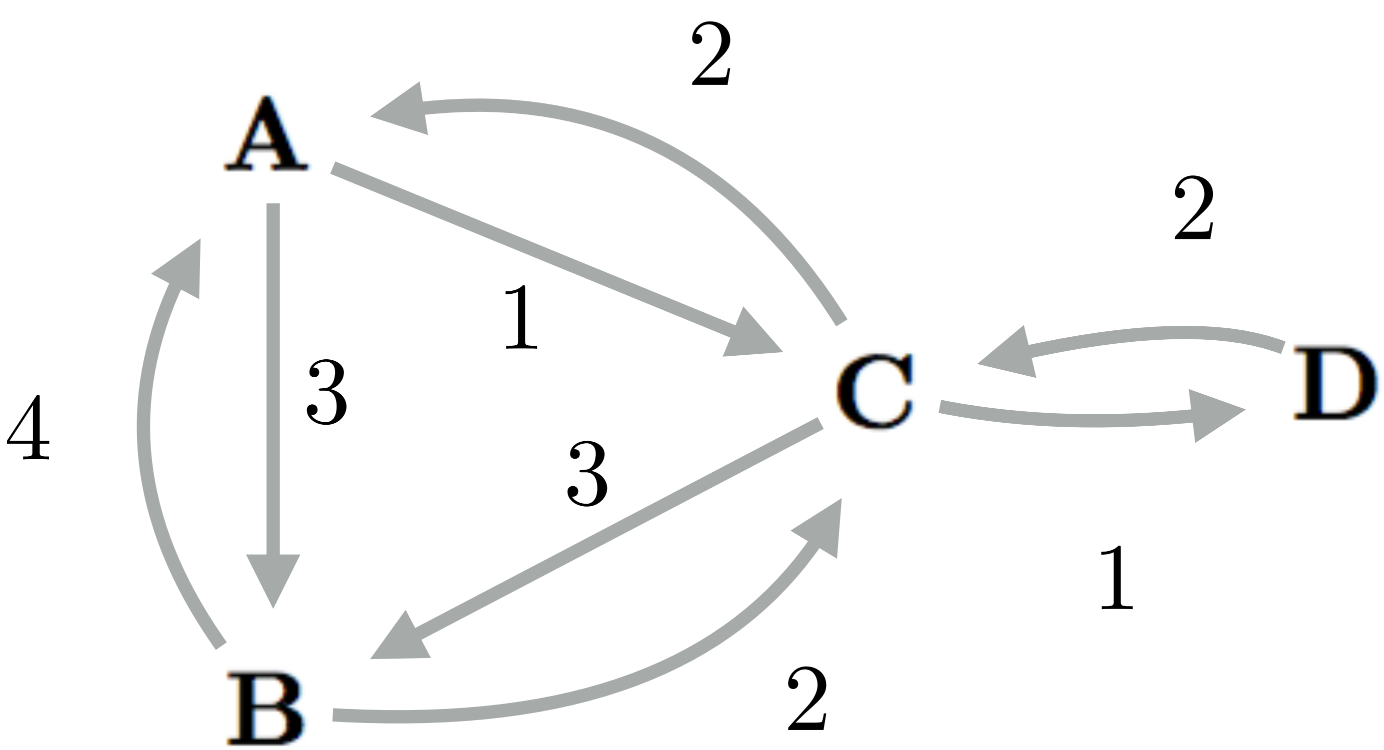
transition rate matrix $\boxed{Q(x, y)}$

$$\textstyle\sum_y Q(x, y) = 0$$
$$Q(x, x) \leq 0$$
$$(\forall y \neq x)\, Q(x, y) \geq 0$$



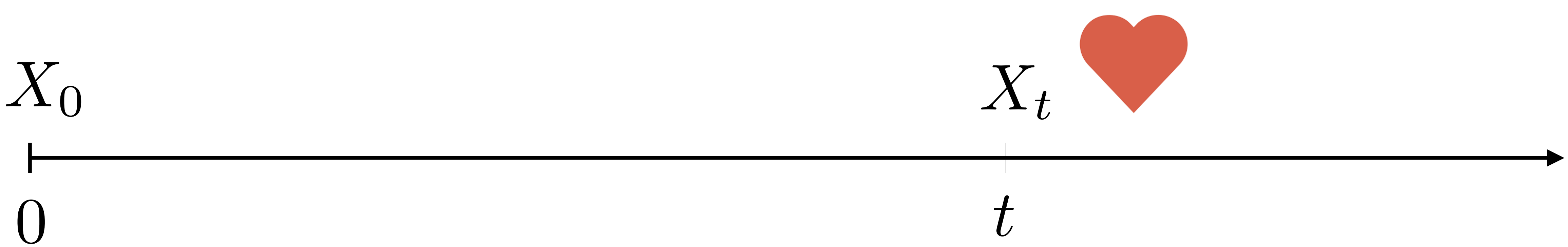
Amorous Bickering Confusion Depression

$$Q = \begin{array}{c} \\ \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{D} \end{array} \begin{array}{c} \begin{array}{cccc} \mathbf{A} & \mathbf{B} & \mathbf{C} & \mathbf{D} \end{array} \\ \begin{bmatrix} -4 & 3 & 1 & 0 \\ 4 & -6 & 2 & 0 \\ 2 & 3 & -6 & 1 \\ 0 & 0 & 2 & -2 \end{bmatrix} \end{array}$$



$$\boxed{Q(x, y)}$$

What is $P(X_t = y | X_0 = x)$?



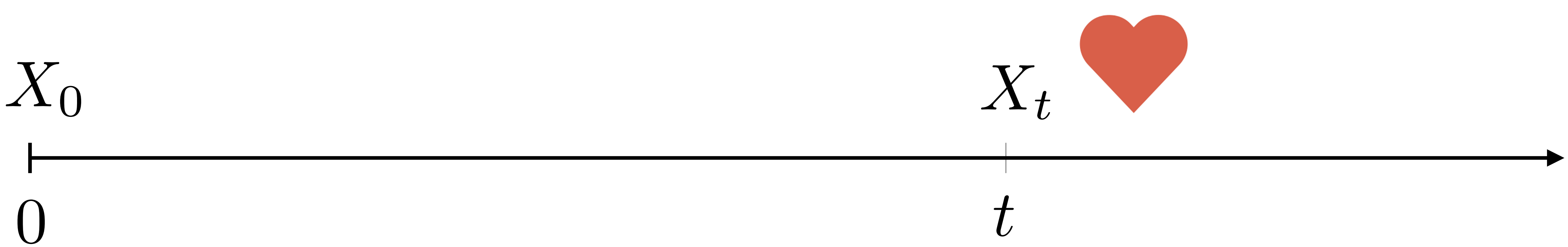
What is $P(X_t = y | X_0 = x)$?

transition matrix: $T_t(x,y) := P(X_t = y | X_0 = x)$

backward Kolmogorov differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} T_t = Q T_t \, , \text{with } T_0 = I$$

$$\Rightarrow \; T_t = e^{Qt} = \lim_{n \to +\infty} (I + \frac{t}{n} Q)^n$$



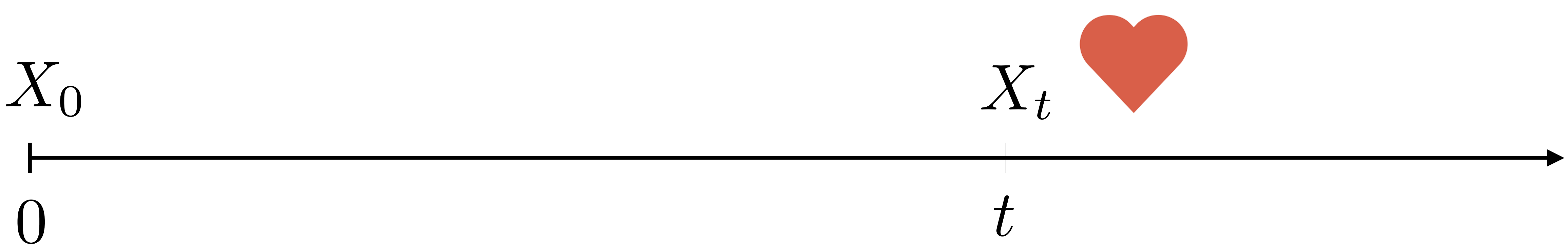
What is $P(X_t = y | X_0 = x)$? $\longrightarrow$ $\boxed{e^{Qt}(x,y)}$

transition matrix: $T_t(x,y) := P(X_t = y | X_0 = x)$

backward Kolmogorov differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t} T_t = Q T_t \,, \text{with } T_0 = I$$

$$\Rightarrow \; T_t = e^{Qt} = \lim_{n \to +\infty} (I + \frac{t}{n} Q)^n$$



What is $P(X_t = y | X_0 = x)$? $\longrightarrow$ $\boxed{e^{Qt}(x, y)}$

What is $E(f(X_t) | X_0 = x)$? $\longrightarrow$ $\boxed{[e^{Qt} f](x)}$

$x = 0$

$f(X_t) = X_t$

X_0 X_t

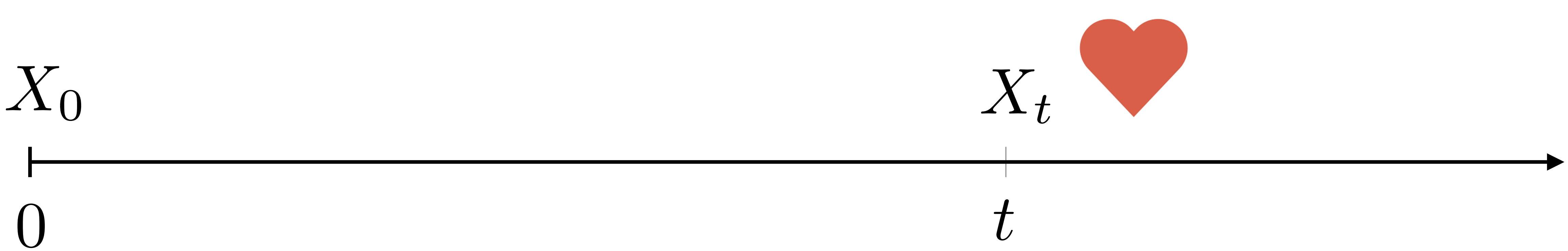
0 t

What is $P(X_t = y | X_0 = x)$? $\longrightarrow$ $\boxed{e^{Qt}(x,y)}$

$$= E(\mathbb{I}_y(X_t) | X_0 = x)$$

$$\mathbb{I}_y(X_t) = \begin{cases} 1 & \text{if } X_t = y \\ 0 & \text{otherwise} \end{cases}$$

What is $E(f(X_t) | X_0 = x)$? $\longrightarrow$ $\boxed{[e^{Qt} f](x)}$



What is $P(X_t = y | X_0 = x)$? $\longrightarrow$ $\boxed{e^{Qt}(x,y)}$

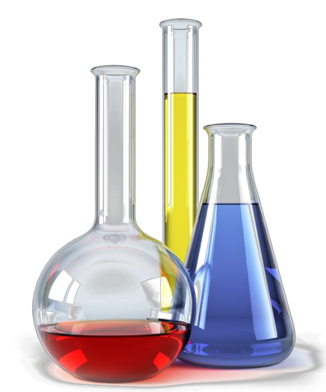
$$\pi_\infty(y) := \lim_{t \to +\infty} P(X_t = y | X_0 = x)$$

$$E_\infty(f) := \lim_{t \to +\infty} E(f(X_t) | X_0 = x)$$

What is $E(f(X_t) | X_0 = x)$? $\longrightarrow$ $\boxed{[e^{Qt} f](x)}$

- Reliability engineering (failure probabilities, …)
- Queuing theory (waiting in line …)
    - optimising supermarket waiting times
    - dimensioning of call centers
    - airport security lines
    - router queues on the internet

- Chemical reactions (time-evolution …)
- Pagerank
- …



So how about imprecision?

So how about imprecision?

What if we don't know Q exactly?

Imprecise continuous-time Markov chain

$$P(X_0 = x) = \pi_0(x)$$

$$\begin{aligned} P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \\ = P(X_{t+\Delta} = y | X_t = x) \\ \approx I(x, y) + \Delta \underbrace{Q_t}_{\|} (x, y) \end{aligned}$$

$$\boxed{Q \in \mathcal{Q}}$$

Imprecise continuous-time Markov chain

$$P(X_0 = x) = \pi_0(x)$$

$$\begin{aligned}P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \\ = P(X_{t+\Delta} = y | X_t = x) \\ \approx I(x,y) + \Delta Q_t(x,y)\end{aligned}$$

$$Q_t = \boxed{Q \in \mathcal{Q}}$$

$$\underline{E}(f) = \min_{Q \in \mathcal{Q}} E(f)$$

$$\overline{E}(f) = \max_{Q \in \mathcal{Q}} E(f) = -\underline{E}(-f)$$

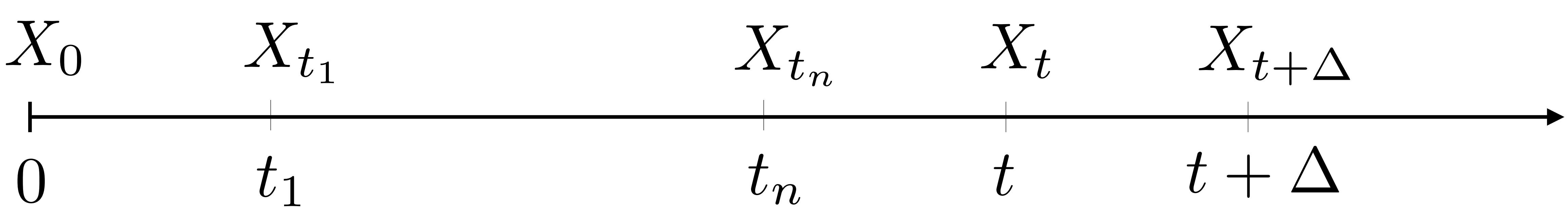
probability bounds are special cases!

Imprecise continuous-time Markov chain

$$P(X_0 = x) = \pi_0(x)$$

$$\begin{aligned} P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) &= P(X_{t+\Delta} = y | X_t = x) \\ &\approx I(x,y) + \Delta \underset{\substack{\| \\ Q \in \mathcal{Q}}}{Q_t}(x,y) \end{aligned}$$

assumption of time-homogeneity



Imprecise continuous-time Markov chain

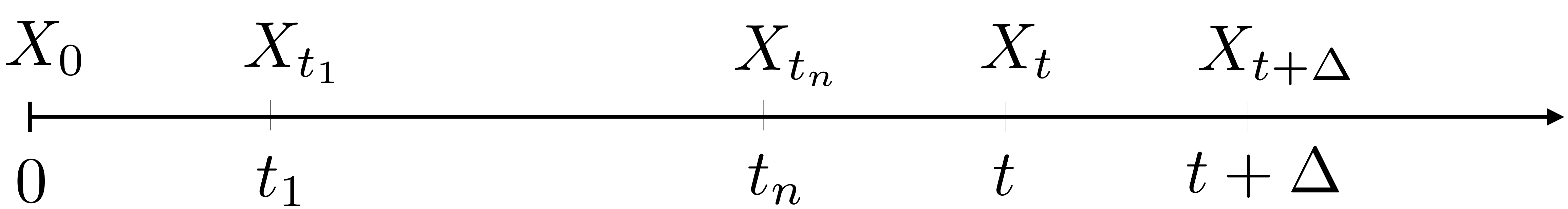
$$P(X_0 = x) = \pi_0(x)$$

$$\begin{aligned} P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \dots, X_{t_n} = x_{t_n}, X_t = x) \\ = P(X_{t+\Delta} = y | X_t = x) \\ \approx I(x,y) + \Delta Q_t(x,y) \end{aligned}$$

$$\|$$

$$Q \in \mathcal{Q}$$

~~assumption of time-homogeneity~~

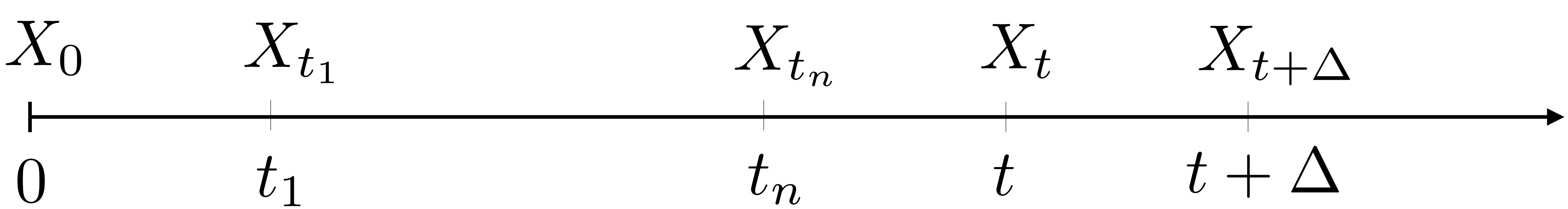


Imprecise continuous-time Markov chain

$$P(X_0 = x) = \pi_0(x)$$

$$\begin{aligned} P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \\ = P(X_{t+\Delta} = y | X_t = x) \\ \approx I(x,y) + \Delta \underset{\mathcal{Q}}{\underset{\cap}{Q_t}}(x,y) \end{aligned}$$

~~assumption of time-homogeneity~~

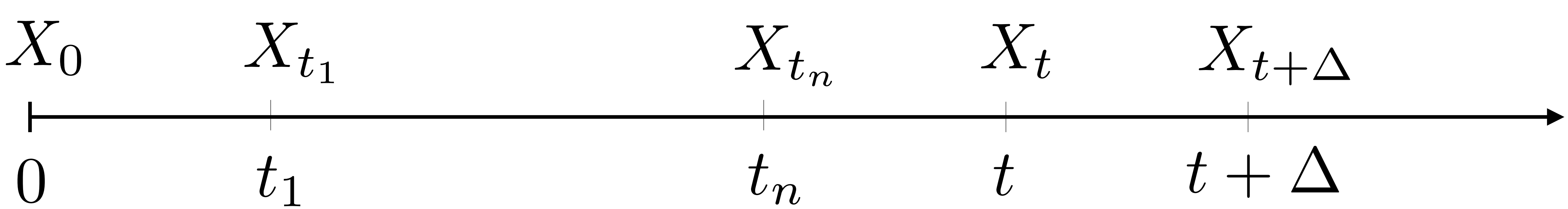


Imprecise continuous-time Markov chain

$$P(X_0 = x) = \pi_0(x)$$

$$\begin{aligned} P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \dots, X_{t_n} = x_{t_n}, X_t = x) \\ = P(X_{t+\Delta} = y | X_t = x) \\ \approx I(x,y) + \Delta \underset{\substack{\cap \\ \mathcal{Q}}}{Q_t}(x,y) \end{aligned}$$

differentiability assumption

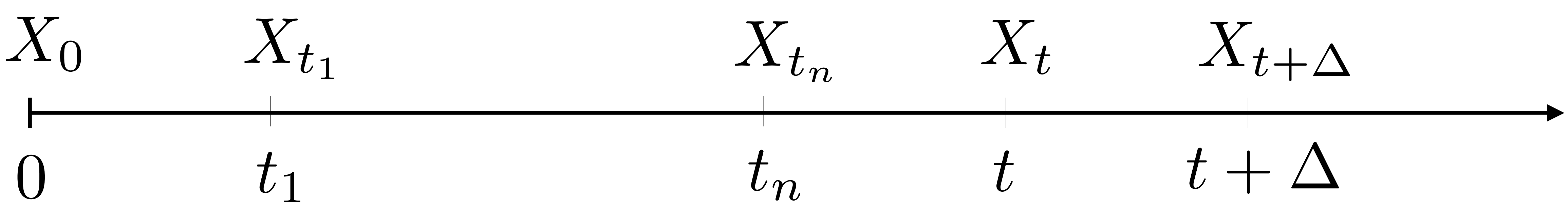


Imprecise continuous-time Markov chain

$$P(X_0 = x) = \pi_0(x)$$

$$\begin{aligned} P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \\ = P(X_{t+\Delta} = y | X_t = x) \\ \approx I(x,y) + \Delta \underset{\substack{\cap \\ \mathcal{Q}}}{Q_t}(x,y) \end{aligned}$$

~~differentiability assumption~~

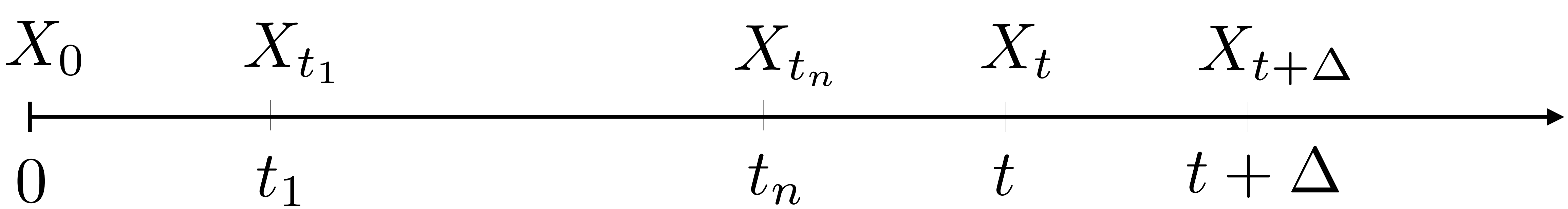


Imprecise continuous-time Markov chain

$$P(X_0 = x) = \pi_0(x)$$

$$\begin{aligned} P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \\ = P(X_{t+\Delta} = y | X_t = x) \\ \approx I(x,y) + \Delta \underset{\substack{\cap \\ \mathcal{Q}}}{Q_{t,\Delta}}(x,y) \end{aligned}$$

~~differentiability assumption~~

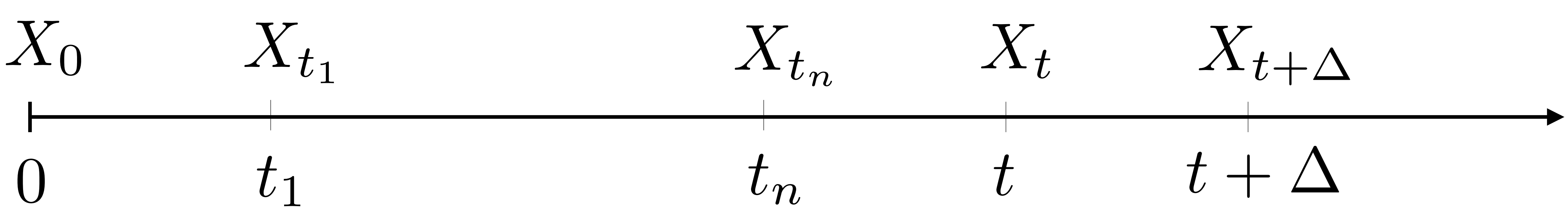


Imprecise continuous-time Markov chain

$$P(X_0 = x) = \pi_0(x)$$

$$\begin{aligned} P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x) \\ = P(X_{t+\Delta} = y | X_t = x) \\ \approx I(x,y) + \Delta \underset{\substack{\rotatebox{90}{$\in$} \\ \mathcal{Q}}}{Q_{t,\Delta}}(x,y) \end{aligned}$$

Markov assumption



Imprecise continuous-time Markov chain

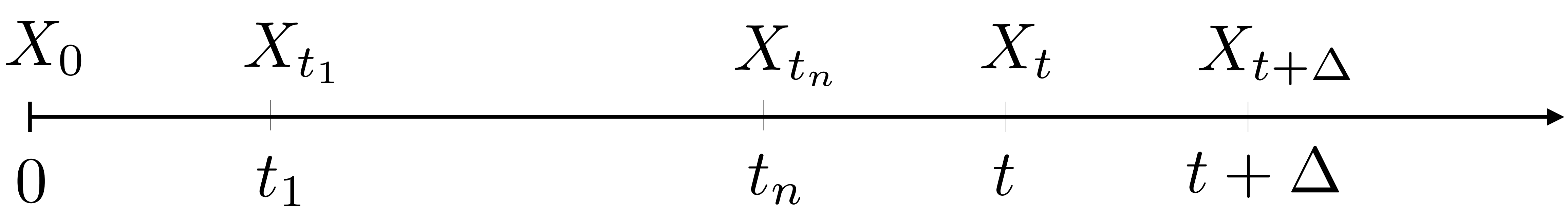
$$P(X_0 = x) = \pi_0(x)$$

$$P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x)$$

$$= P(X_{t+\Delta} = y | X_t = x)$$

~~Markov assumption~~

$$\approx I(x,y) + \Delta \underset{\mathcal{Q}}{\underset{\rotatebox{90}{$\in$}}{Q_{t,\Delta}}}(x,y)$$



Imprecise continuous-time Markov chain

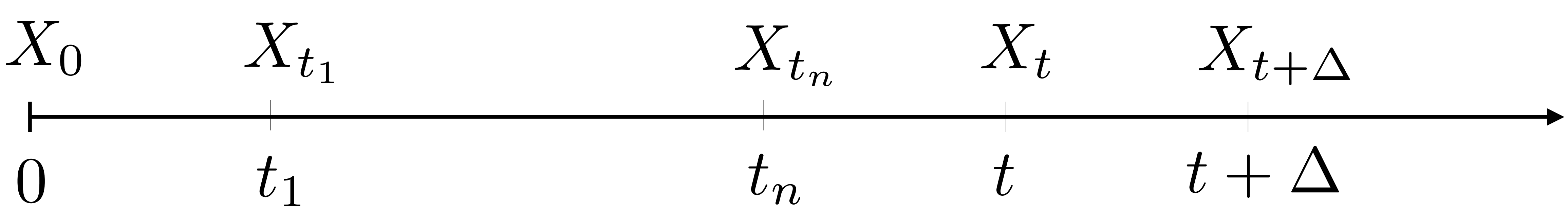
$$P(X_0 = x) = \pi_0(x)$$

$$P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x)$$

$$\approx I(x, y) + \Delta Q_{t,\Delta,x_1,\ldots,x_n}(x, y)$$

$$\in \mathcal{Q}$$

~~Markov assumption~~



Imprecise continuous-time Markov chain

$$P(X_0 = x) = \pi_0(x)$$

$$P(X_{t+\Delta} = y | X_{t_1} = x_{t_1}, \ldots, X_{t_n} = x_{t_n}, X_t = x)$$

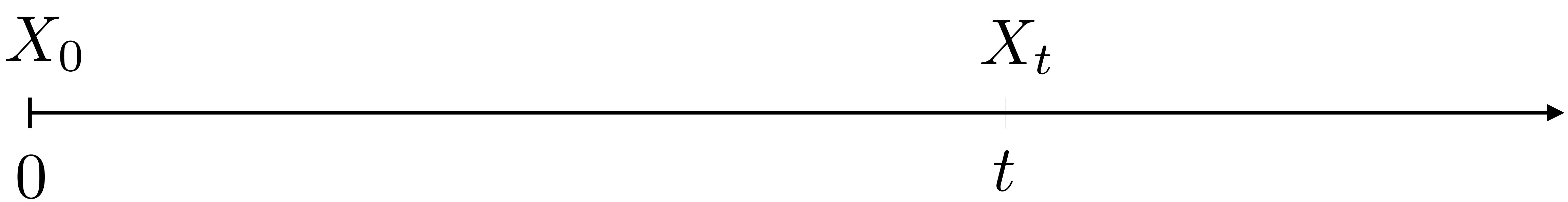
$$\approx I(x,y) + \Delta \underset{\underset{\mathcal{Q}}{\Cap}}{Q_{t,\Delta,x_1,\ldots,x_n}}(x,y)$$

$$\underline{E}(f) = \min_{P \sim \mathcal{Q}} E(f)$$

$$\overline{E}(f) = \max_{P \sim \mathcal{Q}} E(f) = -\underline{E}(-f)$$

probability bounds are again just special cases!

What is $\underline{E}(f(X_t)|X_0 = x)$?



What is $\underline{E}(f(X_t)|X_0 = x)$?

transition operator: $[\underline{T}_t(f)](x) = \underline{E}(f(X_t)|X_0 = x)$

backward Kolmogorov differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\underline{T}_t = \underline{Q}\underline{T}_t, \text{with } \underline{T}_0 = I$$

$$\Longrightarrow \quad \underline{T}_t = e^{\underline{Q}t} = \lim_{n \to +\infty} (I + \frac{t}{n}\underline{Q})^n$$

transition rate operator: $\underline{Q}f(x) = \min_{Q \in \mathcal{Q}} Qf(x)$

What is $\underline{E}(f(X_t)|X_0 = x)$? $\longrightarrow$ $\boxed{[e^{\underline{Q}t}(f)](x)}$

transition operator: $[\underline{T}_t(f)](x) = \underline{E}(f(X_t)|X_0 = x)$

backward Kolmogorov differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\underline{T}_t = \underline{Q}\,\underline{T}_t, \text{with } \underline{T}_0 = I$$

$$\Longrightarrow \quad \underline{T}_t = e^{\underline{Q}t} = \lim_{n \to +\infty}(I + \frac{t}{n}\underline{Q})^n$$

transition rate operator: $\underline{Q}f(x) = \min_{Q \in \mathcal{Q}} Qf(x)$

Imprecise continuous-time Markov chain ?

Imprecise continuous-time Markov chain ✓

$$\underline{E}(f(X_{t+\Delta}|X_{t_1}=x_{t_1},\ldots,X_{t_n}=x_{t_n},X_t=x)$$

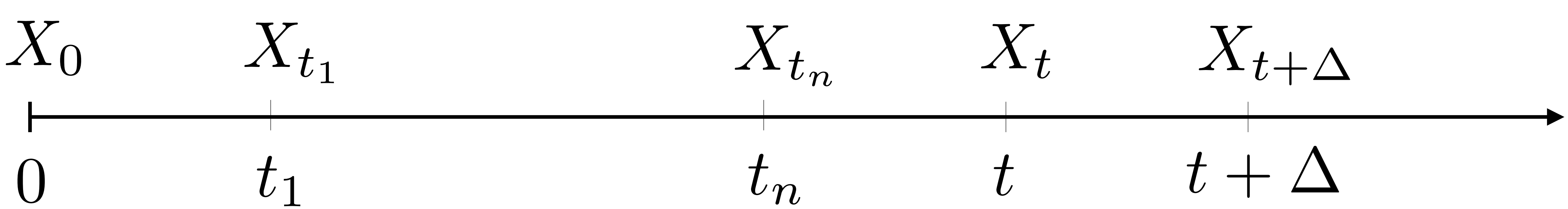
$$= \underline{E}(f(X_{t+\Delta}|X_t=x)$$

$$= \underline{E}(f(X_{\Delta}|X_0=x)$$

$$= [\underline{T}_{\Delta}(f)](x)$$

Markov property

time-homogeneity



Imprecise continuous-time Markov chain

$$\underline{E}(f(X_{t+\Delta}|X_{t_1} = x_{t_1}, \dots, X_{t_n} = x_{t_n}, X_t = x)$$

$$= \underline{E}(f(X_{t+\Delta}|X_t = x)$$

$$= \underline{E}(f(X_{\Delta}|X_0 = x)$$

$$= [\underline{T}_{\Delta}(f)](x)$$

Markov property

time-homogeneity

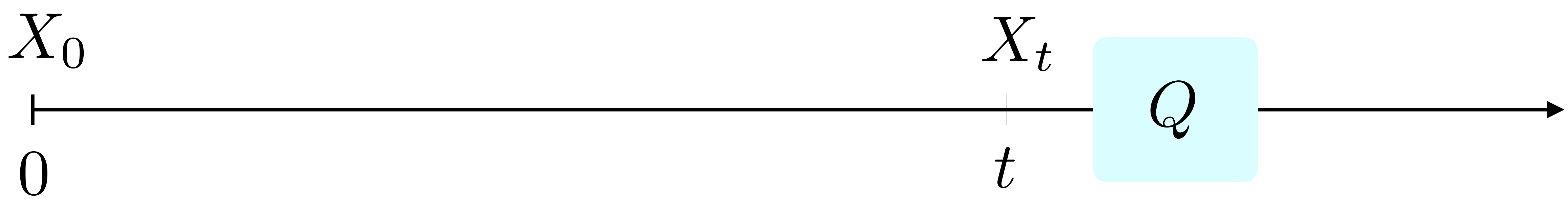
$$\underline{E}_{\infty}(f) = \lim_{t\to+\infty} \underline{E}(f(X_t)|X_0 = x) = \lim_{t\to+\infty} [\underline{T}_{\Delta}(f)](x)$$

What if I don’t have any imprecision?

What if I don't have any imprecision?

I know Q exactly!

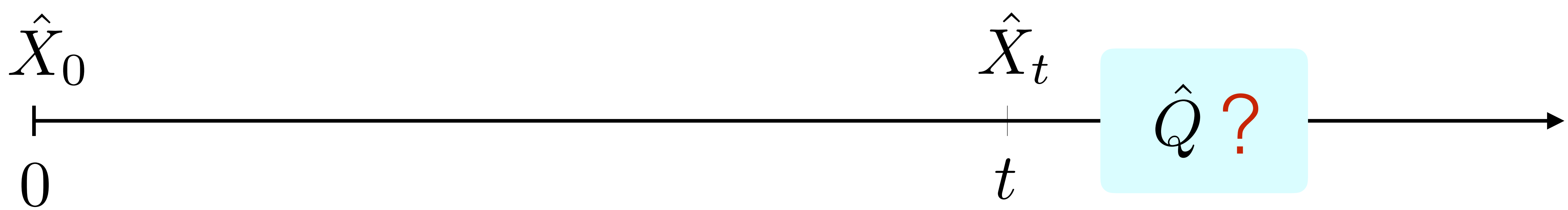
Solving the scaling problem



$X_t \in \mathcal{X}$ original Markov chain

$\Lambda\colon \mathcal{X} \to \hat{\mathcal{X}}$ lumping map

$\hat{X}_t = \Lambda(X_t) \in \hat{\mathcal{X}}$ lumped process



Amorous — GOOD

Bickering

Confusion — NOT GOOD

Depression

$$Q = \begin{array}{c} \\ \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{D} \end{array} \begin{array}{c} \begin{array}{cccc} \mathbf{A} & \mathbf{B} & \mathbf{C} & \mathbf{D} \end{array} \\ \begin{bmatrix} -4 & 3 & 1 & 0 \\ 4 & -6 & 2 & 0 \\ 2 & 3 & -6 & 1 \\ 0 & 0 & 2 & -2 \end{bmatrix} \end{array}$$

Solving the scaling problem

$$\hat{Q}_{t,x_{t_1},\ldots,x_{t_n}} \in \hat{\mathcal{Q}}$$

$\hat{X}_0$ $\hat{X}_{t_1}$ $\hat{X}_{t_n}$ $\hat{X}_t$

0 t_1 t_n t

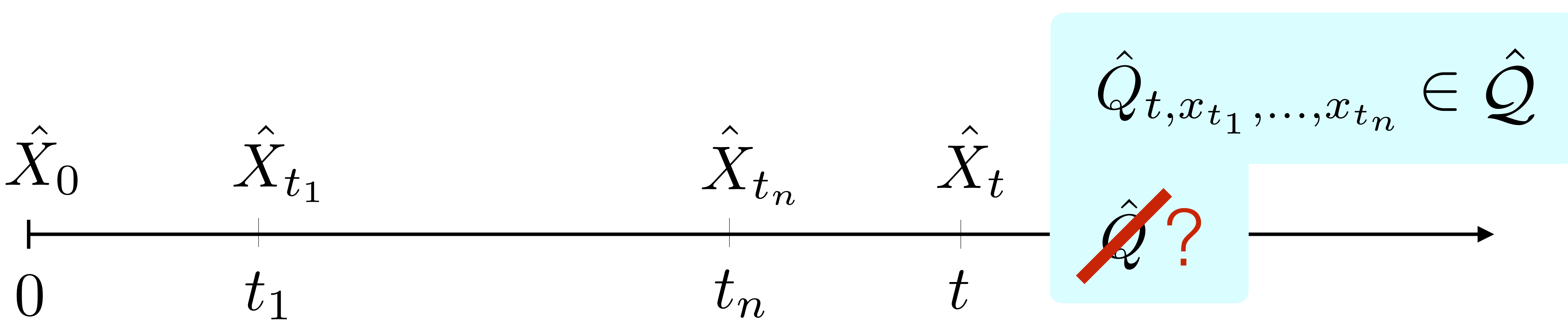
$\cancel{\hat{Q}}$?

Solving the scaling problem

$$E(\hat{f}(\hat{X}_t)|\hat{X}_0 = \hat{x}) \geq \left[e^{\underline{\hat{Q}}t}\hat{f}\right](\hat{x})$$

$$= \left[\lim_{n\to+\infty}(I + \frac{t}{n}\underline{\hat{Q}})^n \hat{f}\right](\hat{x})$$

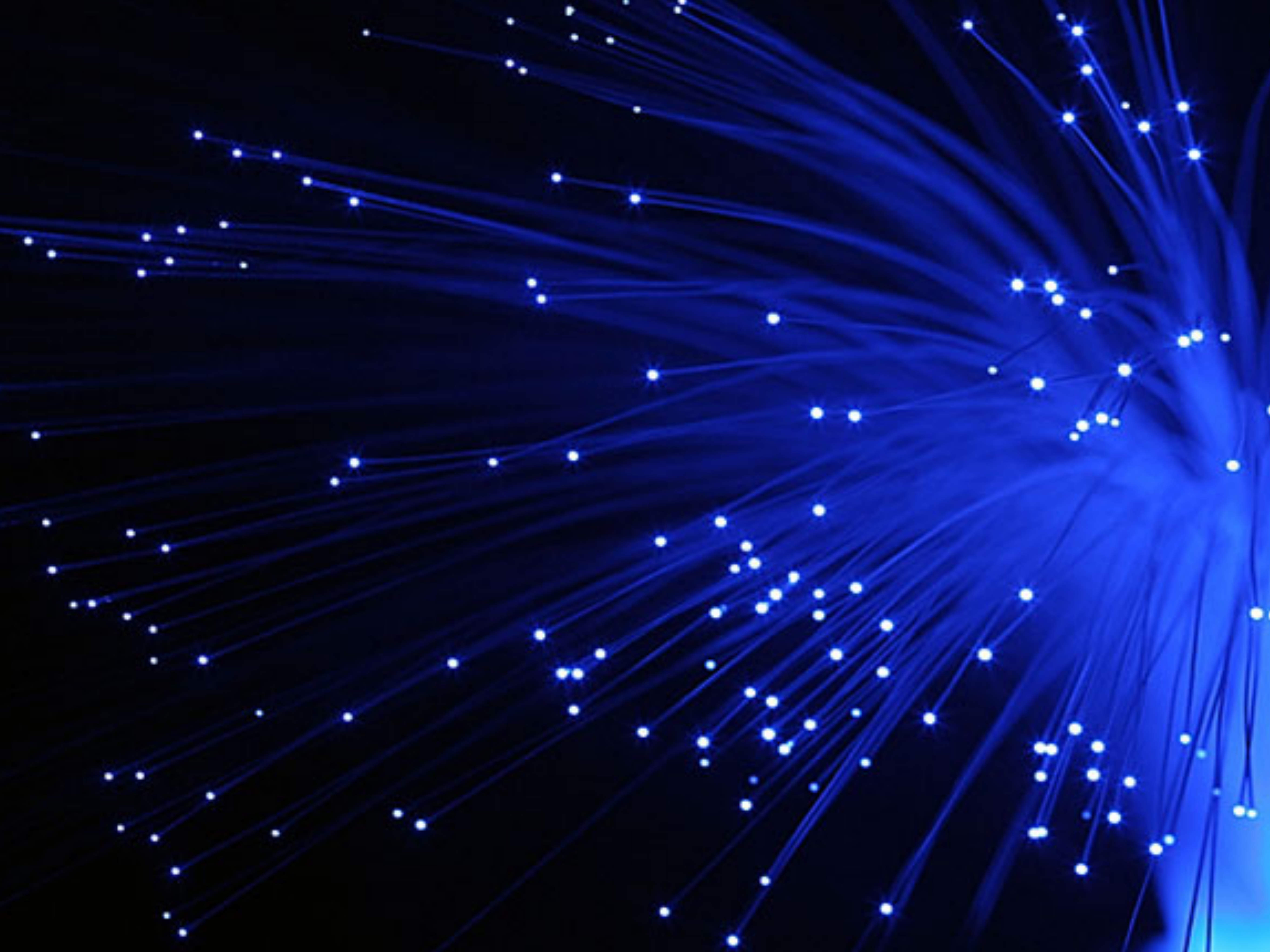
$$\left[\underline{\hat{Q}}\hat{f}\right](\hat{x}) = \min\left\{\sum_{\hat{y}\in\hat{\mathcal{X}}} \hat{f}(\hat{y}) \sum_{y\sim\hat{y}} Q(x,y) \colon x \sim \hat{x}\right\}$$



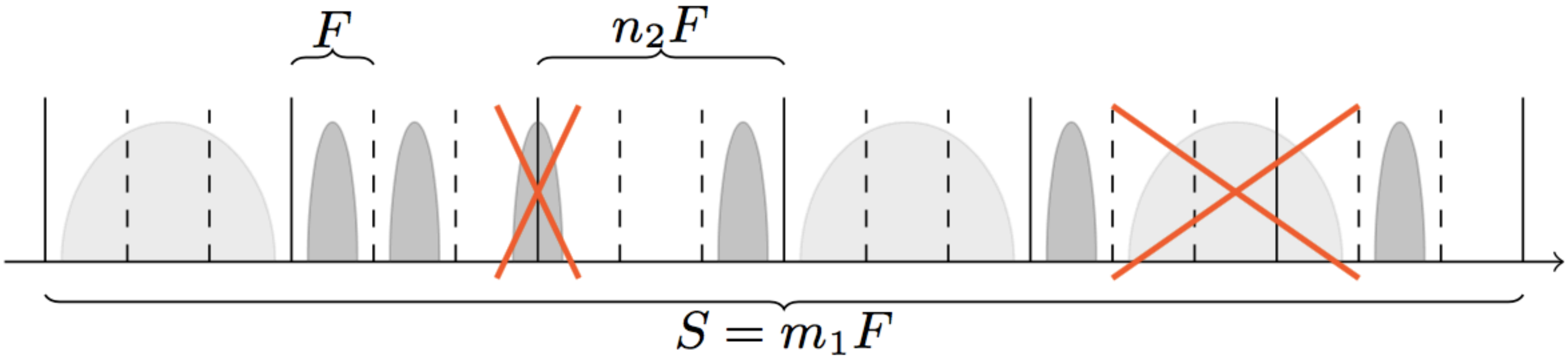
- Reliability engineering (failure probabilities, …)
- Queuing theory (waiting in line …)
  - optimising supermarket waiting times
  - dimensioning of call centers
  - airport security lines
  - router queues on the internet

- Chemical reactions (time-evolution …)

- Pagerank
- …



Message passing in optical links



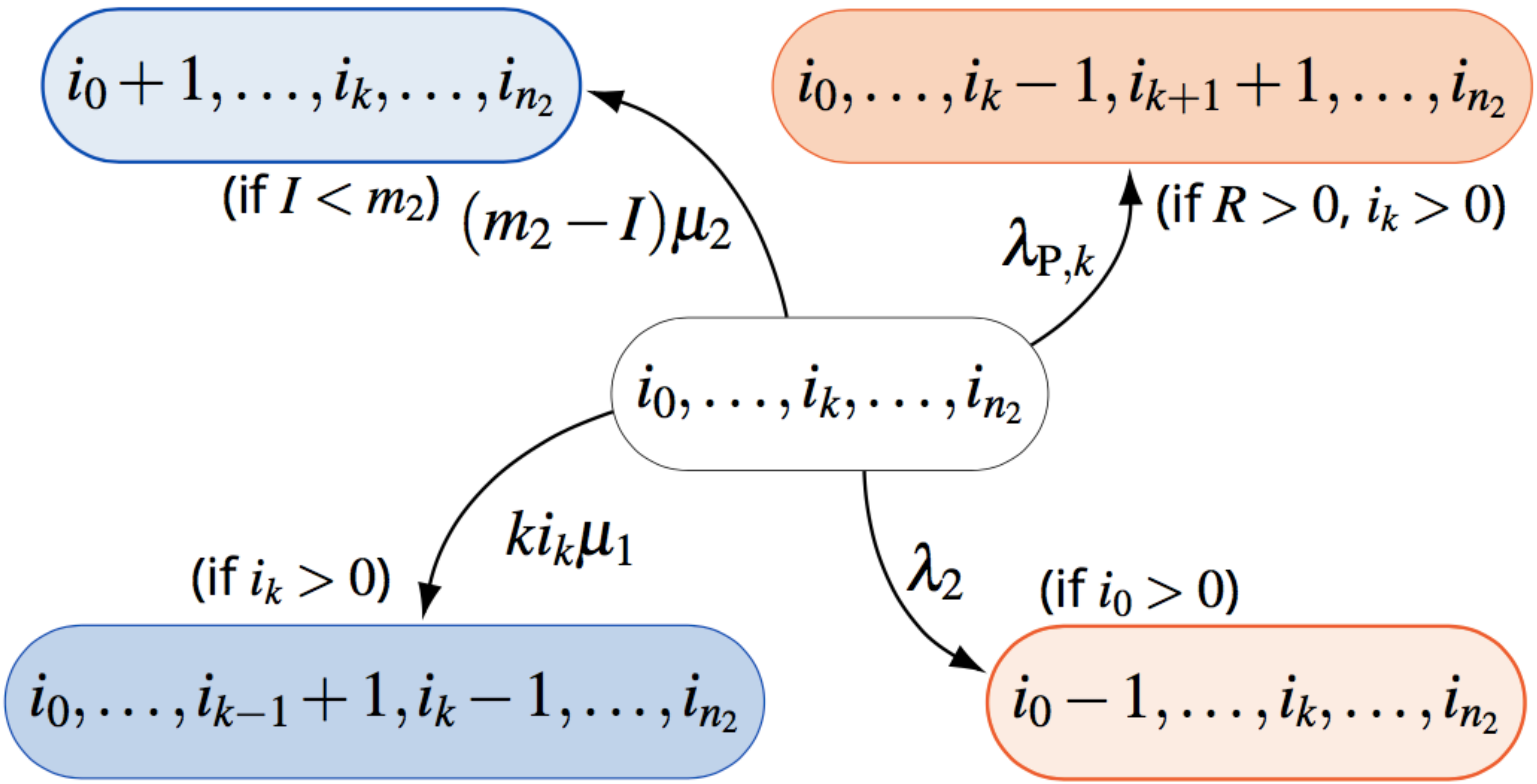
m_1 channels $\qquad$ $m_2 = \dfrac{m_1}{n_2}$ superchannels

type I messages require 1 channel

type II messages require 1 superchannel (n_2 channels)

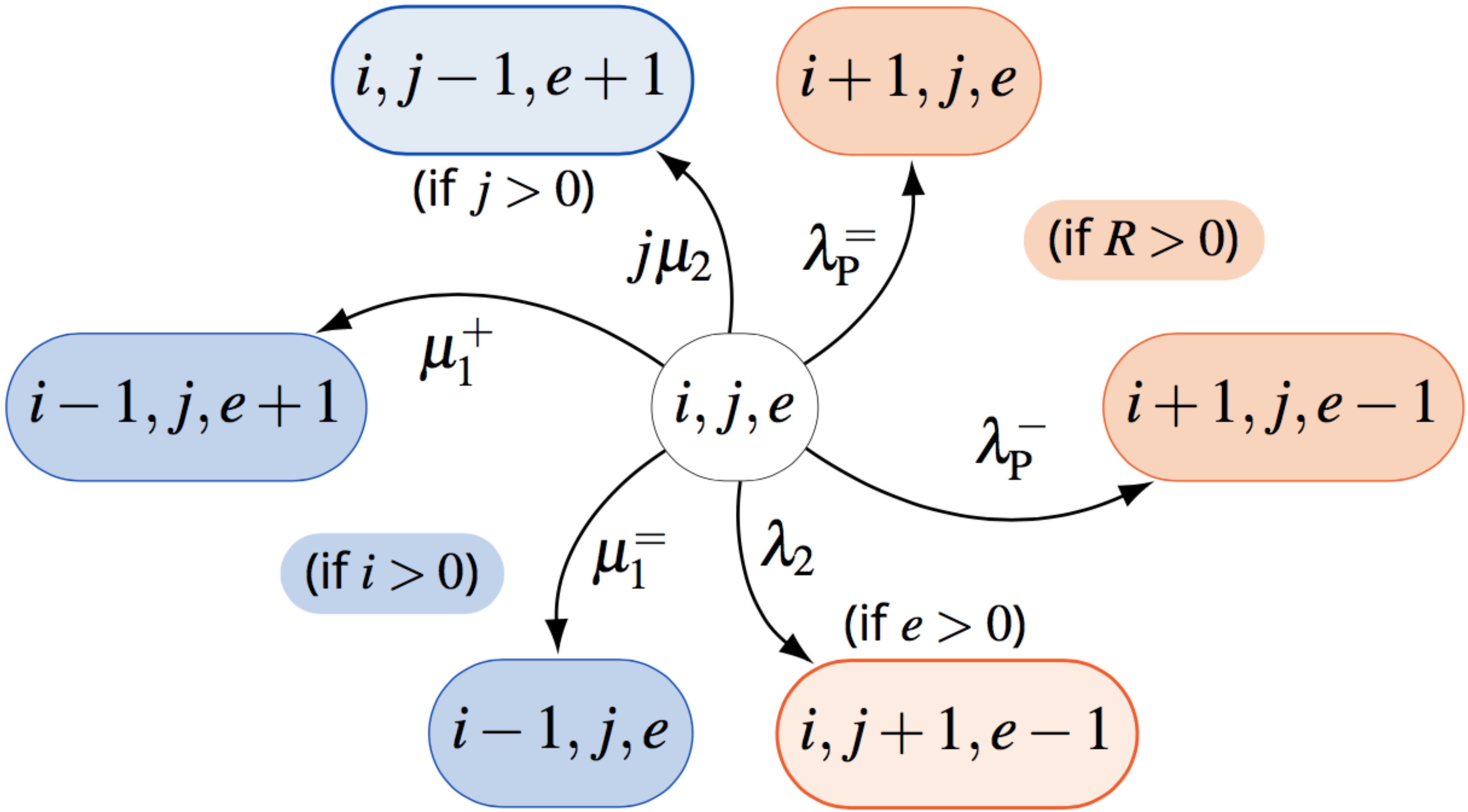
We want to know the blocking probability of messages for a given policy, and optimise it

$$\mathscr{X}_{\text{det}} := \left\{ (i_0, \ldots, i_{n_2}) \in \mathbb{N}^{(n_2+1)} : \sum_{k=0}^{n_2} i_k \le m_2 \right\}$$

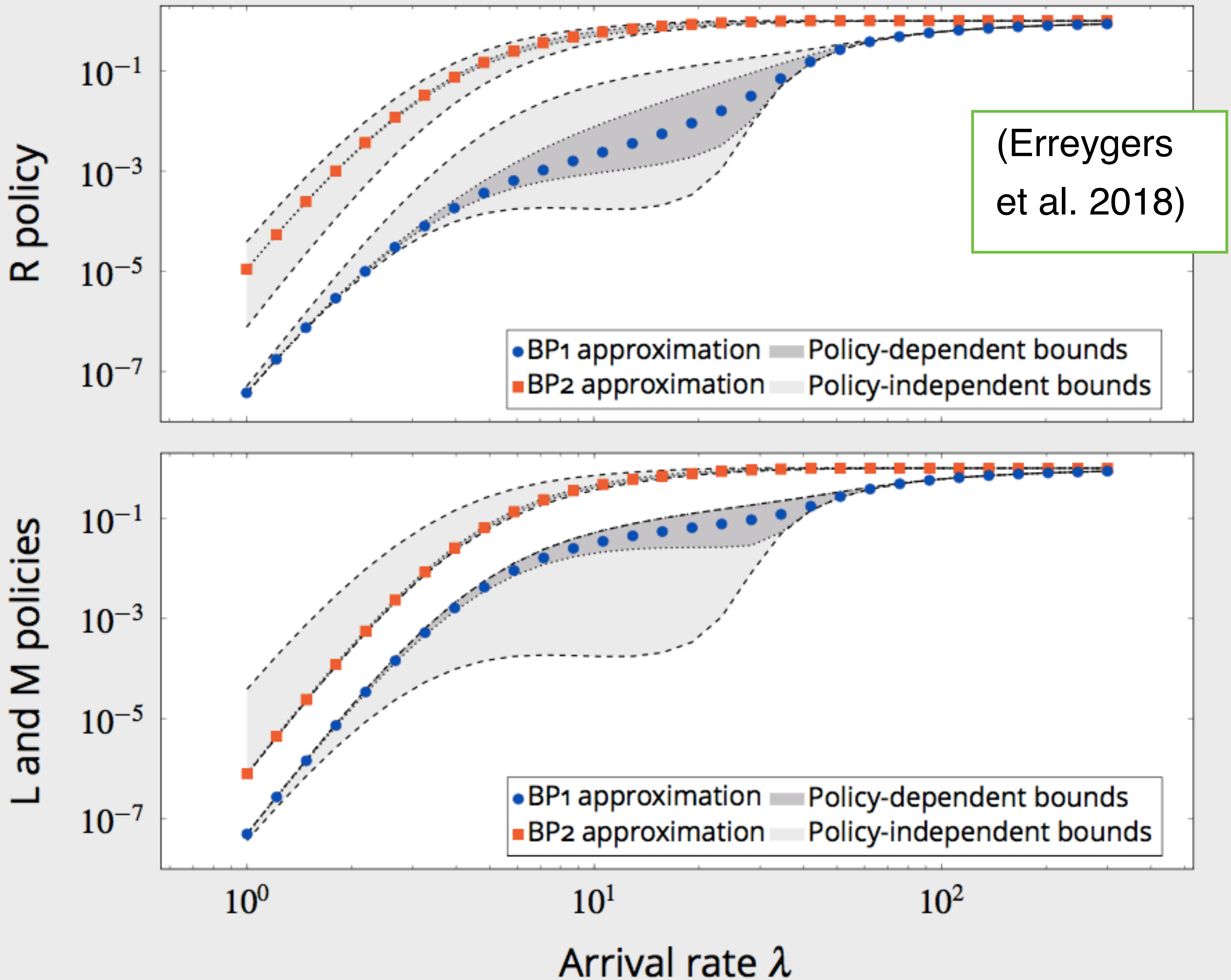


$I := \sum_{k=0}^{n_2} i_k \qquad R := \sum_{k=0}^{n_2-1} i_k (n_2 - k)$

$$\mathscr{X}_{\text{red}} := \{(i,j,e) \in \mathbb{N}^3 : m_2 \le i+j+e, i+(j+e)n_2 \le m_1\}$$



$$R := m_1 - i - jn_2$$



(Erreygers et al. 2018)
R policy
L and M policies
10^{-1}
10^{-3}
10^{-5}
10^{-7}
BP1 approximation
BP2 approximation
Policy-dependent bounds
Policy-independent bounds
10^0
10^1
10^2
Arrival rate λ

Advantages of imprecise Markov chains

- ✓ Partially specified Q (and π_0) are allowed
- ✓ Time-homogeneity can be relaxed
- ✓ The Markov assumption can be relaxed
- ✓ Efficient computations remain possible
- ✓ State space explosion can be dealt with