Bayesian Inference of Information Transfer in Networked Multi-Agent Systems

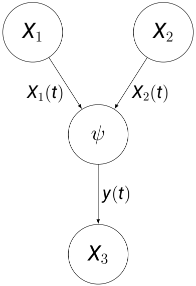
TECHNISCHE UNIVERSITÄT DARMSTADT

Master-Thesis

Presented by: Gizem Ekinci
Supervisors : Dominik Linzner
Anam Tahir

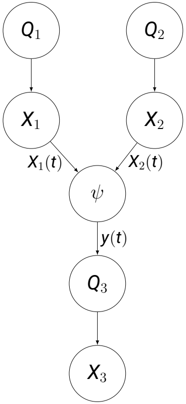
Introduction

- Problem: Agent making decisions based on incoming messages, but observes only a summary of them
- Objective: To infer this observation model from agent's behaviour
- Motivation: Cell to cell communication and cellular decision making [1]
  - A cell may emit some message based on its gene activation state. Consider two cells emitting messages, a third cell receiving a translation of these messages. Third cell may have evolved to achieve some task in coordination with other cells.
    e.g. the messages containing information about consumption of a particular molecule which needs to be partitioned for survival.
- Assuming that the behaviour of agent has been shaped by evolution (close) to optimality



Problem Formulation

Continuous-time Bayesian network (CTBN)



- X_1 and X_2 homogeneous continuous-time Markov processes with Q_1 and Q_2 transition intensity matrices

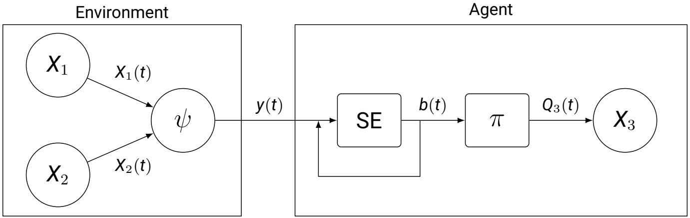
$$Q_n \sim \mathrm{Gam}(\boldsymbol{\alpha}_n, \boldsymbol{\beta}_n)$$

- X_3 conditional continuous-time Markov process with set of actions $a \in \{a_0, a_1\}$ and set of transition intensity matrices $\boldsymbol{Q}_3 = \left\{Q_{3|a_0}, Q_{3|a_1}\right\}$
- X_P: joint process of X_1 and X_2, with factorizing state space $\mathcal{X}_P = \mathcal{X}_1 \times \mathcal{X}_2$
- Observation model
  - $\psi(x_P) = p(y(t) \mid X_P(t) = x_P)$
  - ψ denoting the matrix with rows $\{\psi(x_P)\}_{x_P \in \mathcal{X}_P}$
- $\mathcal{X}_n = \{0, 1\}$
- $X_n^{[0,T]}$: discrete valued trajectory in time interval $[0, T]$

Problem Formulation

Partially observable Markov decision process (POMDP)

TECHNISCHE UNIVERSITÄT DARMSTADT



- Belief state
  - $b(x_P;t) = \Pr(X_P(t) = x_P \mid y_1, \ldots, y_t)$
  - $b(t)$ denoting the row vector with $\{b(x_P;t)\}_{x_P \in \mathcal{X}_P}$

- Optimal policy of the agent
  - $\pi(b(t)) = a(t) = \begin{cases} a_0 & \text{if } wb(t)^\intercal > 0.5 \\ a_1 & \text{otherwise} \end{cases}$
  - $Q_3(t) = \begin{cases} Q_{3|a_0} & \text{if } a(t) = a_0 \\ Q_{3|a_1} & \text{otherwise} \end{cases}$

BCS

Exact Belief State Update
Filtering CTMPs

- Continuous-time solution of belief state through filtering for CTMPs, used as a baseline
- The posterior probability can be described by a system of ODEs

$$\frac{db(t)}{dt} = b(t)\, Q_P \tag{1}$$

with the solution,

$$b(t) = b(0) \exp(tQ_P) \tag{2}$$

where the initial condition $b(0)$ is row vector with $\{b(x_P; t = 0)\}_{x_P \in \mathcal{X}_P}$ [2].

- Q_P is the joint transition intensity matrix of X_1 and X_2 and given by amalgamation operation between Q_1 and Q_2 [3].

$$Q_P = Q_1 * Q_2 \tag{3}$$

Exact Belief State Update

Filtering CTMPs

- The belief update at discrete times of observation $y_L = y(t_L)$ can be obtained as

$$
\begin{aligned}
b(x_P; t_L) &= \Pr(X_P(t_L) = x_P, \mid y_1, ..., y_L) \\
&= \frac{\Pr(y_1, ..., y_L, X_P(t_L) = x_P)}{\Pr(y_1, ..., y_L)} \\
&= \frac{\Pr(y_L \mid y_1, ..., y_{L-1}, X_P(t_L) = x_P)}{\Pr(y_L \mid y_1, ..., y_{L-1})} \frac{\Pr(y_1, ..., y_{L-1}, X_P(t_L) = x_P)}{\Pr(y_1, ..., y_{L-1})} \\
&= Z_L^{-1} \; \Pr(y_L \mid X_P(t_L) = x_P) \; \Pr(X_P(t_L) = x_P \mid y_1, ..., y_{L-1}) \\
&= Z_L^{-1} \, p(y_L \mid x_P) \, b(x_P; t_L^-)
\end{aligned}
\tag{4}
$$

where $Z_L = \sum_{x_P \in \mathcal{X}_P} p(y_L \mid x_P) \, b(x_P; t_L^-)$ is the normalization factor [2].

Belief State Update Using Particle Filter

- The assumption that the complete information of parent dynamics is available is unrealistic.
- The agent may rather have some prior beliefs over them.
- With exact update method, these parameters are assumed to be available for the inference as well.
- To simulate a more realistic model and be able to marginalize out these parameters from inference problem
  - Replacing the exact belief update with marginal particle filter approximation

Transition Intensity Marginalization

over Q_1 and Q_2

TECHNISCHE UNIVERSITÄT DARMSTADT

- Given the priors and sample trajectory, the intensity matrices can be replaced by estimates. [4]
  - Q_n with non-diagonal entries $q^n_{i,j} \sim \mathrm{Gam}(\alpha^n_{i,j}, \beta^n_{i,j}), \;\; n \in \{1, 2\}$
  - $X_n^{[0,T]}$ with summary statistics $\Upsilon_n(x_i)$ and $r_n(x_i, x_j)$, where $\Upsilon_n(x_i)$ is the total time spent in state x_i, $r_n(x_i, x_j)$ is the number of transitions from state x_i to state x_j

$$p(X_n^{[0,T]} \mid Q) = \left(\prod_i q_i^{r_n(x_i)} \exp\left(-q_i \Upsilon_n(x_i)\right)\right) \left(\prod_i \prod_{j \neq i} \left(\frac{q_{i,j}}{q_i}\right)^{r_n(x_i,x_j)}\right)$$
$$= \prod_{j \neq i} \exp(-q_{i,j} \Upsilon_n(x_i))\, q_{i,j}^{r_n(x_i,x_j)} \tag{5}$$

  - Using Bayes' rule and the likelihood of trajectory in Eq.5, the estimates can be evaluated analytically as follows:

$$\mathbb{E}\left[q^n_{i,j} | X_n^{[0,T]}\right] = \frac{\alpha^n_{i,j} + r_n(x_i, x_j)}{\beta^n_{i,j} + \Upsilon_n(x_i)} \tag{6}$$

Marginal Particle Filter

TECHNISCHE UNIVERSITÄT DARMSTADT

- The particles to represent the belief state are drawn from a marginalized counterparts of parent processes [4].
- With every new observation, the particles are propagated through the marginal process.
  - The processing of the particles are done one after another.
  - After each particle, the summary statistics are updated and the parameters are re-estimated using the Eq.6.
  - The belief state is obtained as the distribution of states over the particles,

$$b(x_P; t) = \frac{1}{M} \sum_{m=1}^{M} \delta_{k_m(t), x_P} \tag{7}$$

where M is the number of particles, $k_i \in \mathbf{k}$ is the set of particles, and δ is the Kronecker delta.

Likelihood Model of the System

TECHNISCHE UNIVERSITÄT DARMSTADT

- Let $S^{[0,T]}$ be a sample of trajectories in the dataset, such that $S^{[0,T]} = \left\{X_1^{[0,T]}, X_2^{[0,T]}, X_3^{[0,T]}\right\}$, and the set of parameters to the system $\theta = \{Q_1, Q_2, \boldsymbol{Q}_3, \pi, \psi\}$.
- The likelihood of the sample $S^{[0,T]}$ can be written as

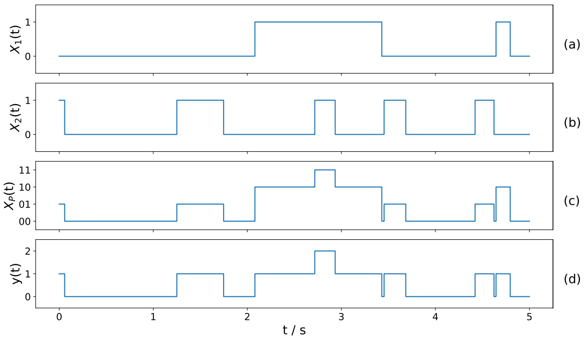
$$\begin{aligned} p(S^{[0,T]} \mid \theta) &= p(X_1^{[0,T]}, X_2^{[0,T]}, X_3^{[0,T]} \mid Q_1, Q_2, \boldsymbol{Q}_3, \pi, \psi) \\ &= p(X_3^{[0,T]} \mid X_1^{[0,T]}, X_2^{[0,T]}, \boldsymbol{Q}_3, \pi, \psi)\, p(X_1^{[0,T]} \mid Q_1)\, p(X_2^{[0,T]} \mid Q_2) \\ &= p(X_3^{[0,T]} \mid Q_3^{[0,T]})\, p(X_1^{[0,T]} \mid Q_1)\, p(X_2^{[0,T]} \mid Q_2) \end{aligned} \tag{8}$$

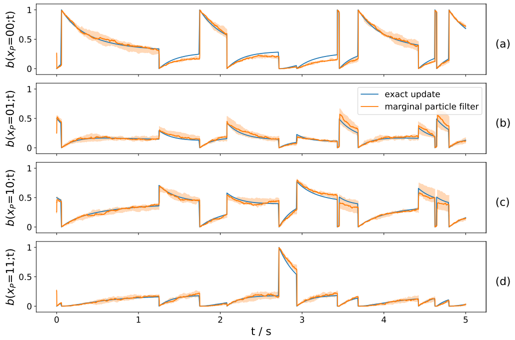
- Marginalized over Q_1 and Q_2

$$\begin{aligned} p(S^{[0,T]} \mid \pi, \Phi) = p(X_3^{[0,T]} \mid Q_3^{[0,T]}) &\prod_{x_1 \in \mathcal{X}_1} \frac{\beta_{x_1}^{\alpha_{x_1}}}{\Gamma(\alpha_{x_1})} \left(\Upsilon(x_1) + \beta_{x_1}\right)^{-r(x_1)-\alpha_{x_1}} \Gamma(r(x_1) + \alpha_{x_1}) \\ &\prod_{x_2 \in \mathcal{X}_2} \frac{\beta_{x_2}^{\alpha_{x_2}}}{\Gamma(\alpha_{x_2})} \left(\Upsilon(x_2) + \beta_{x_2}\right)^{-r(x_2)-\alpha_{x_2}} \Gamma(r(x_2) + \alpha_{x_2}) \end{aligned} \tag{9}$$

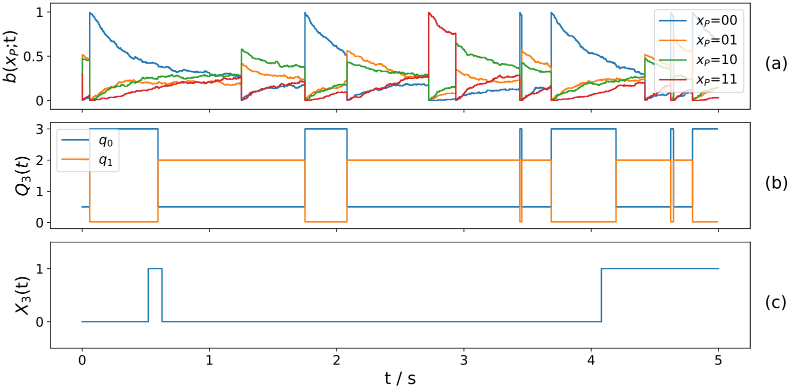
Simulation

Sampling trajectories using Gillespie algorithm



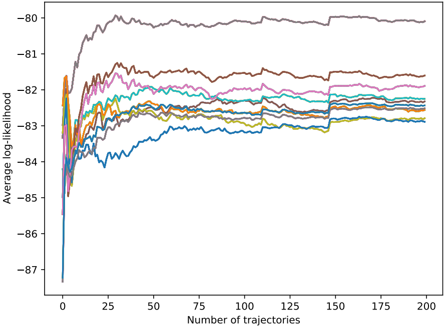


Simulation
Continued



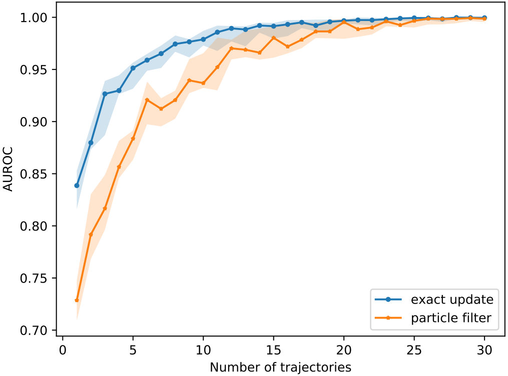
Limitation of Equivalence Classes

- Although we have 81 observation models, we can separate them to only 10 classes.
- This is caused by either identical effect on the belief state or on the behavior.



Results

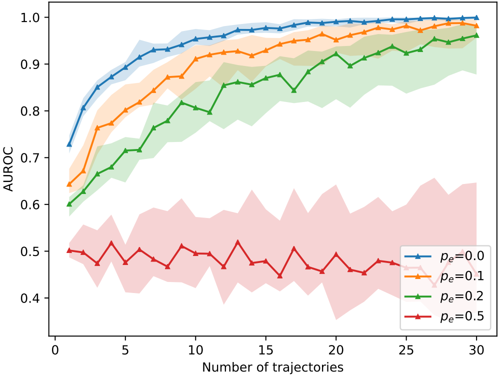
Area under Receiver Operating Characteristic curve (AUROC)



- For the classification problem, we obtain the posterior probabilities from the estimated likelihoods and uniform prior.
- We train the classifiers with a given number of trajectories, through bootstrapping.
- With increasing number of trajectories, AUROC approaches to 1.

Results under Noise

Area under Receiver Operating Characteristic curve (AUROC)



- p_e denotes the probability of producing erroneous observation.
- Noisy observation model can be interpreted as a noisy communication channel with an error probability of p_e.
- The noise parameter is assumed to be available to the agent, i.e. it is not estimated.

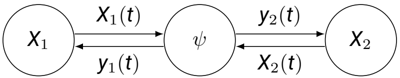
Conclusion

TECHNISCHE UNIVERSITÄT DARMSTADT

- A realistic system is achieved using particle filtering with marginalized CTBN. Given Gamma-priors of Q_1 and Q_2, the exact update method is well approximated by the marginal particle filter.
- In classification, the marginal particle filter yields a slightly lower performance. Nevertheless, in both methods, as the number of samples increases, the metric approaches to 1.
- The performance decreases as the noise introduced to the true observation model increases. With the increasing number of trajectories, the metric converges to 1, showing robustness.
- The main limitation is equivalence classes.

Outlook

- Eliminate the equivalence classes
  - Joint inference of observation model and policy
- Application of the model and solution approach to a more complex environment to evaluate the performance further
  - Non-binary messages, more than two parent nodes etc.
- Employing the method in different environments to get insights into the interactions of agents and environments
  - Inferring the communication protocols that lead to the success or failure of the agents in Foerster's multi-step MNIST game [6]
- Inference of observation model in an interactive multi-agent system



References

[1] T. J. Perkins and P. S. Swain, “Strategies for cellular decision-making,” *Molecular systems biology*, vol. 5, no. 1, p. 326, 2009.

[2] L. Huang, L. Pauleve, C. Zechner, M. Unger, A. S. Hansen, and H. Koeppl, “Supporting information for reconstructing dynamic molecular states from single-cell time series,” Aug 2016.

[3] U. Nodelman, C. R. Shelton, and D. Koller, “Continuous time Bayesian networks,” in *Proceedings of the 18th Conference in Uncertainty in Artificial Intelligence*, pp. 378–387, 2002.

[4] L. Studer, L. Paulevé, C. Zechner, M. Reumann, M. R. Martínez, and H. Koeppl, “Marginalized continuous time Bayesian networks for network reconstruction from incomplete observations,” in *Proceedings of the 30th Conference on ArtificialIntelligence (AAAI 2016)*, pp. 2051–2057, 2016.

[5] U. Nodelman, C. R. Shelton, and D. Koller, “Learning continuous time Bayesian networks,” in *Proceedings of the 19th Conference in Uncertainty in Artificial Intelligence*, pp. 451–458, 2003.

[6] J. Foerster, I. A. Assael, N. De Freitas, and S. Whiteson, “Learning to communicate with deep multi-agent reinforcement learning,” in *Advances in neural information processing systems*, pp. 2137–2145, 2016.

Thank you!

Parameters
Parameters

- Intensity matrices

$$Q_1 = \begin{bmatrix} -1.117 & 1.117 \\ 0.836 & -0.836 \end{bmatrix}$$

$$Q_2 = \begin{bmatrix} -1.1 & 1.1 \\ 2.445 & -2.445 \end{bmatrix}$$

$$\boldsymbol{Q}_3 = \left\{ \begin{bmatrix} -0.5 & 0.5 \\ 2 & -2 \end{bmatrix}, \begin{bmatrix} -3 & 3 \\ 0.02 & -0.02 \end{bmatrix} \right\}$$

- Weights of the policy

$$w = \begin{bmatrix} 0.02 & 0.833 & 0.778 & 0.87 \end{bmatrix}$$

- Gamma priors for parent dynamics such that $Q_n \sim \mathrm{Gam}(\boldsymbol{\alpha}^n, \boldsymbol{\beta}^n)$ for $n \in \{1, 2\}$, and $\boldsymbol{\alpha}^n = [\alpha_0^n, \alpha_1^n]$ and $\boldsymbol{\beta}^n = [\beta_0^n, \beta_1^n]$

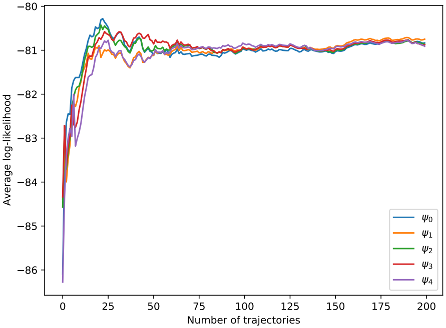
$$\boldsymbol{\alpha}^1 = [5, 10] \quad \boldsymbol{\beta}^1 = [5, 20]$$
$$\boldsymbol{\alpha}^2 = [10, 10] \quad \boldsymbol{\beta}^2 = [10, 5]$$

- Number of particles $M = 200$
- True observation model for the results

$$\psi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Equivalence Classes

- Marginal particle filtering over 5 observation model from same equivalence class

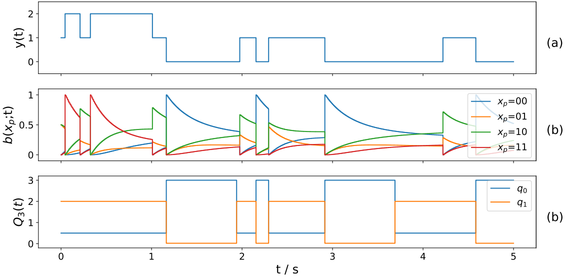


Equivalence Classes

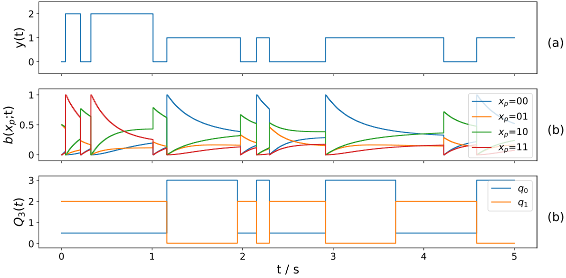
Identical Effect on Belief State

TECHNISCHE UNIVERSITÄT DARMSTADT

$$\psi_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



$$\psi_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

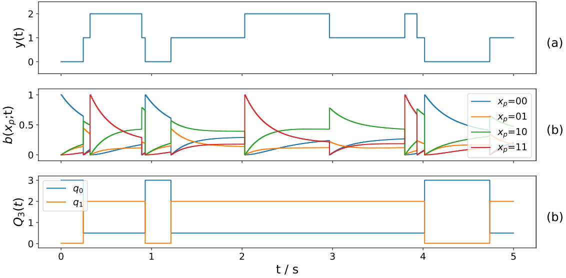


Equivalence Classes

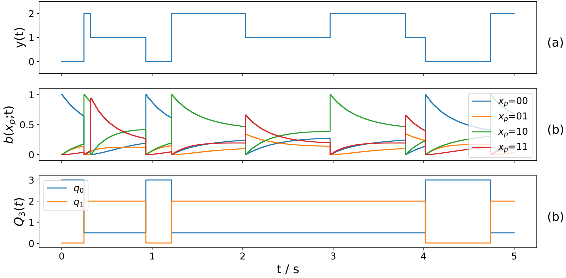
Identical Effect on Behaviour

TECHNISCHE UNIVERSITÄT DARMSTADT

$$\psi_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



$$\psi_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$



Marginal Particle Filter

Algorithm 1: Marginal particle filter for belief state update [4]

Input : Observation y_l at time t_l, set of particles $\mathbf{k}^{l-1}$, estimated $\hat{Q}$
Output: New set of particles $\mathbf{k}^l$, $\mathbf{b}^{[t_{l-1},t_l]}$

for $k_m \in \mathbf{k}^{l-1}$ **do**
  $k_m = \left\{x_m, \hat{Q}\right\} \leftarrow$ *Propagate particle through marginal process from* t_{l-1} *to* t_l
  $\hat{Q} \leftarrow$ *sufficient statistics added from* $k_m[t_{l-1}, t_l]$
  `// observation likelihood assigned as particle weight`
  $w_m \leftarrow p(y_l \mid X_P(t_l) = x_m)$
end for
`// belief state from` t_{l-1} `to` t_l
$\mathbf{b}^{[t_{l-1},t_l]} \leftarrow \left\{\frac{1}{M}\sum_{m=1}^{M}\delta_{k_m^{[t_{l-1},t_l]},x_P}\right\}_{x_P\in\mathcal{X}_P}$
`// normalize weights`
$w_m \leftarrow \frac{w_m}{\sum_m w_m}$
`// resample particles`
for $k_m \in \mathbf{k}^l$ **do**
  $k_m \leftarrow$ *Sample from* $\mathbf{k}^l$ *with probabilities* w_m *with replacement*
end for

Markov Processes

- Consider a homogenous Markov process X with values $\chi = \{x_0, x_1, ..., x_n\}$. The transition intensity matrix Q of such process has the following form:

$$Q = \begin{bmatrix} -q_{x_0} & q_{x_0x_1} & \cdots & q_{x_0x_n} \\ q_{x_1x_0} & -q_{x_1} & \cdots & q_{x_1x_n} \\ \vdots & \vdots & \ddots & \cdots \\ q_{x_nx_0} & q_{x_nx_1} & \cdots & -q_{x_n} \end{bmatrix} \tag{10}$$

  where $q_x = \sum_{x' \neq x, x' \in \chi} q_{xx'}$.
- The amount of time that X stays in a transitioned state is exponentially distributed, and the probability density function of staying in a state x is given by [3]

$$f(t) = q_x \, exp(-q_x t). \tag{11}$$

Likelihood Functions

Homogenous continuous-time Markov process

- The likelihood of a single transition $d = \langle x, t, x' \rangle$, where transition happens from x to x' after spending time amount of time t:

$$P(d \mid Q) = (q_x\, exp(-q_x t)) \left(\frac{q_{xx'}}{q_x} \right) \tag{12}$$

- The likelihood of trajectory $X^{[0,T]}$ can be decomposed as a product of likelihood of single transitions.

$$\begin{aligned} P(X^{[0,T]} \mid Q) &= \left(\prod_x q_x^{M[x]}\, exp(-q_x T[x]) \right) \left(\prod_x \prod_{x' \neq x} \frac{q_{xx'}}{q_x}^{M[x,x']} \right) \\ &= \prod_x exp(-q_x T[x]) \prod_{x' \neq x} q_{xx'}^{M[x,x']} \end{aligned} \tag{13}$$

where $T[x]$ is the total time spent in state x, $M[x, x']$ is the number of transitions from state x to state x', $M[x]$ is total number of transitions leaving state x [5].