

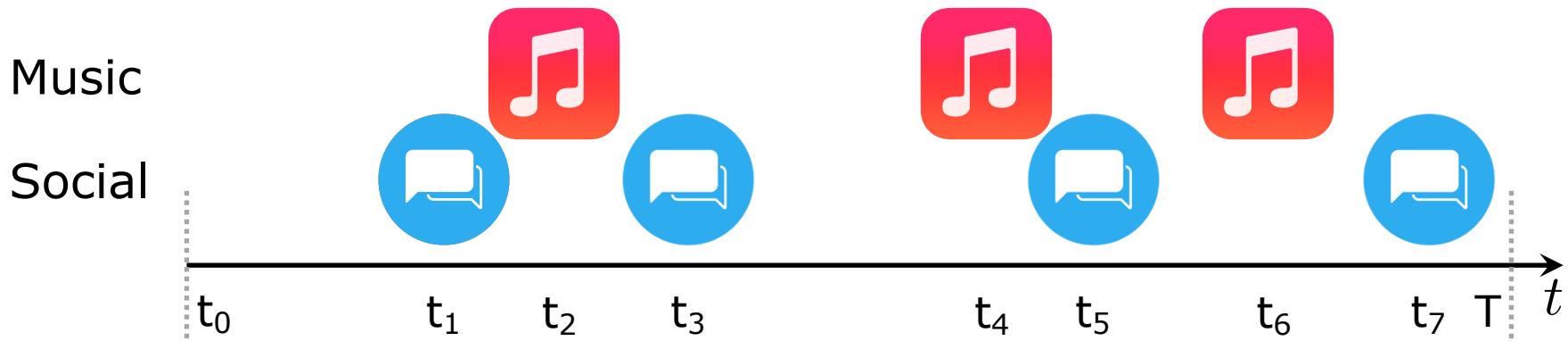
Modeling Irregular Time Series

Hongyuan Mei

2018–2021 Bloomberg Data Science Ph.D. Fellow

Johns Hopkins University

Multivariate Time Series



$\propto \lambda_{\text{music}}$



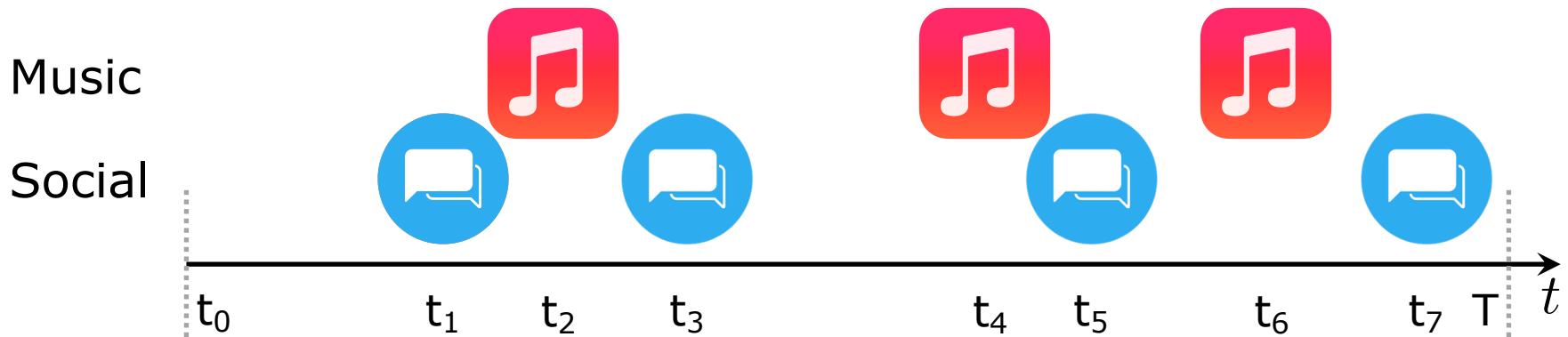
$\propto \lambda_{\text{social}}$

$(\lambda_{\text{music}} + \lambda_{\text{social}})dt = \text{prob of some event at } t$

λ_{social}

λ_{music}

Multivariate Time Series: MLE



$$\begin{aligned} & \lambda_{\text{social}} \exp(-\lambda (t_1 - t_0)) \rightarrow (1-\lambda dt) \dots (1-\lambda dt) \\ & \times \lambda_{\text{music}} \exp(-\lambda (t_2 - t_1)) \quad \lambda = \lambda_{\text{music}} + \lambda_{\text{social}} \\ & \times \lambda_{\text{social}} \exp(-\lambda (t_3 - t_2)) \\ & \dots \dots \\ & = \lambda_{\text{social}}^4 \lambda_{\text{music}}^3 \exp(-\lambda (T - t_0)) \end{aligned}$$

$\lambda_{\text{social}} = 4 / (T - t_0)$

$\lambda_{\text{music}} = 3 / (T - t_0)$

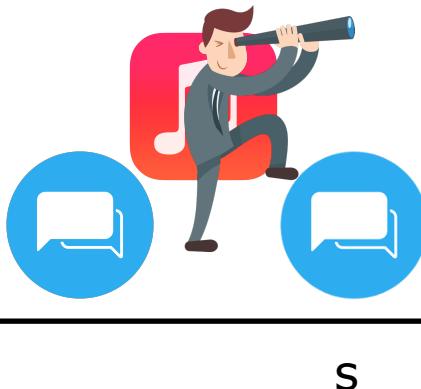
of events
of *that* type

Multivariate Time Series: MBR

Music

Social

time



$$\lambda_{\text{social}} = 4 / (T - t_0)$$

$$\lambda_{\text{music}} = 3 / (T - t_0)$$

$$t_{\text{pred}} = \int_s^\infty t p(t) dt \quad p(t) = \lambda \exp(-\lambda (t - s))$$

$$\lambda = \lambda_{\text{music}} + \lambda_{\text{social}} = 7 / (T - t_0)$$

type?

$$k = \operatorname{argmax}_k p(k | t) = \operatorname{argmax}_k \lambda_k$$

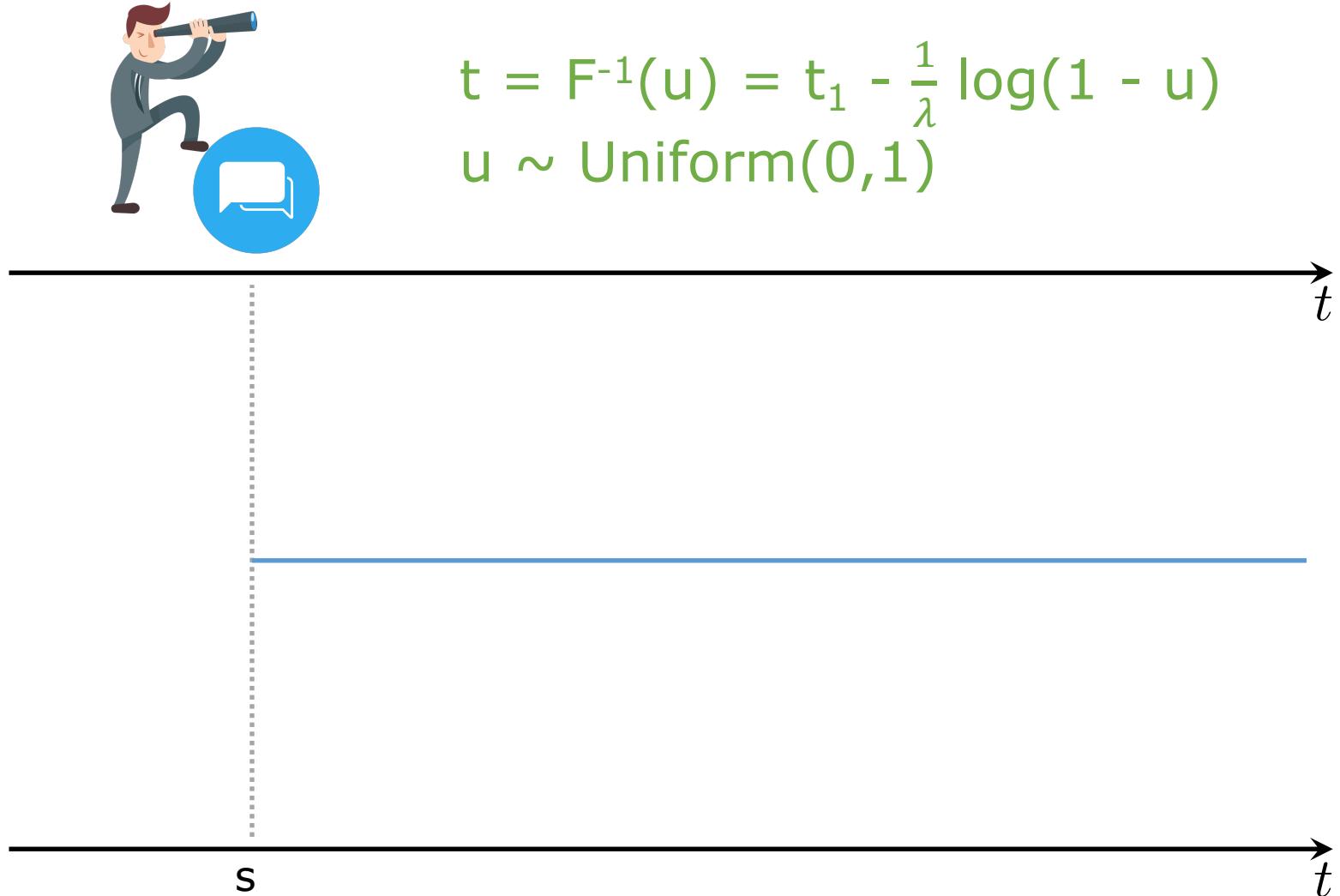
$$\operatorname{argmin}_k \sum_j p(j | t) \mathbb{I}(k \neq j) \quad j, k \in \{\text{music}, \text{social}\}$$

choose type k that makes $\sum_{j \neq k} p(j | t)$ small

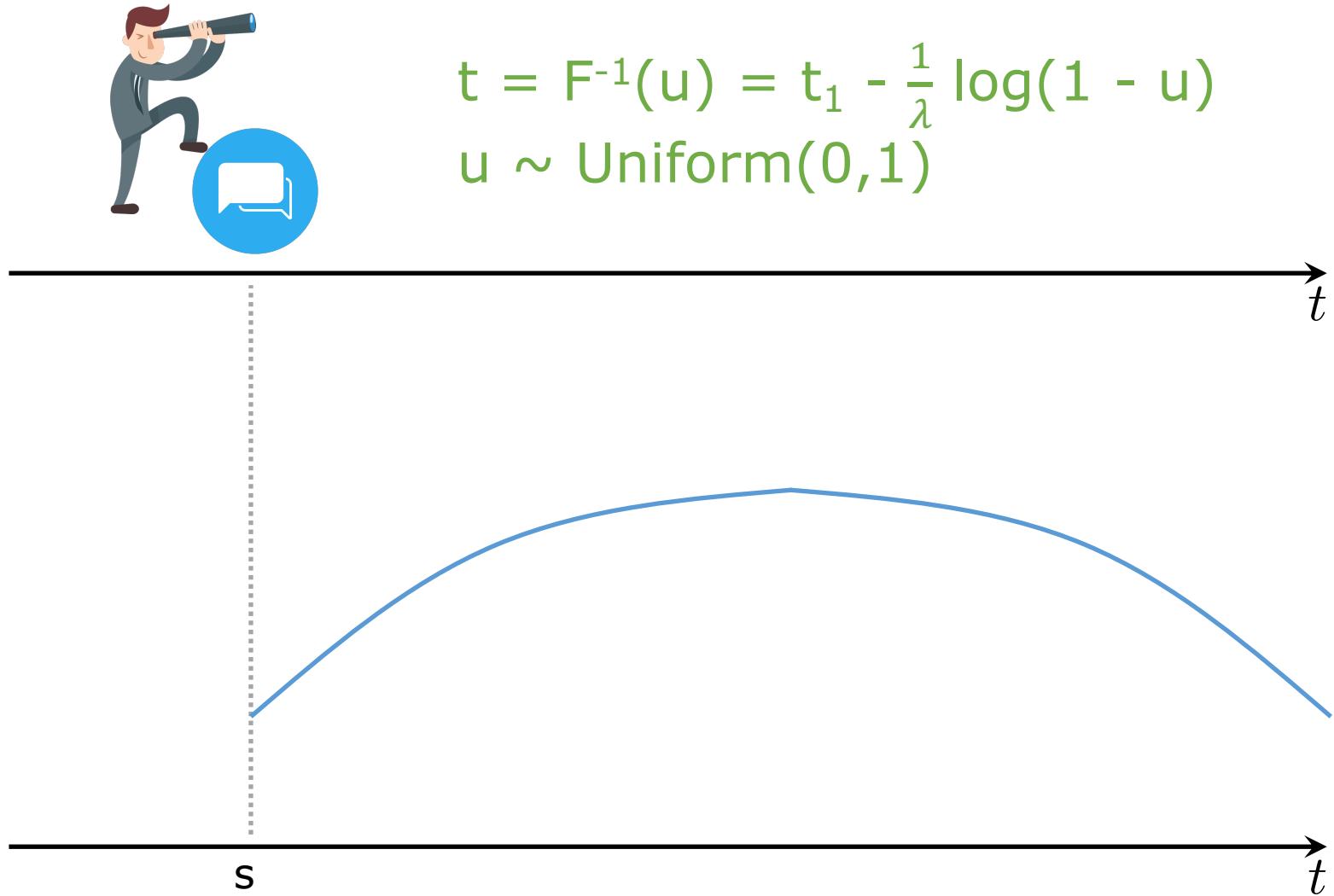
Lecture Structure

- Lecture-1: concepts
 - Intensity, point process, MLE, thinning, ...
- Lecture-2: fancy models
 - Hawkes, NHP, Neural ODE, ...
- Lecture-3: advanced topics
 - NCE, Datalog, marks, ...
 - Useful techniques to other kind of data

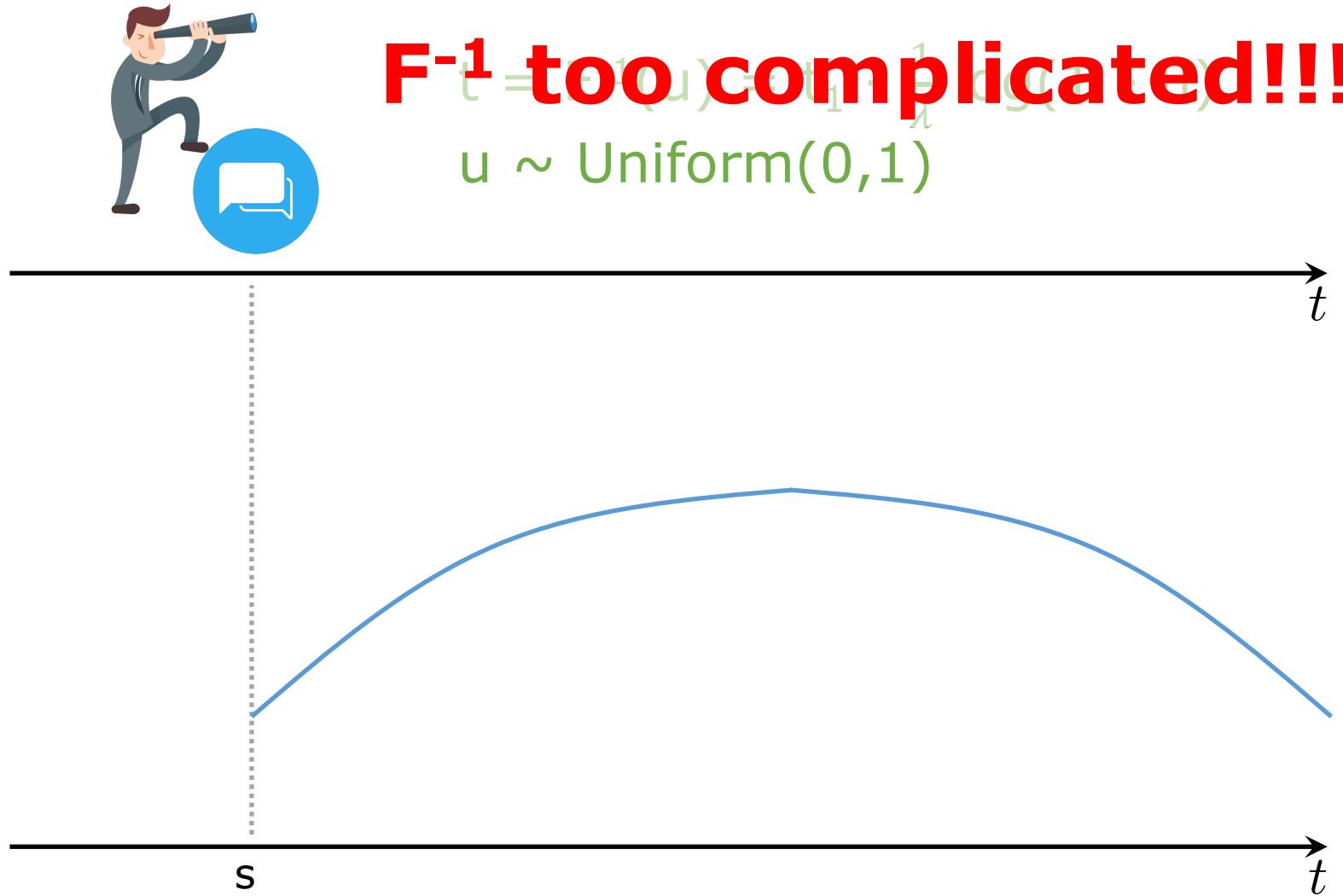
When intensities are not constant...



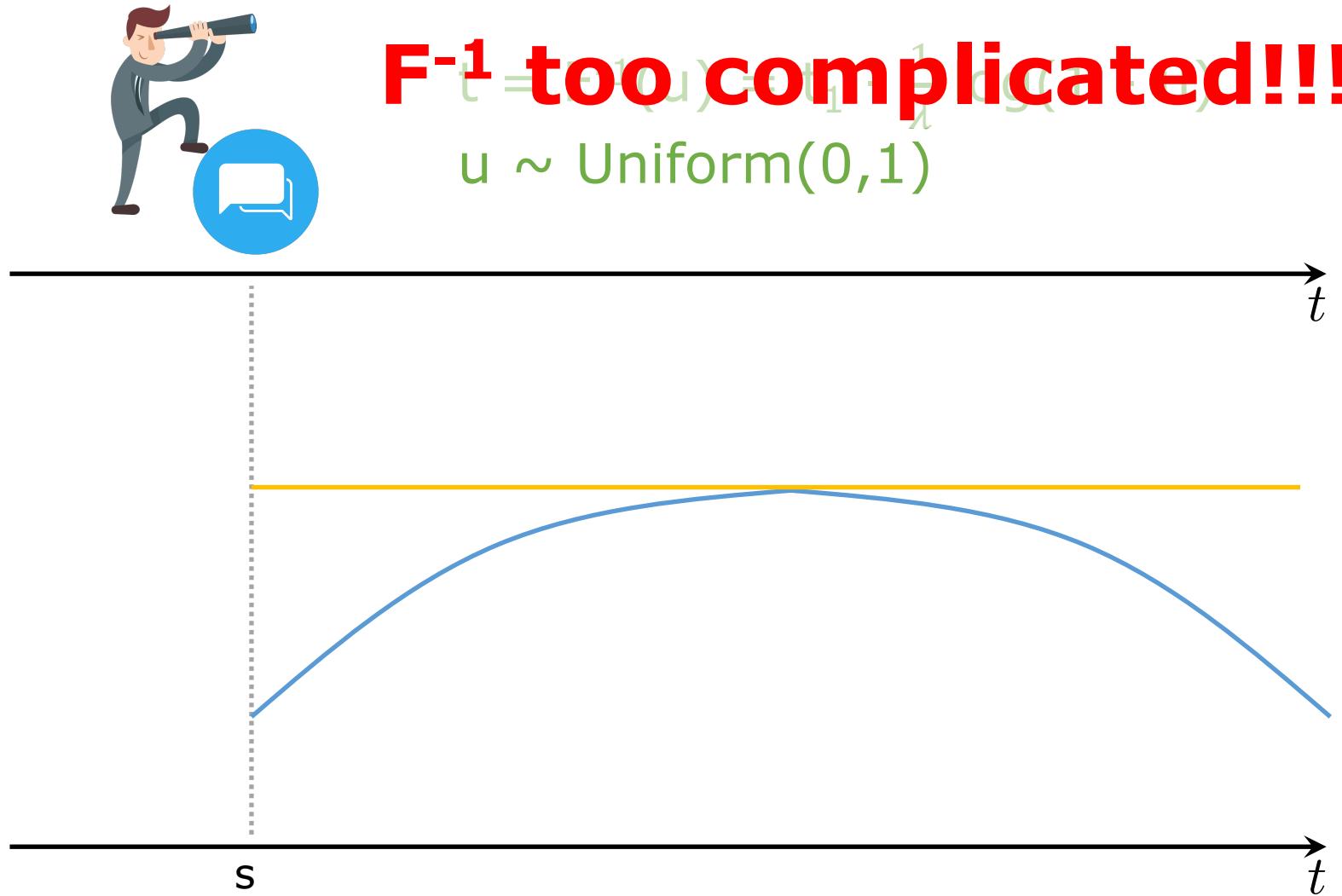
When intensities are not constant...



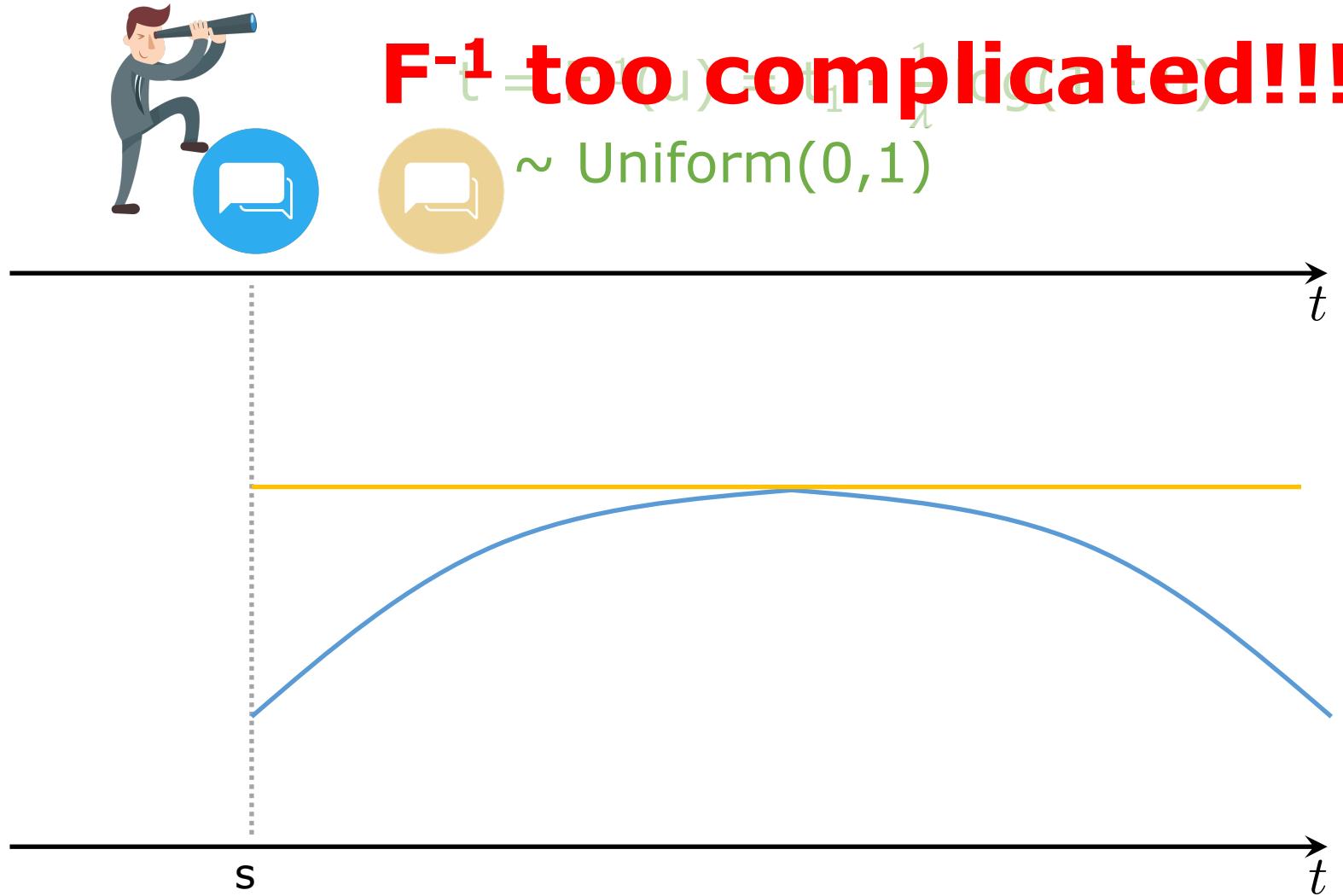
When intensities are not constant...



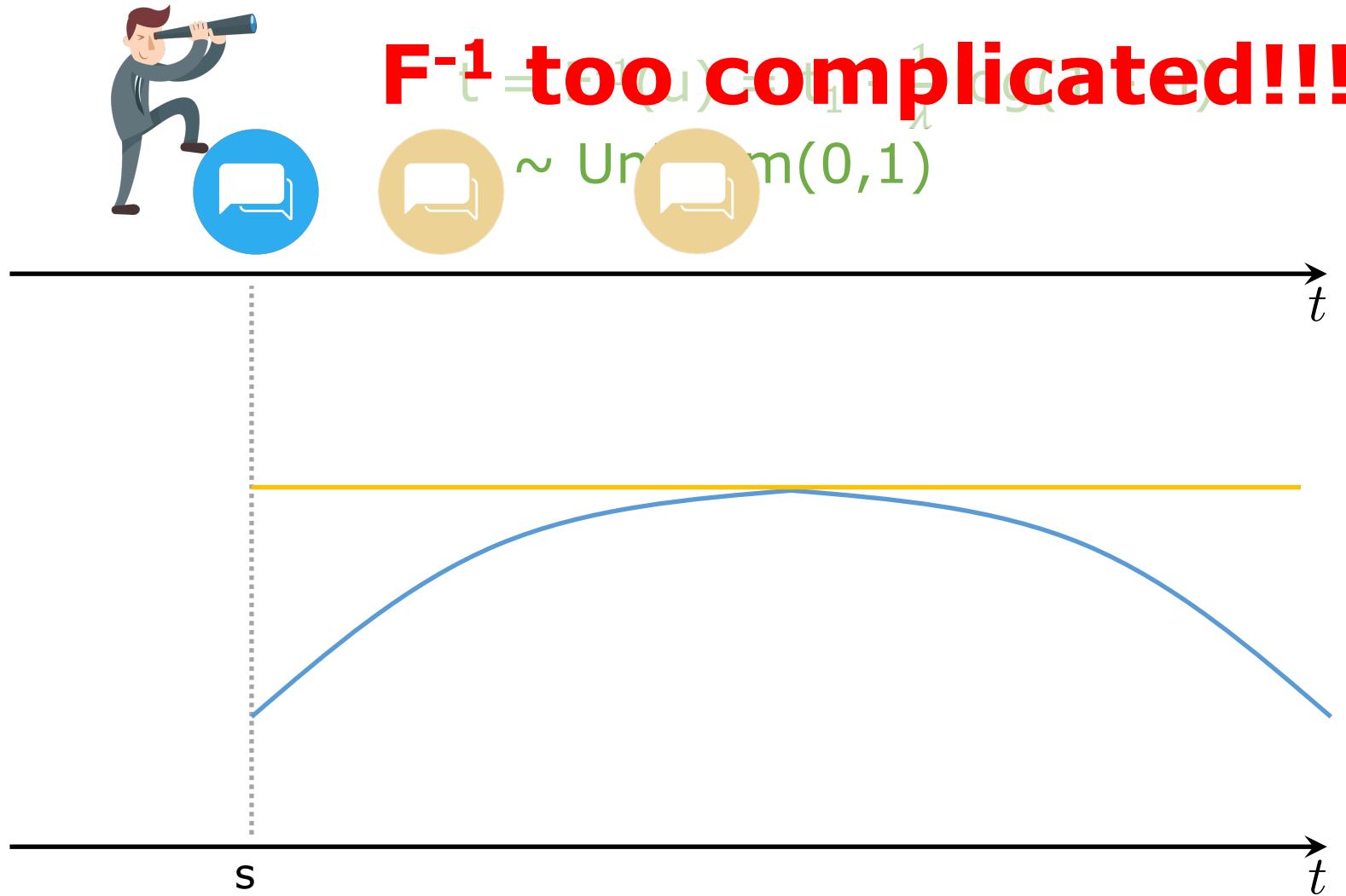
When intensities are not constant...



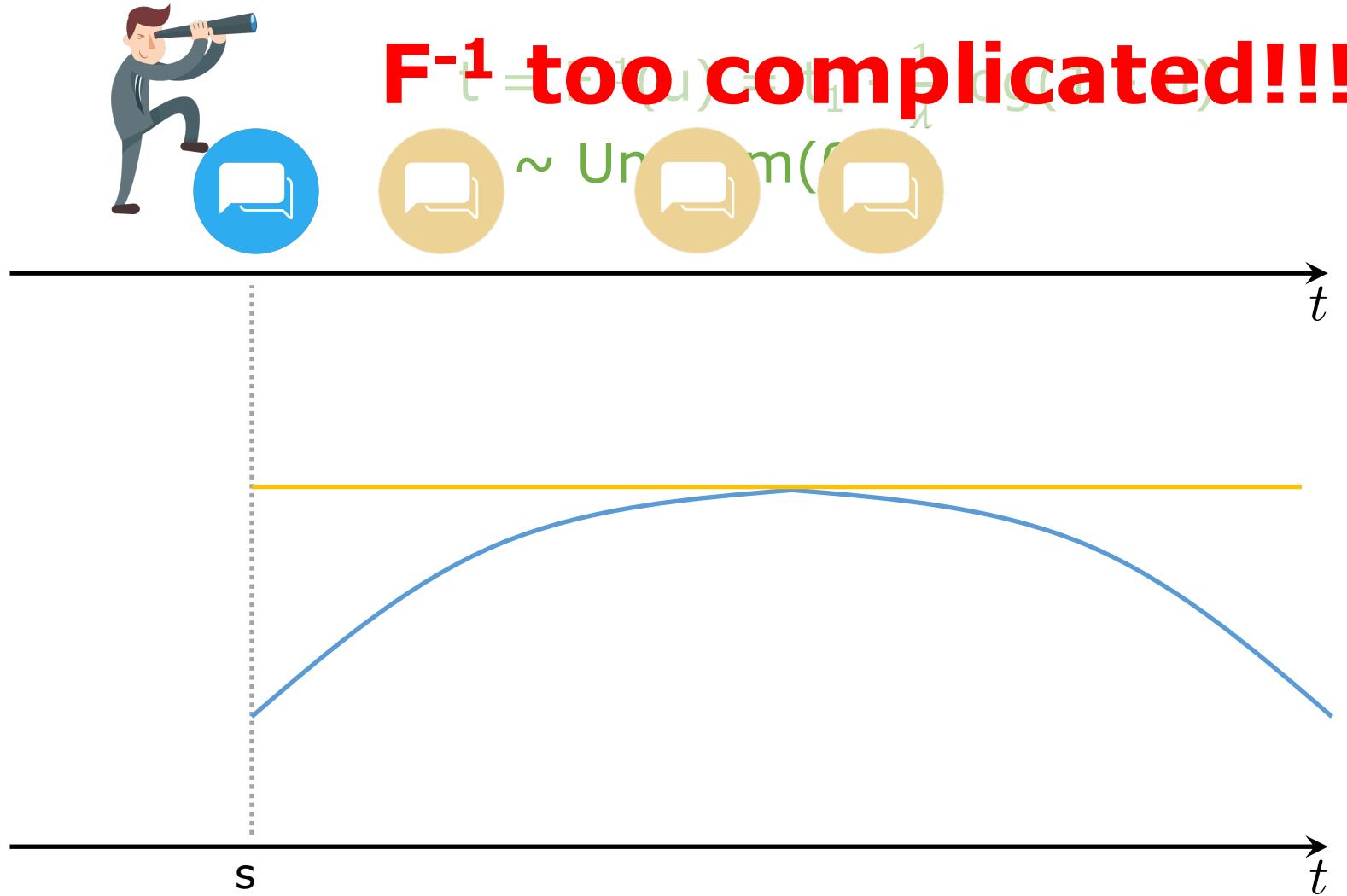
When intensities are not constant...



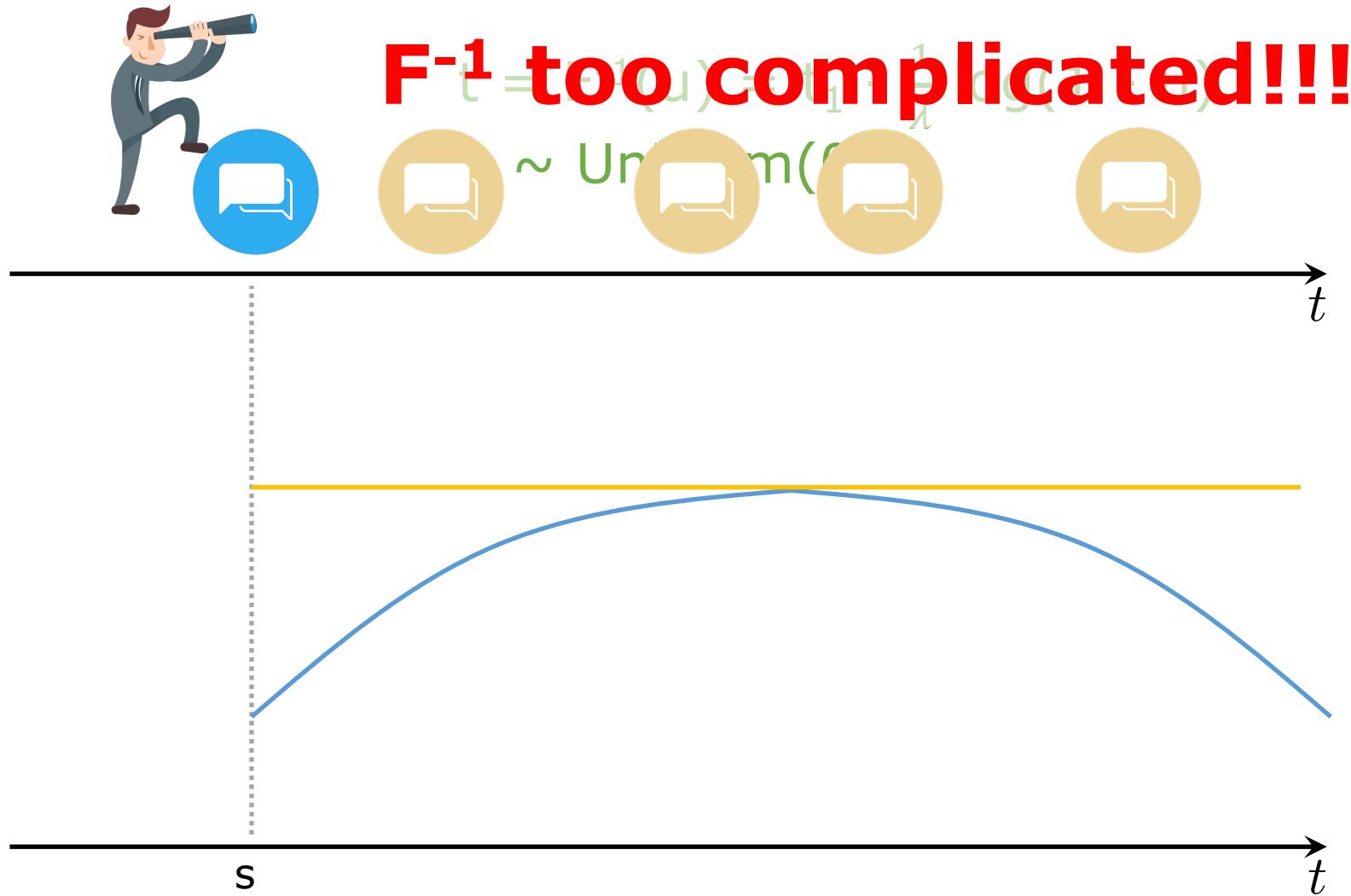
When intensities are not constant...



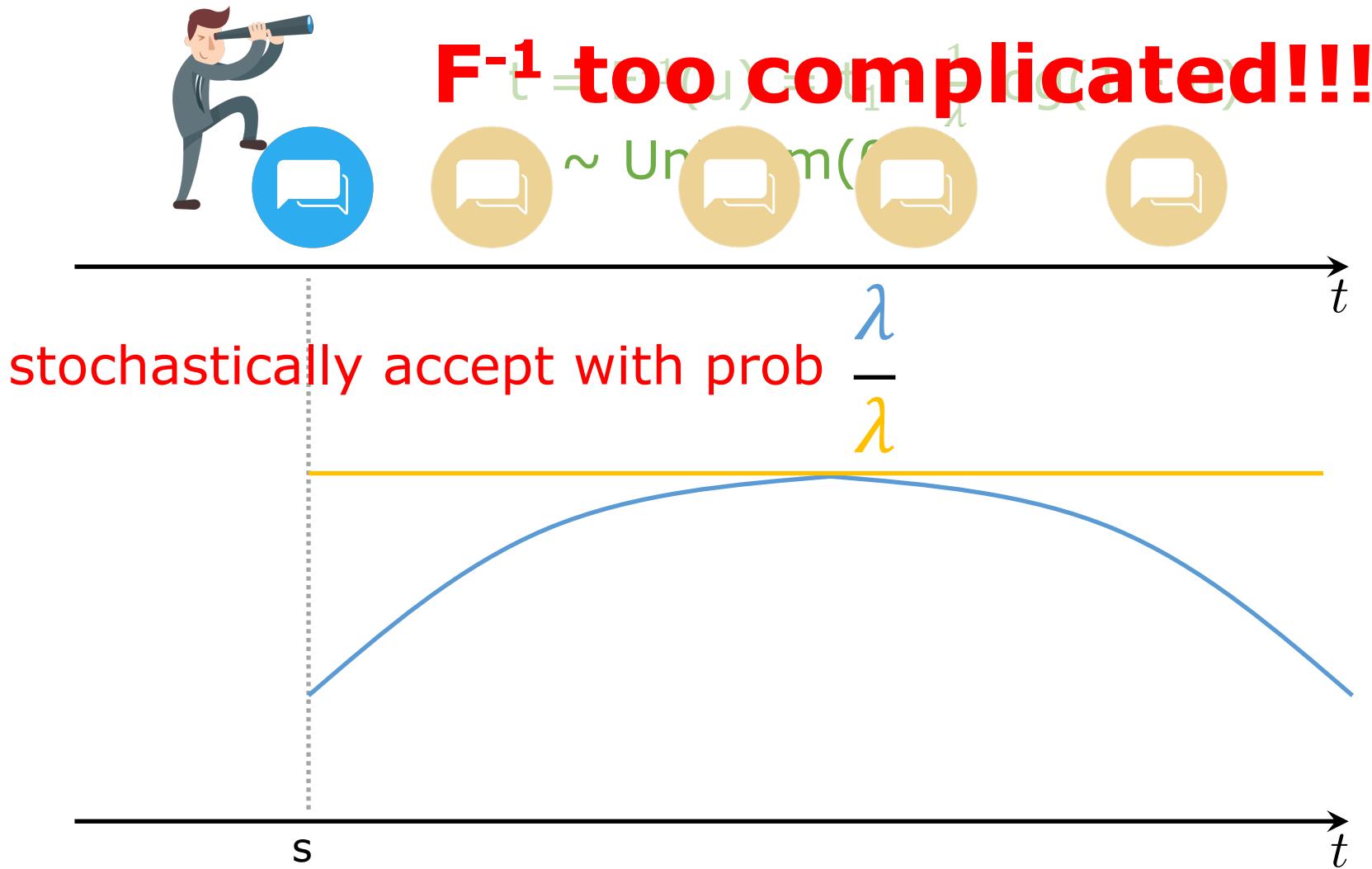
When intensities are not constant...



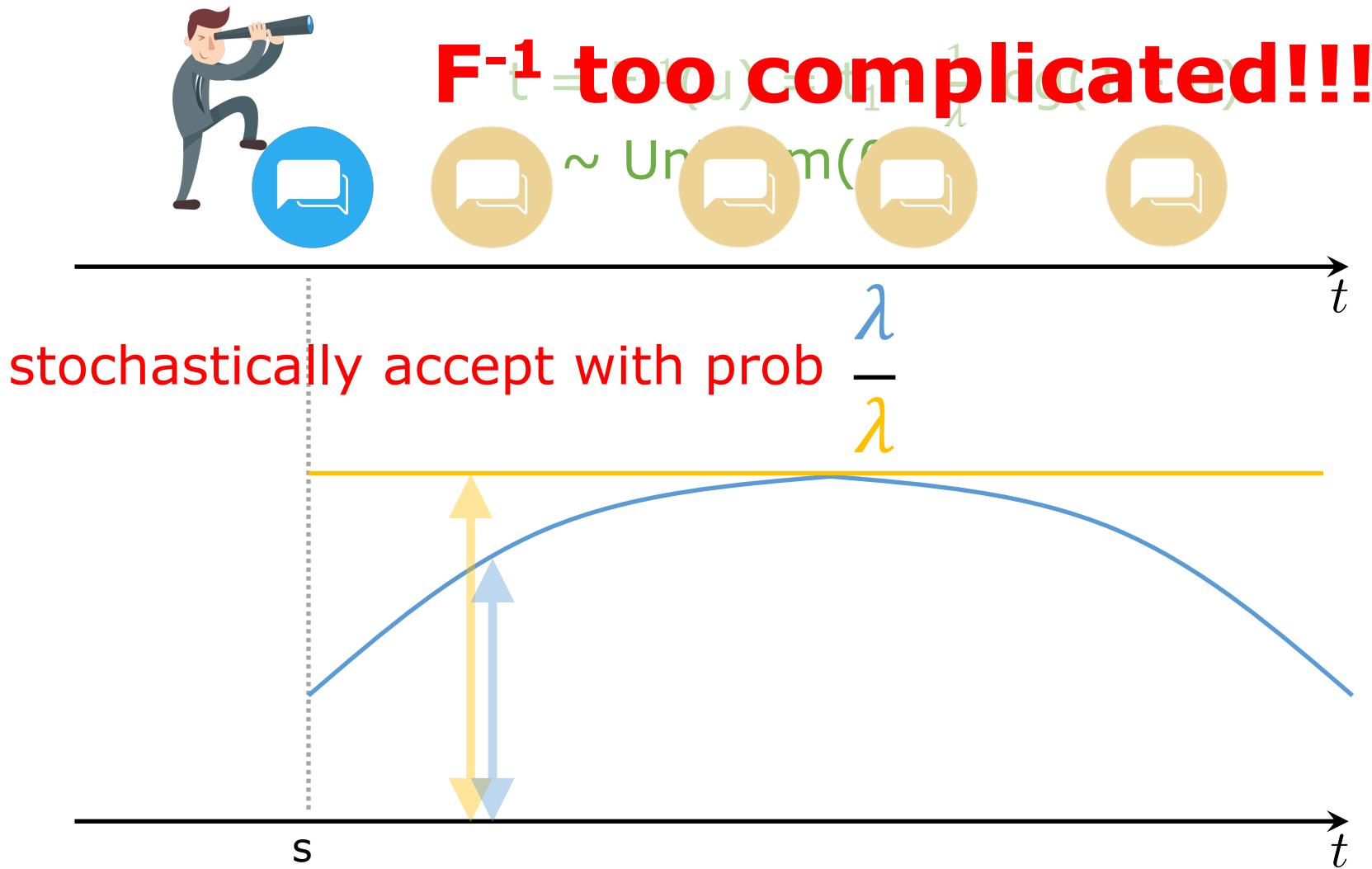
When intensities are not constant...



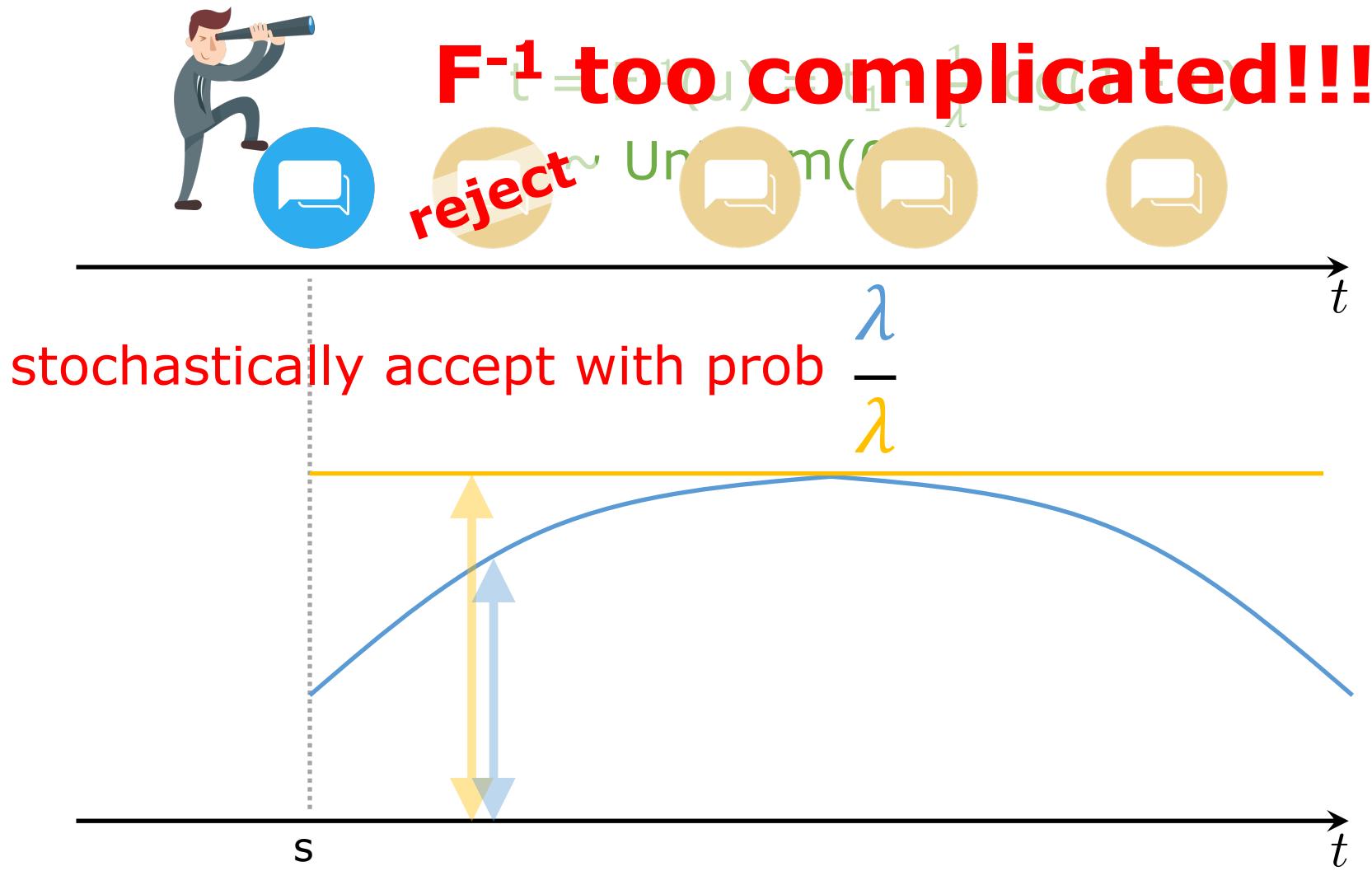
When intensities are not constant...



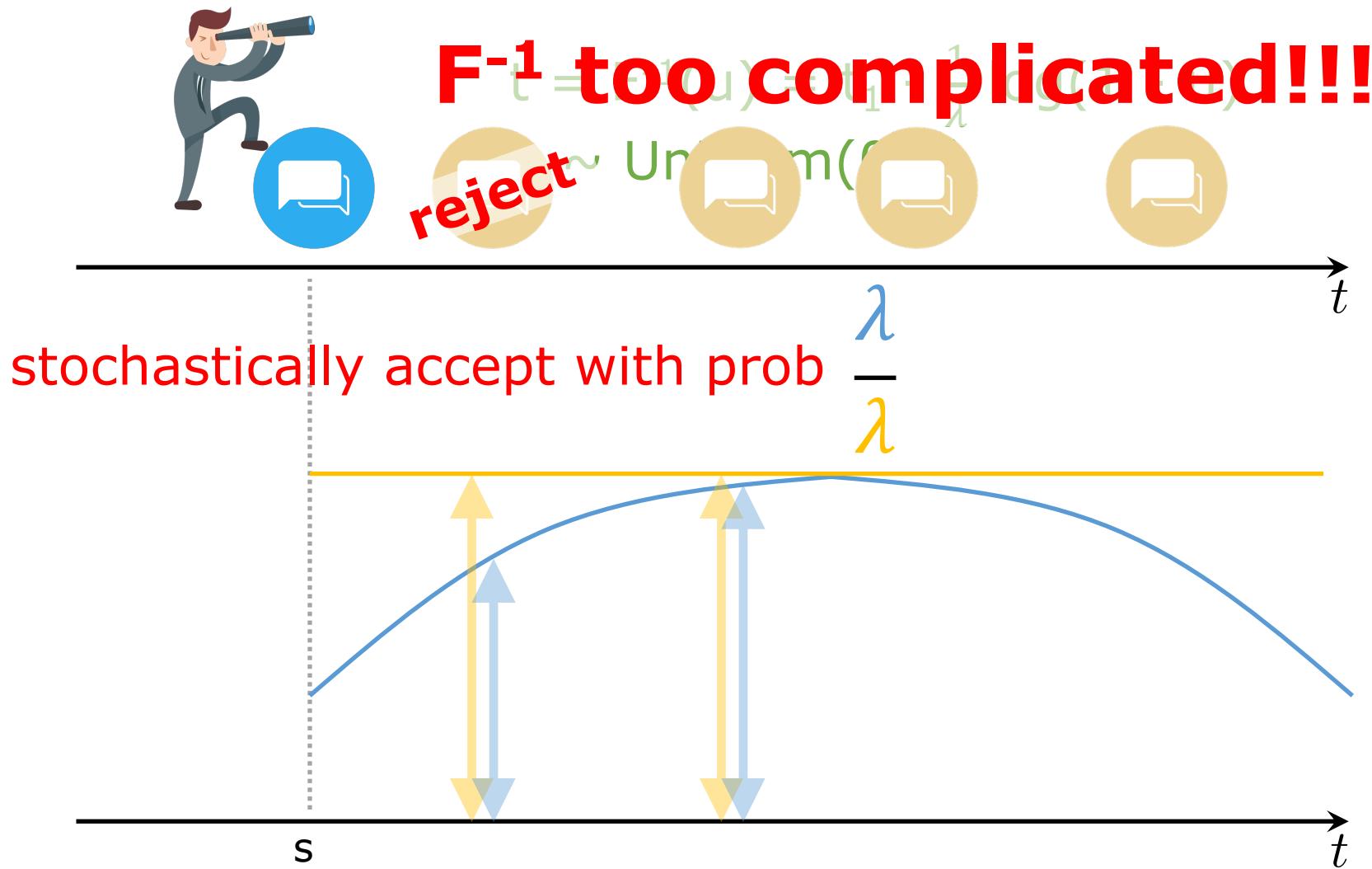
When intensities are not constant...



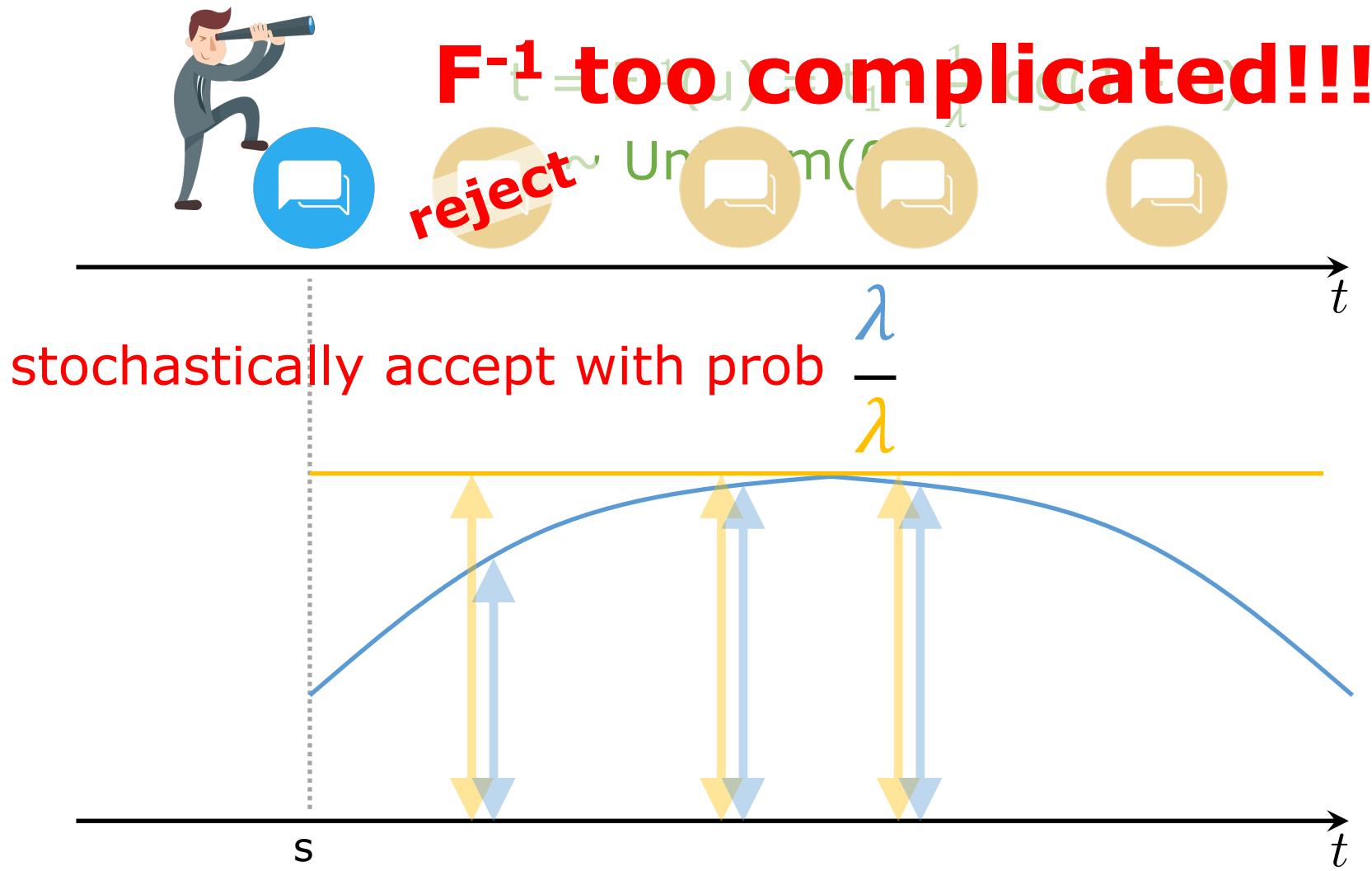
When intensities are not constant...



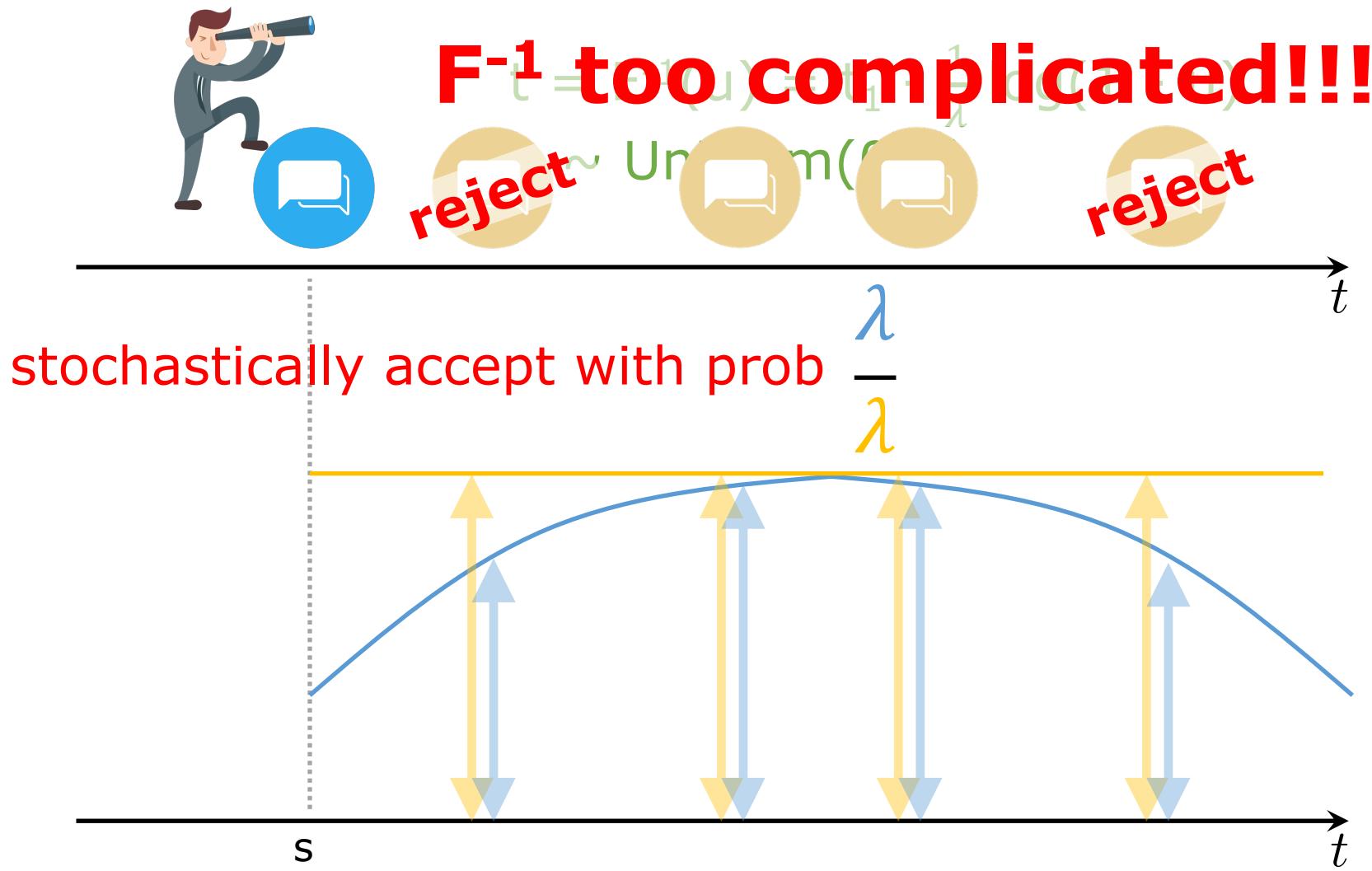
When intensities are not constant...



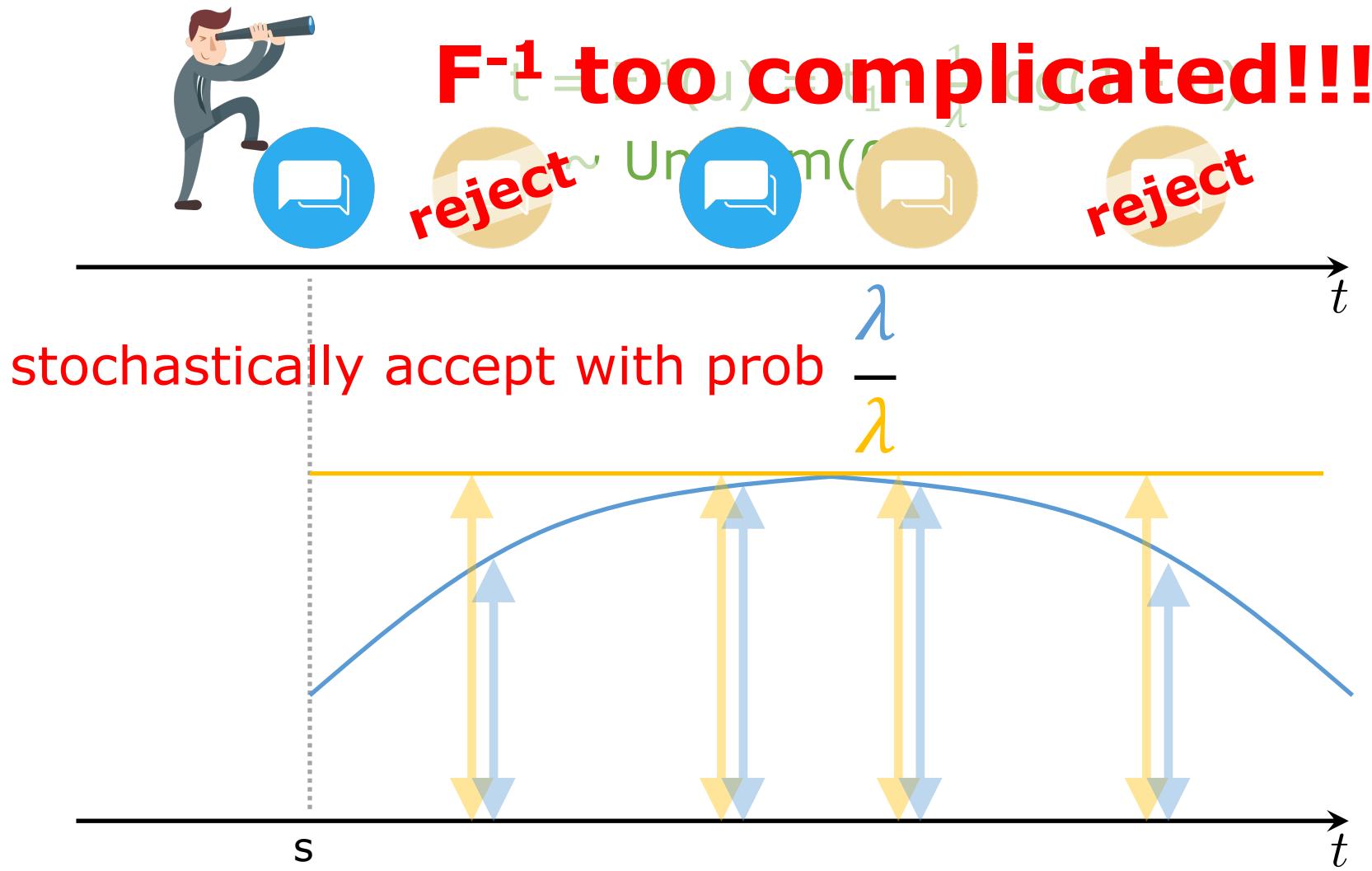
When intensities are not constant...



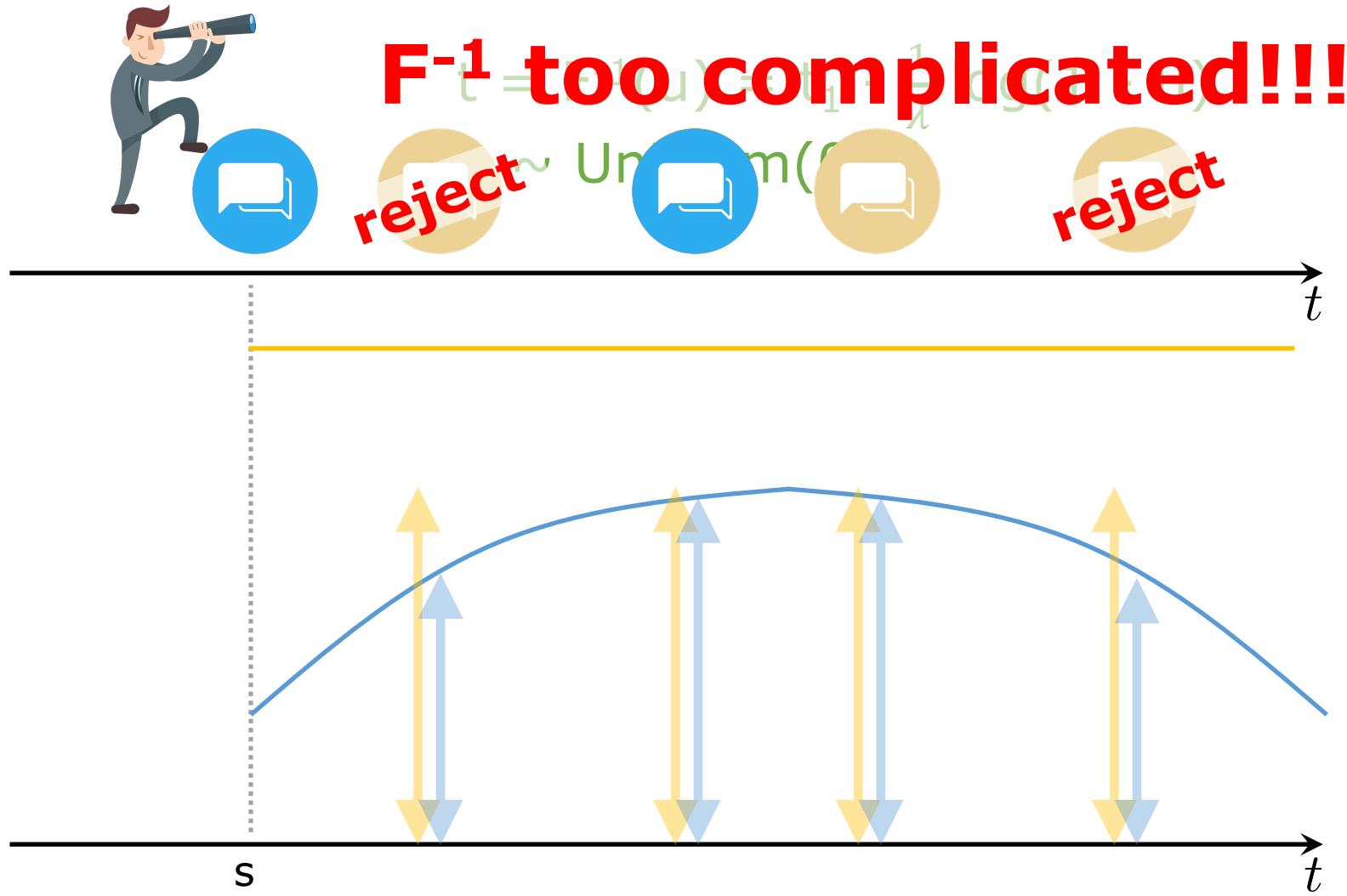
When intensities are not constant...



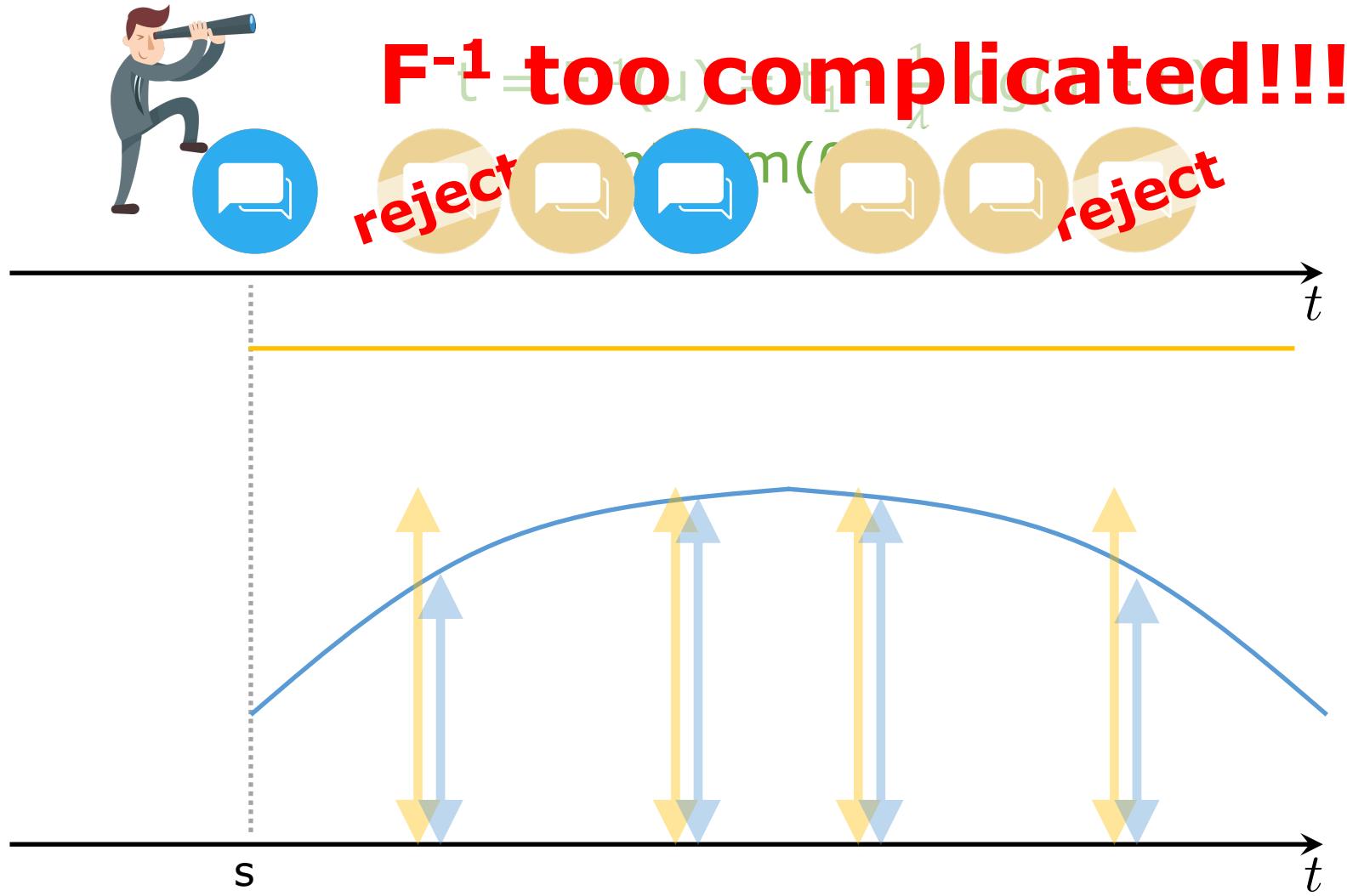
When intensities are not constant...



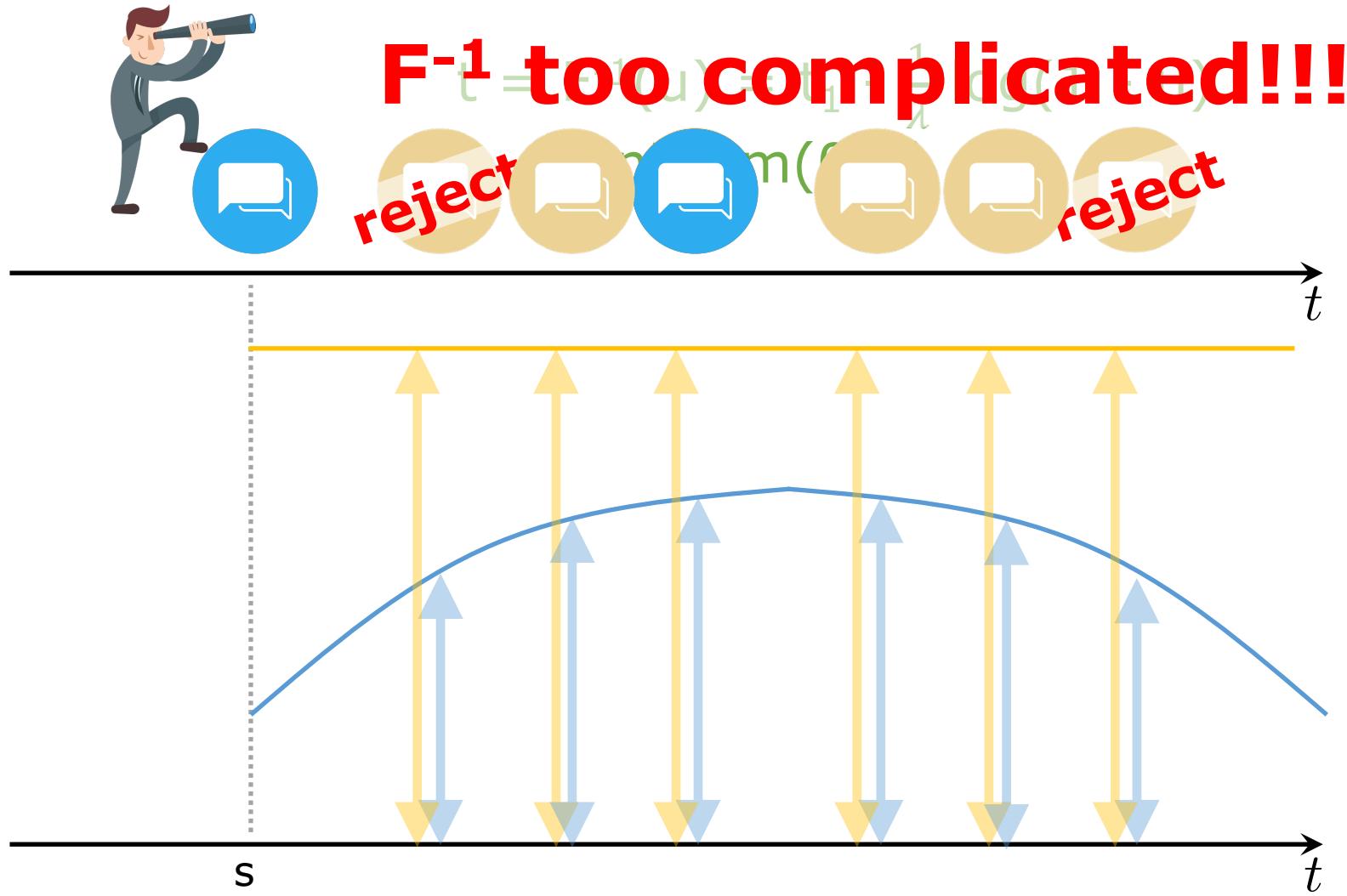
When intensities are not constant...



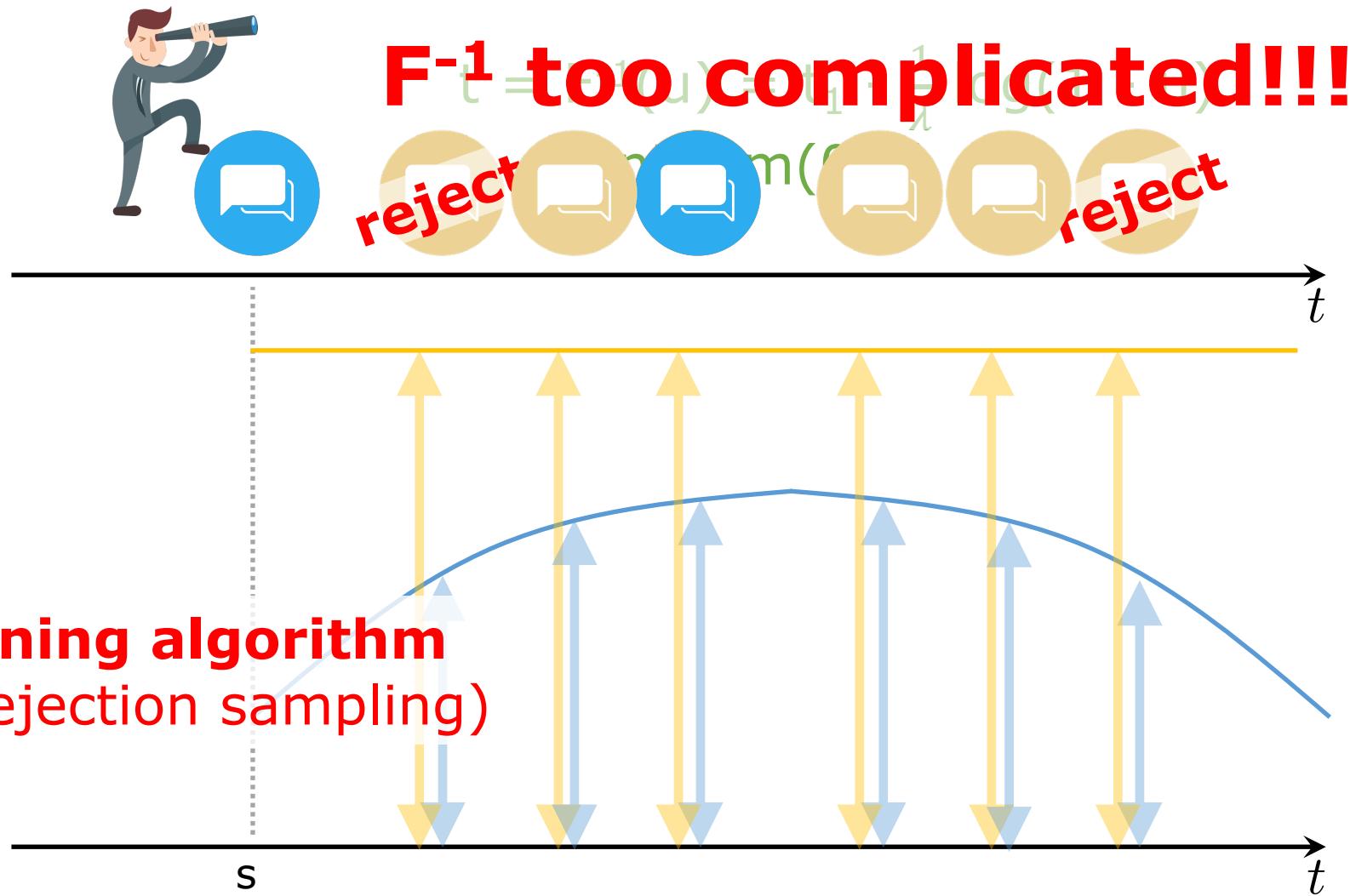
When intensities are not constant...



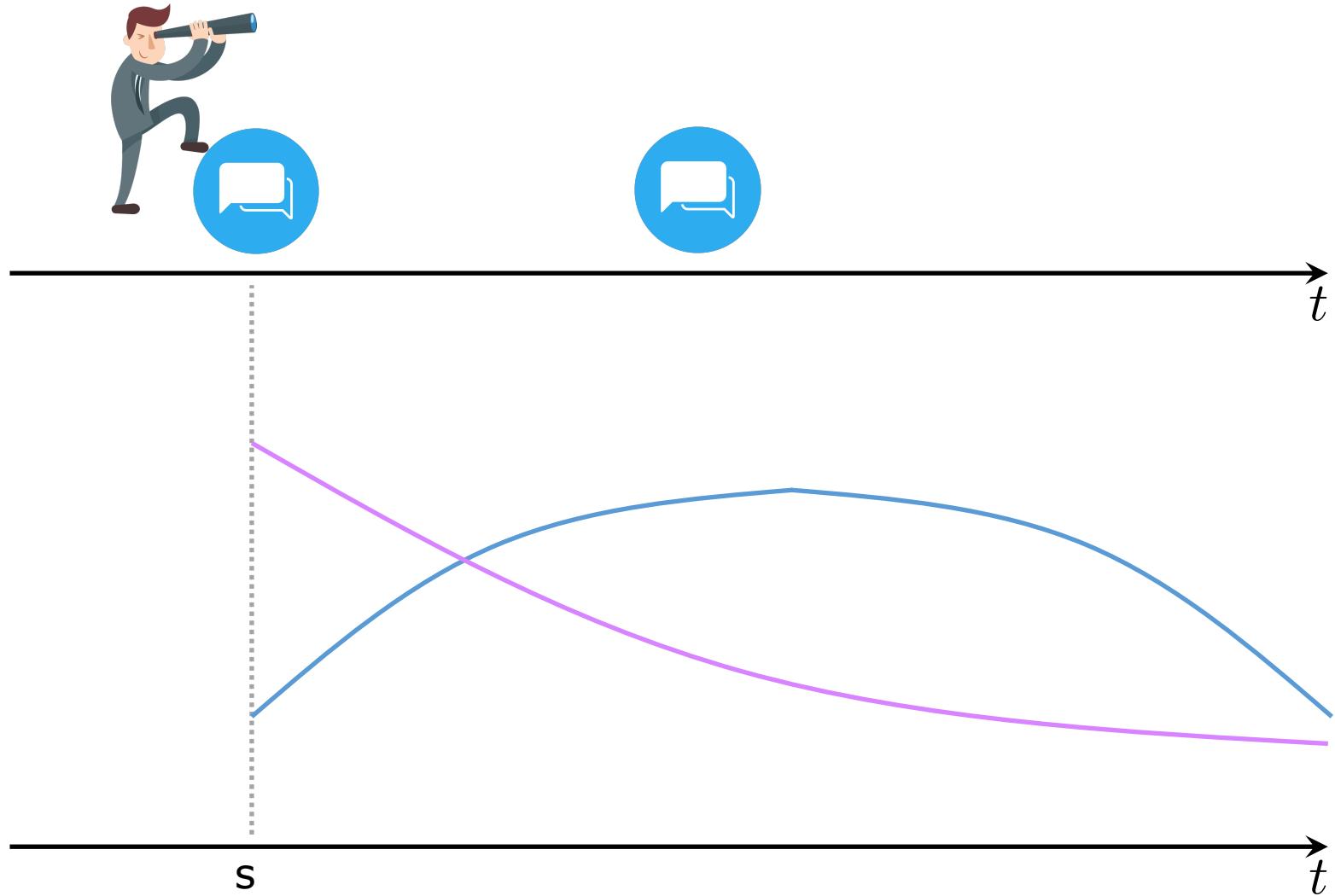
When intensities are not constant...



When intensities are not constant...

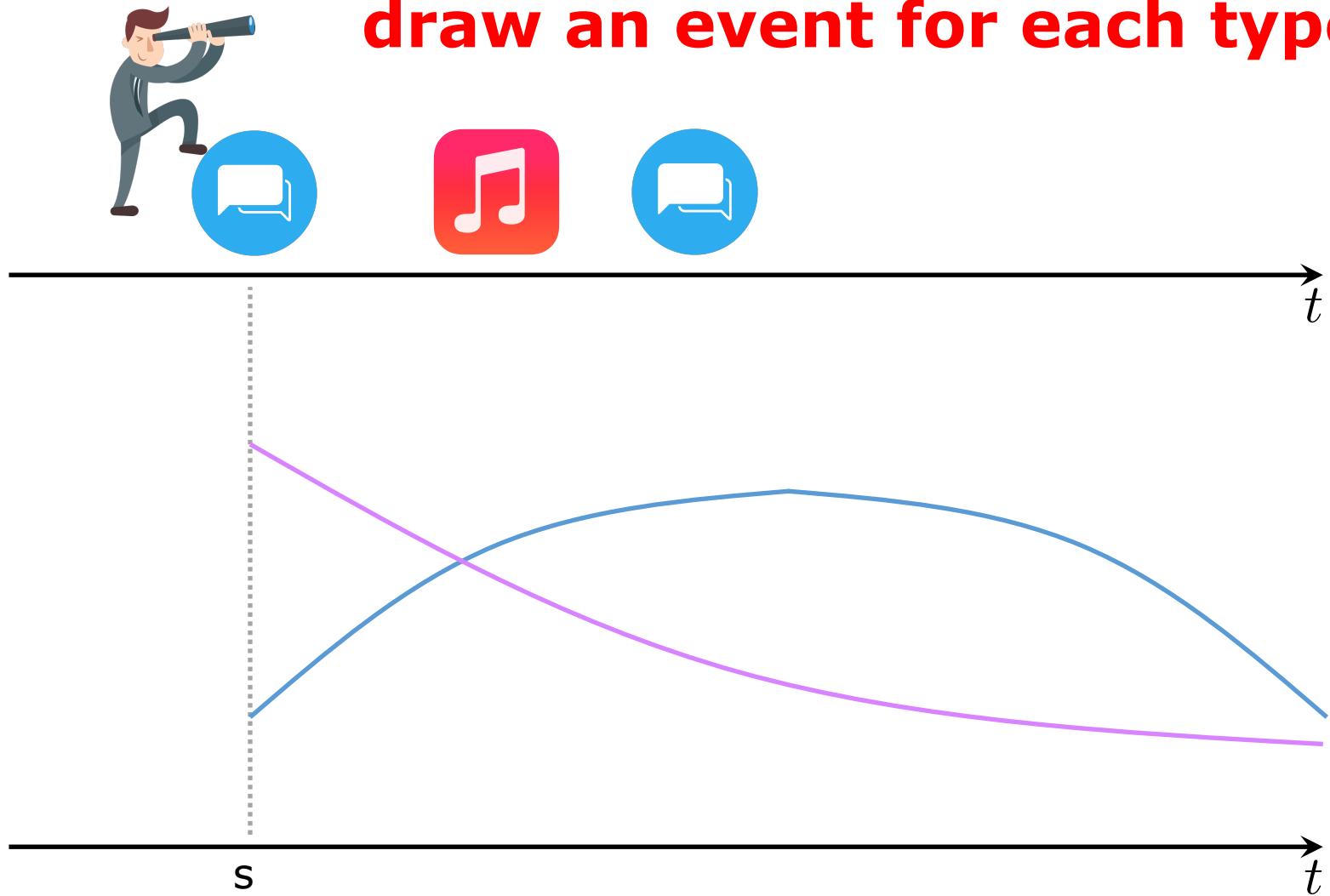


Different types *race* against one another

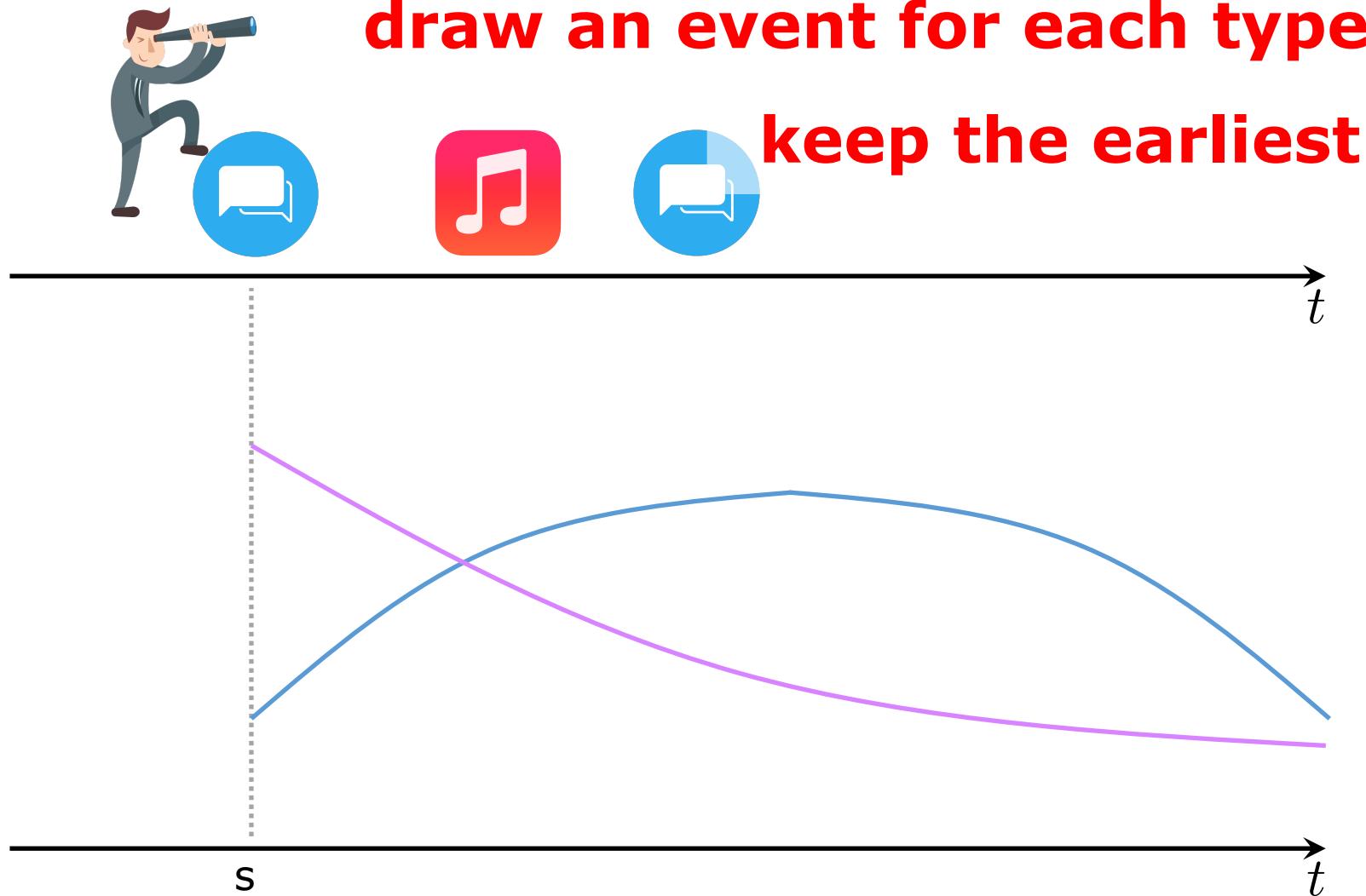


Different types *race* against one another

draw an event for each type

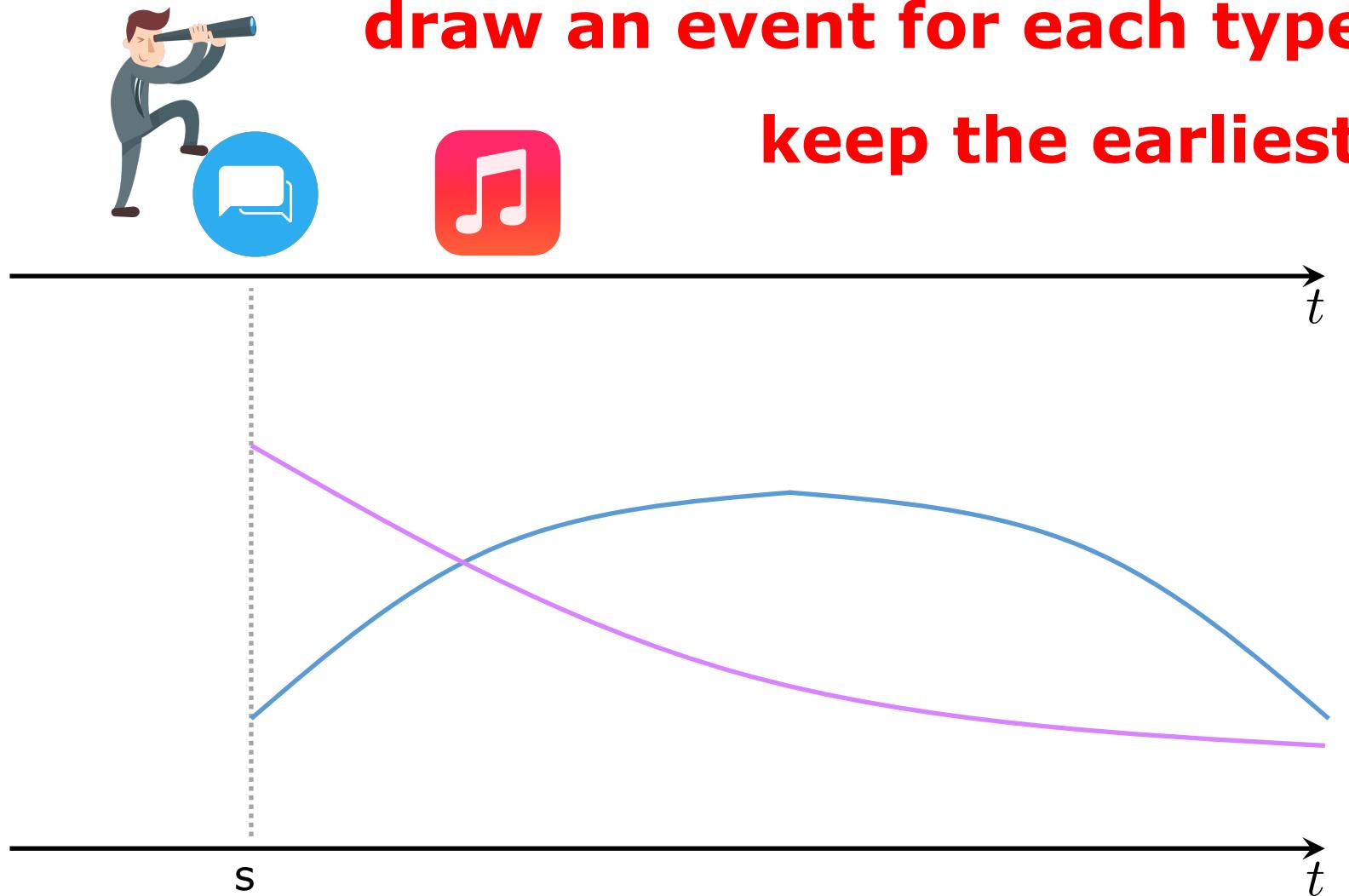


Different types *race* against one another

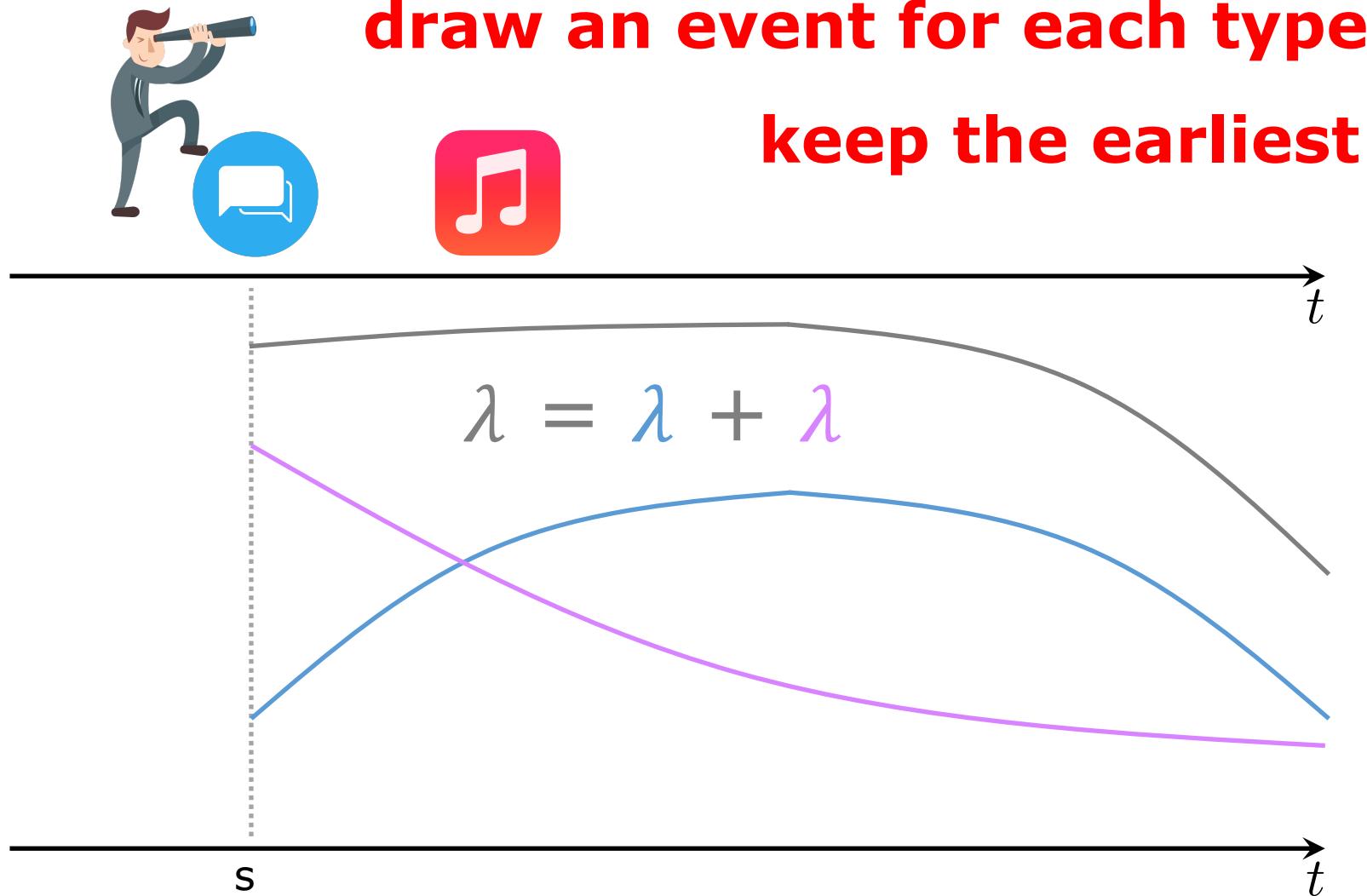


Different types *race* against one another

draw an event for each type

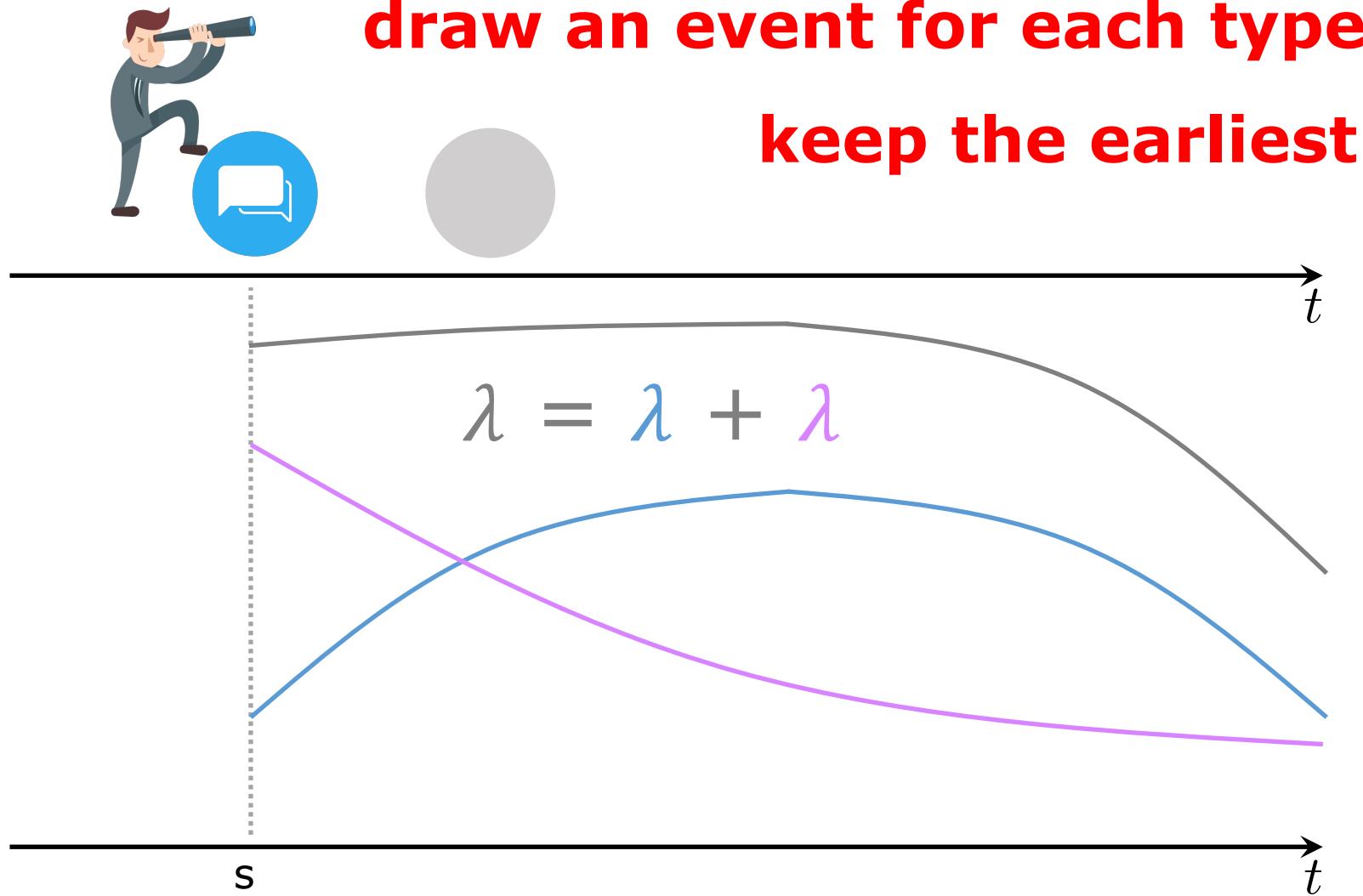


Different types *race* against one another

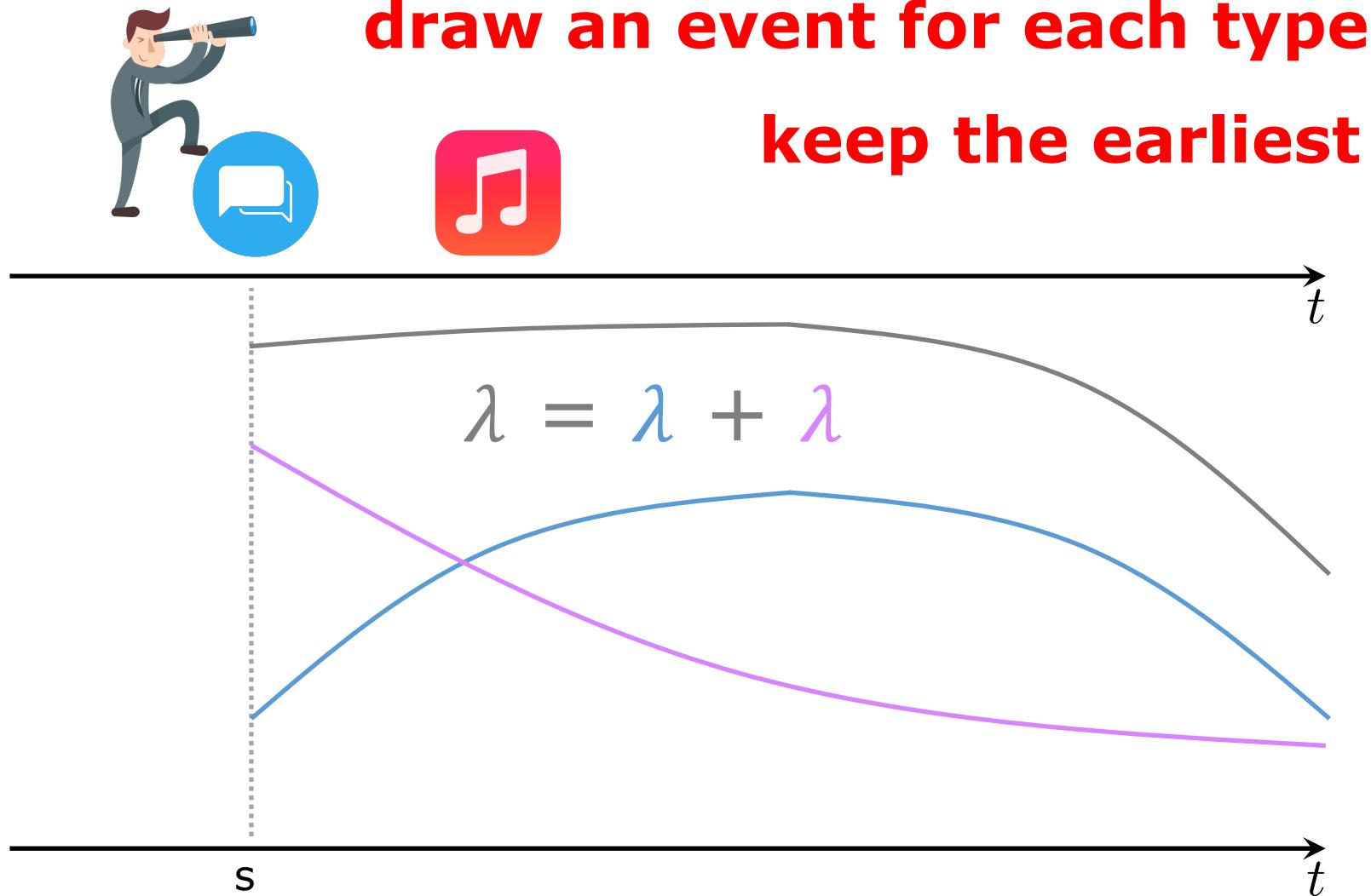


Different types *race* against one another

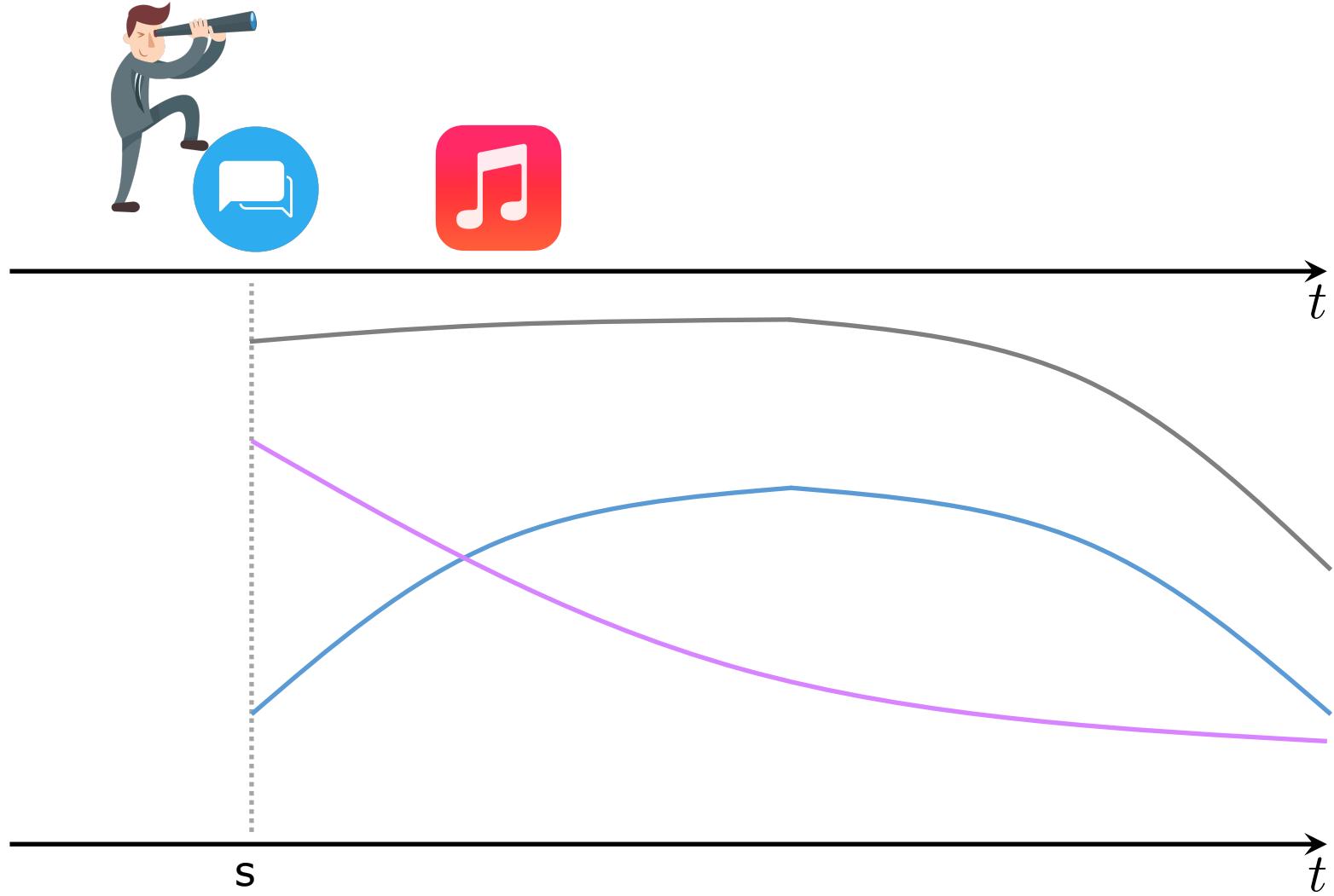
draw an event for each type



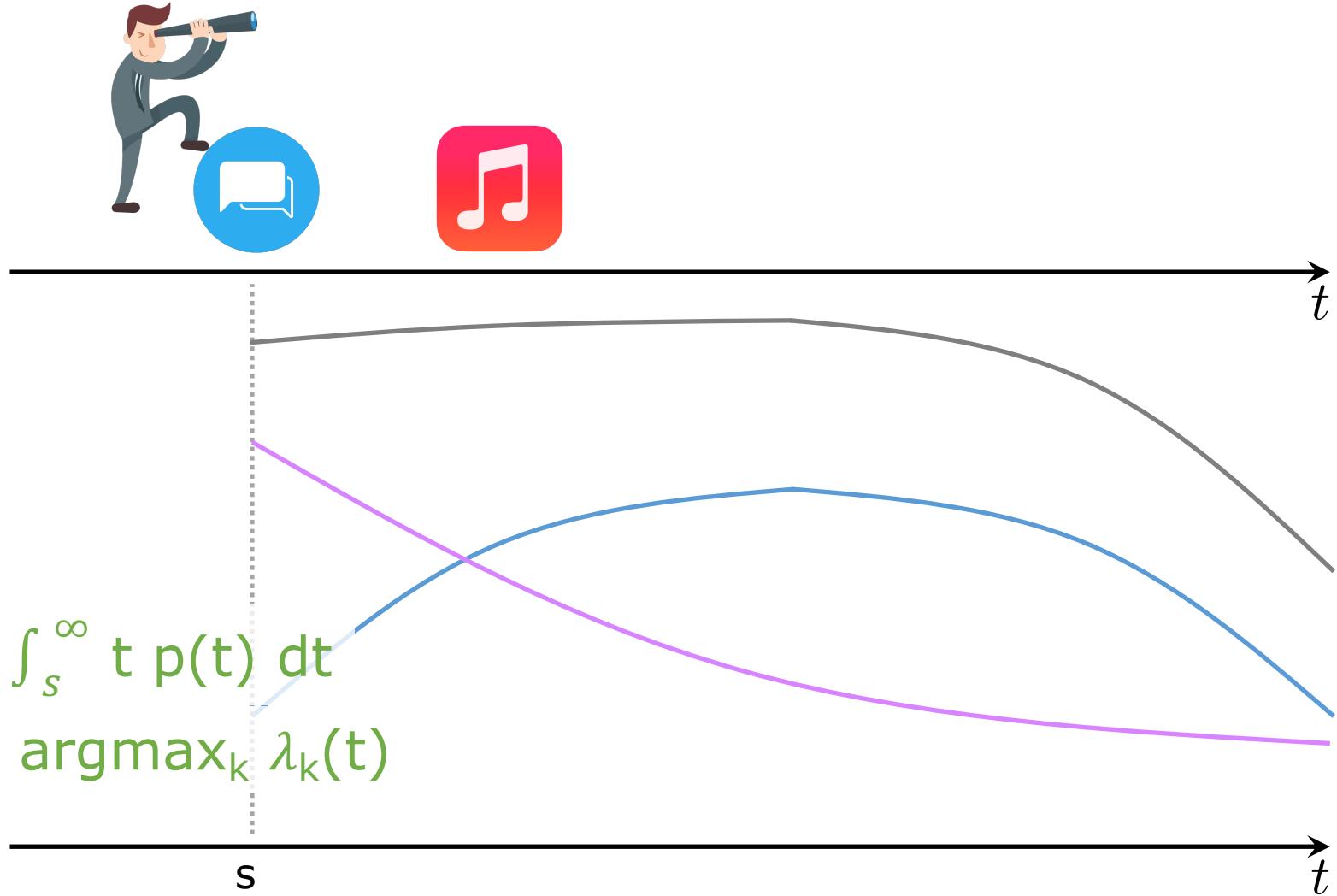
Different types *race* against one another



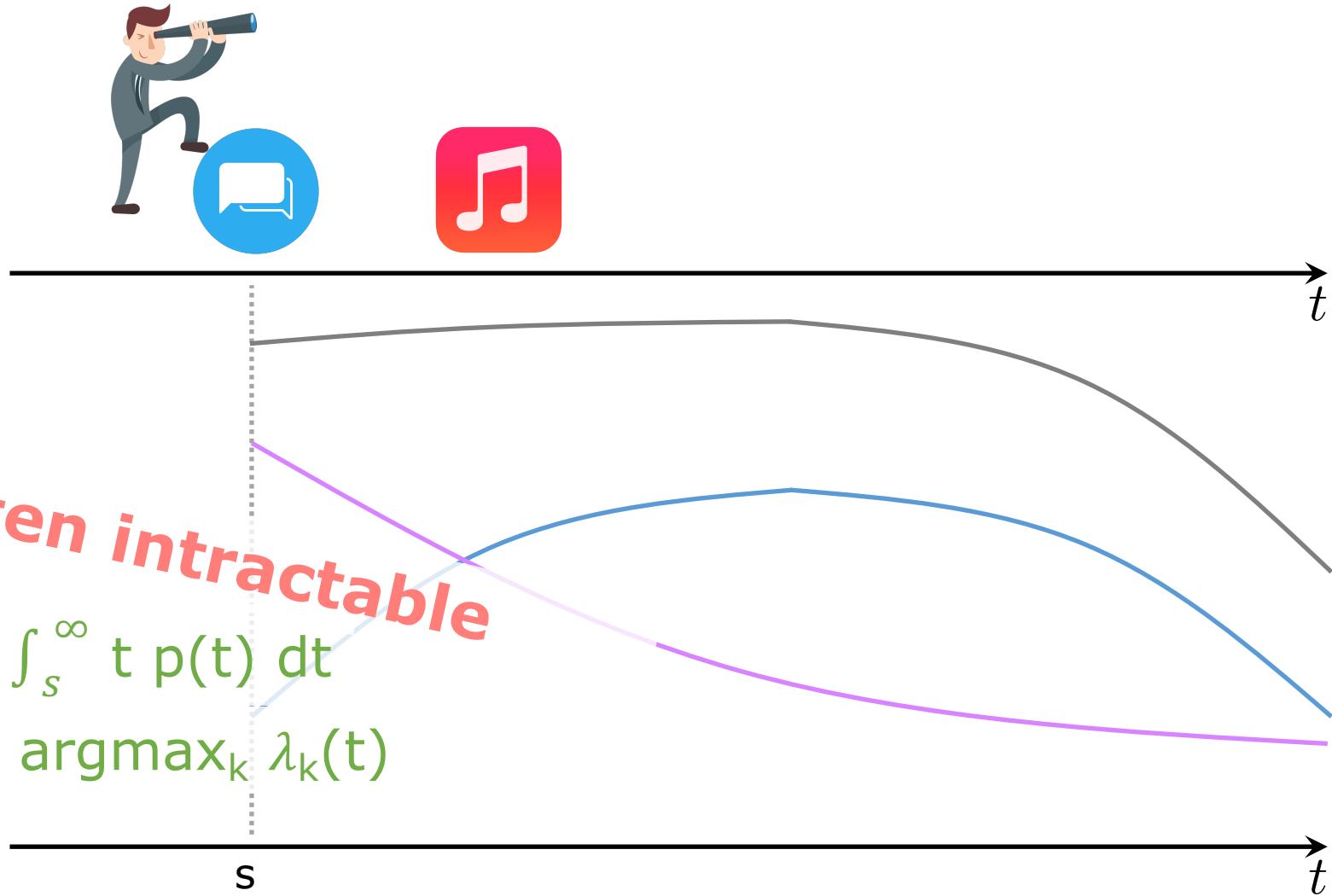
Predict future by sampling



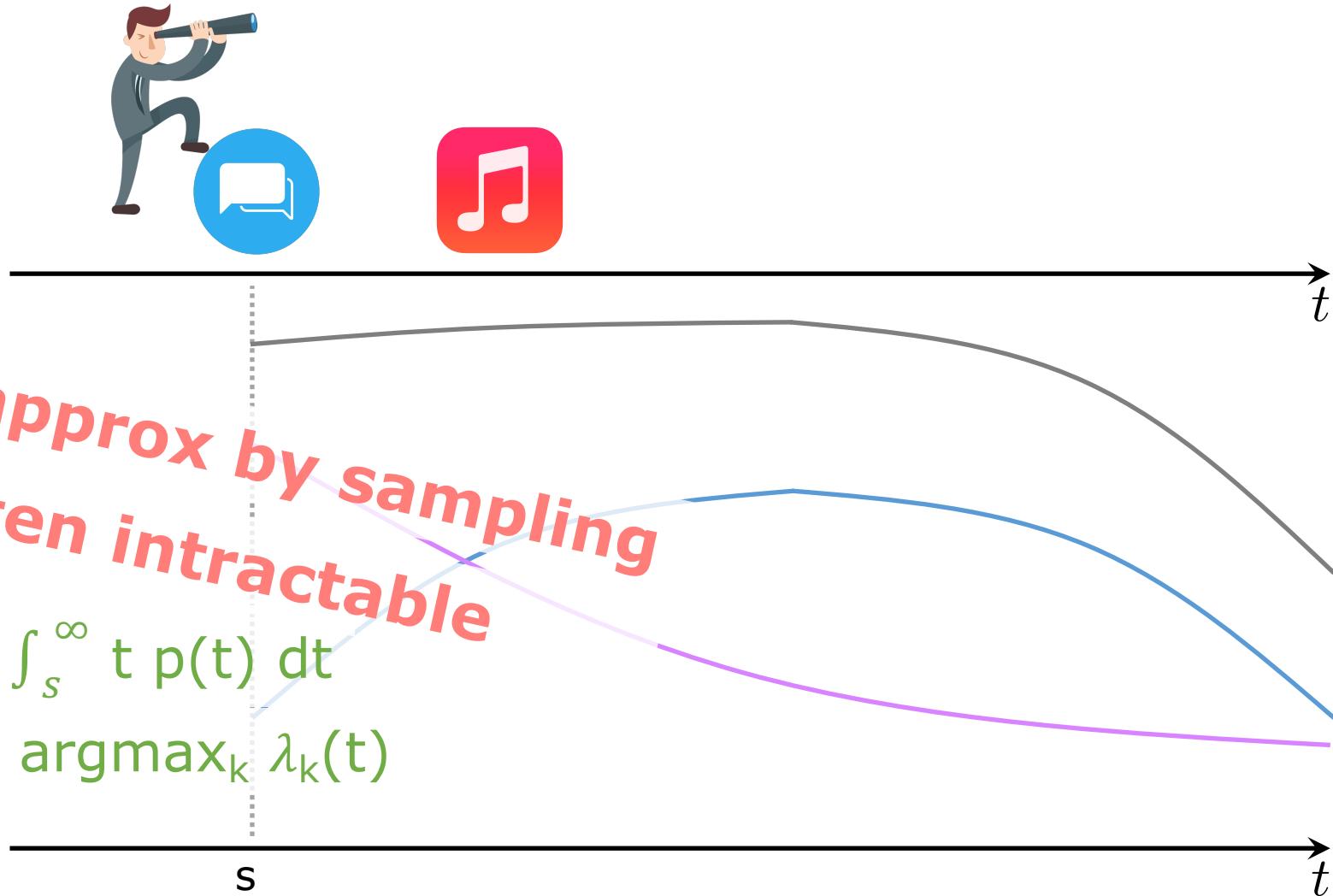
Predict future by sampling



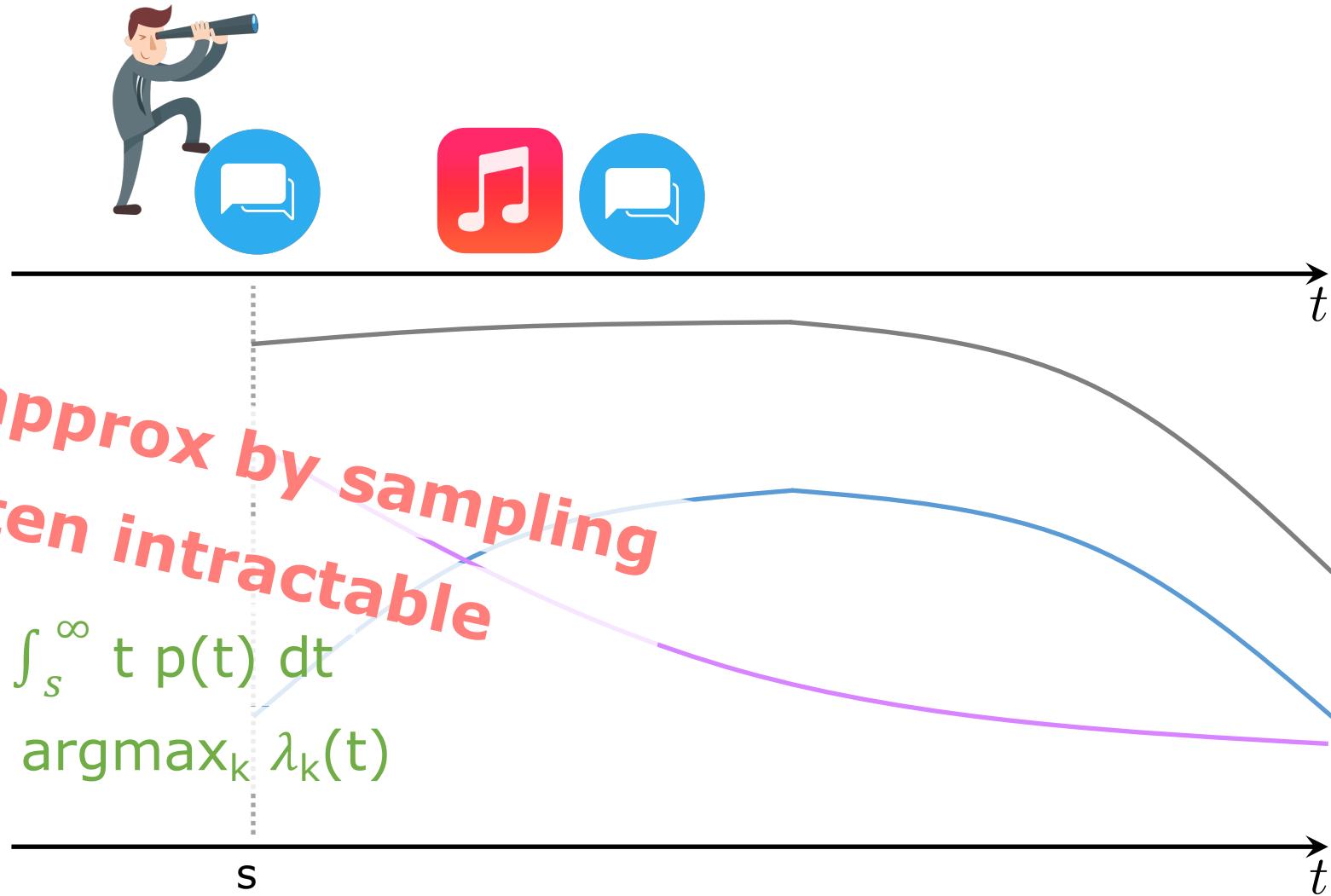
Predict future by sampling



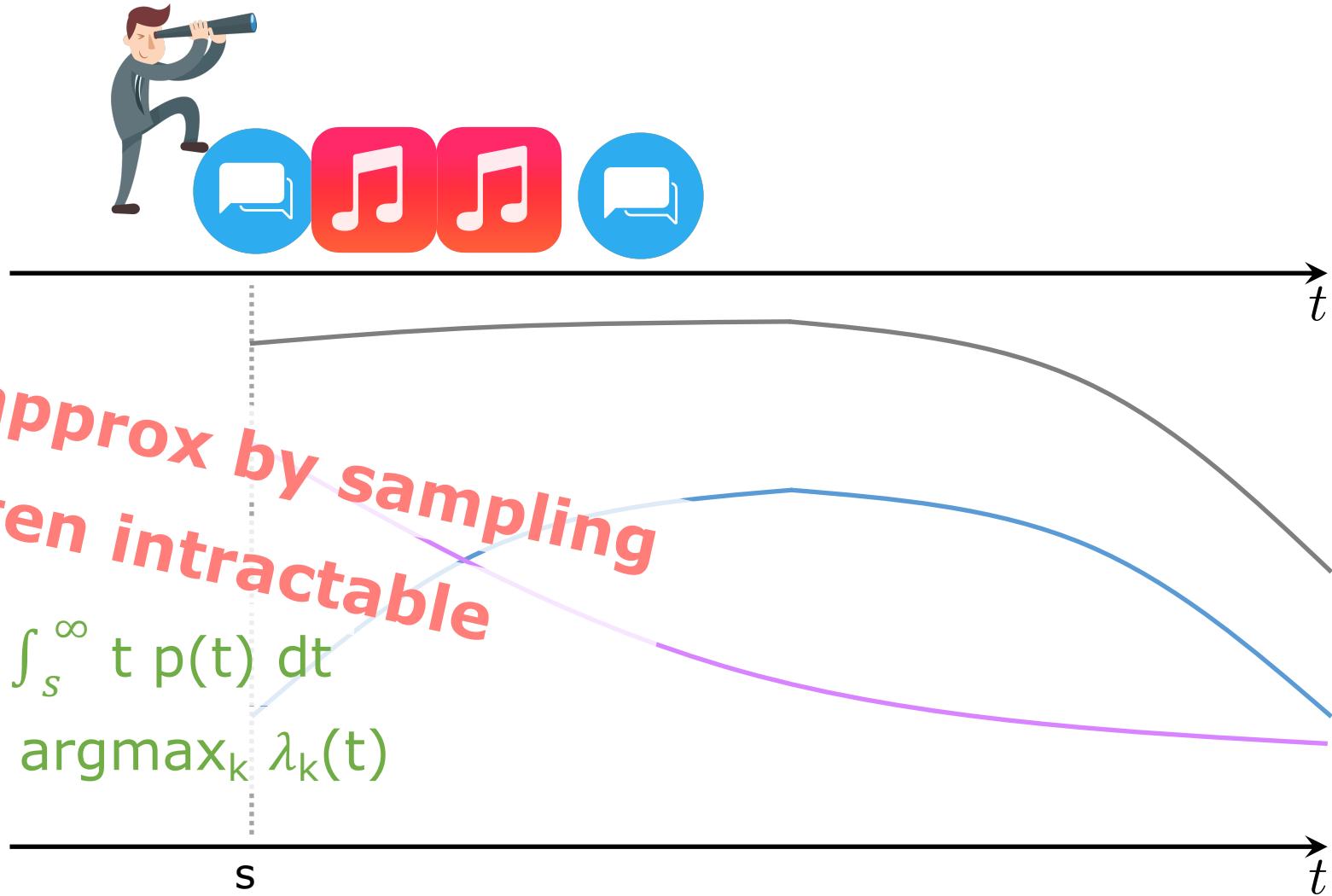
Predict future by sampling



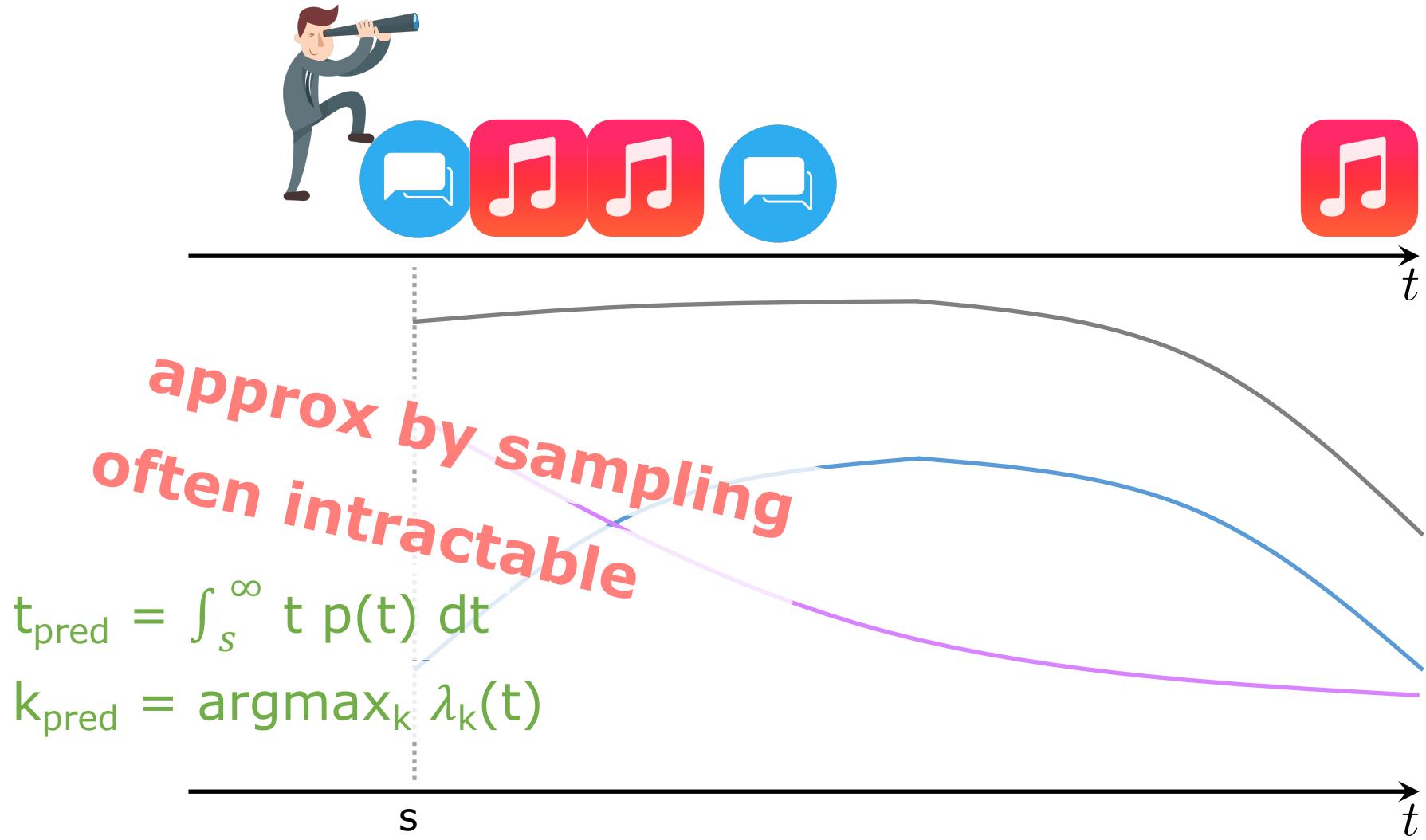
Predict future by sampling



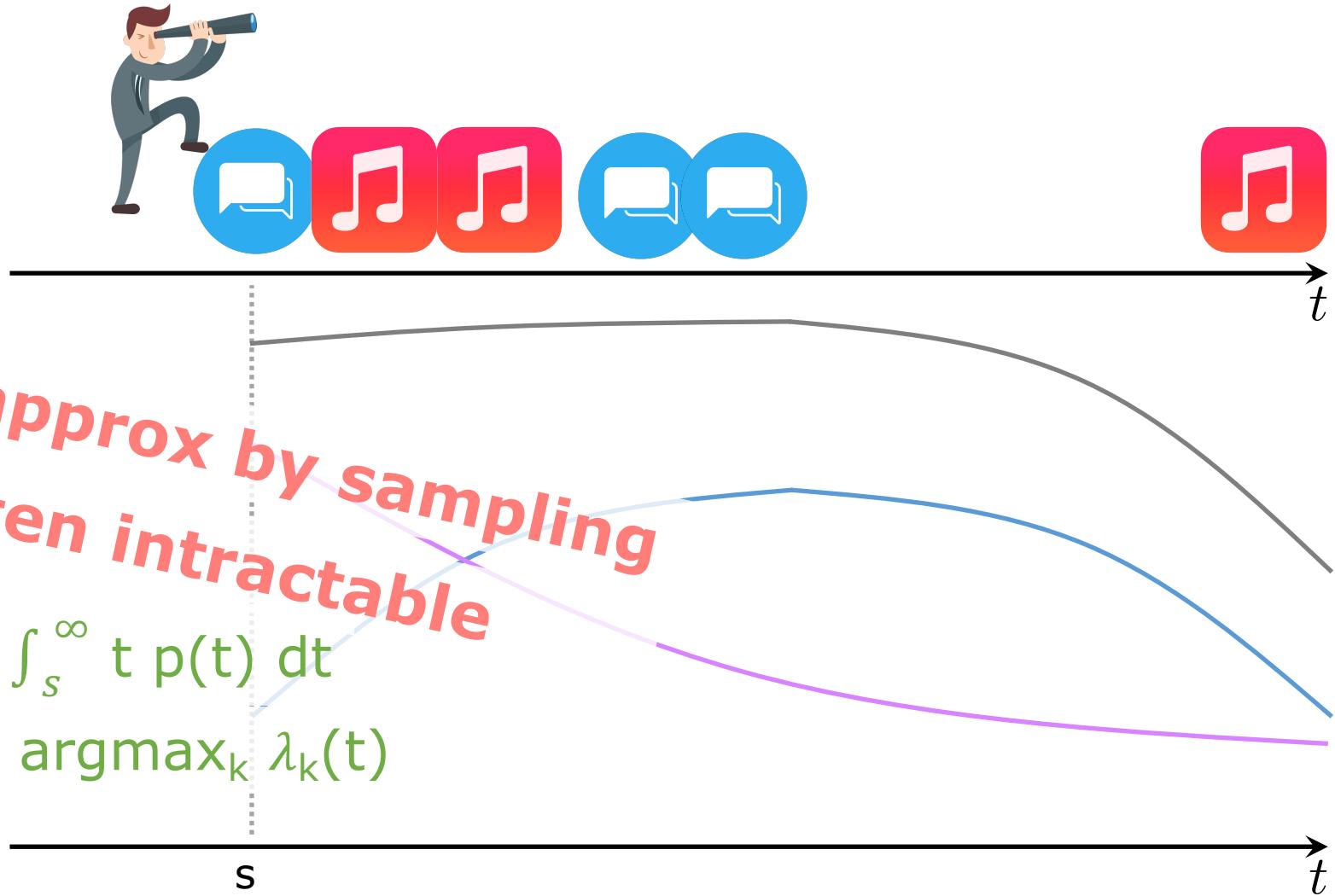
Predict future by sampling



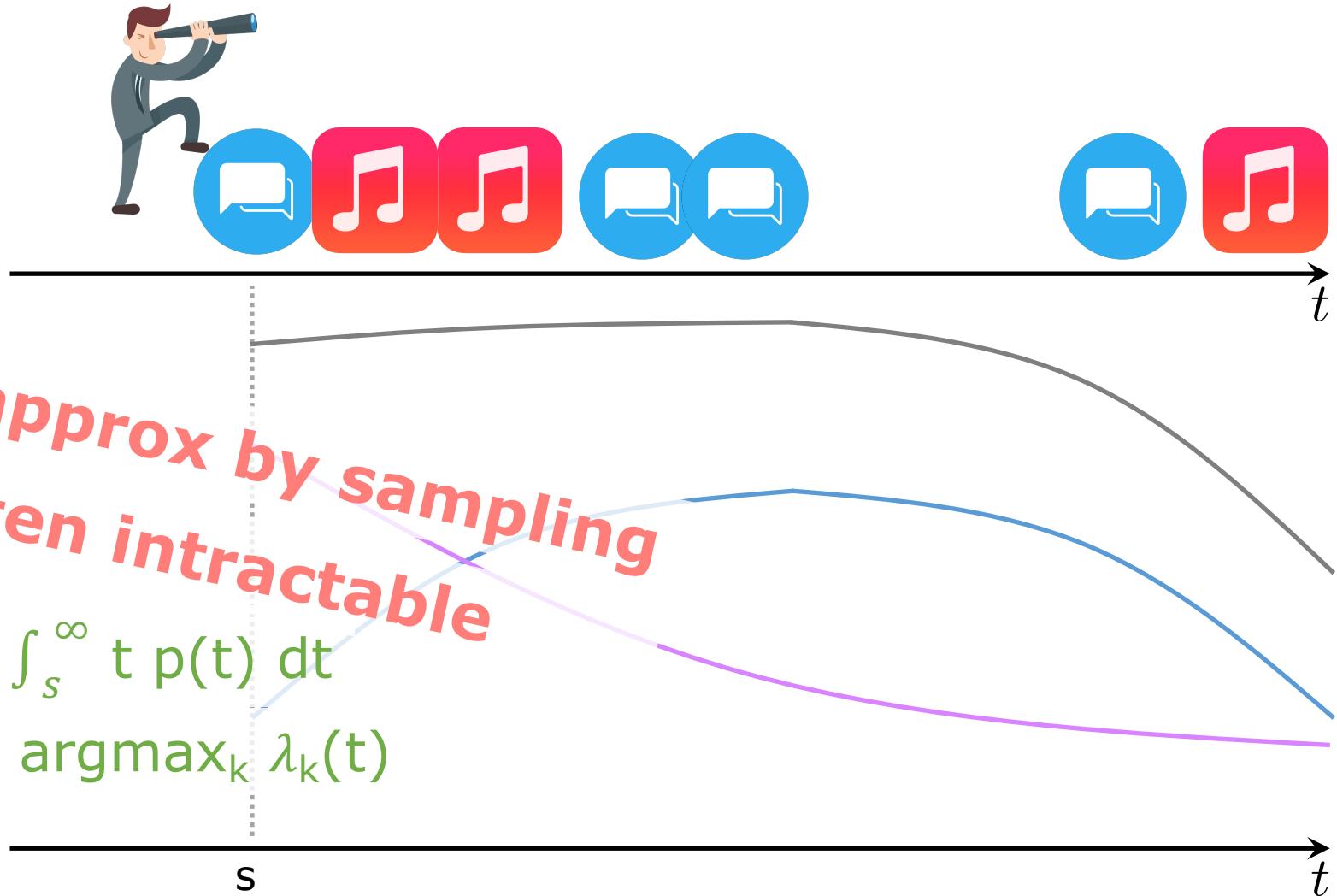
Predict future by sampling



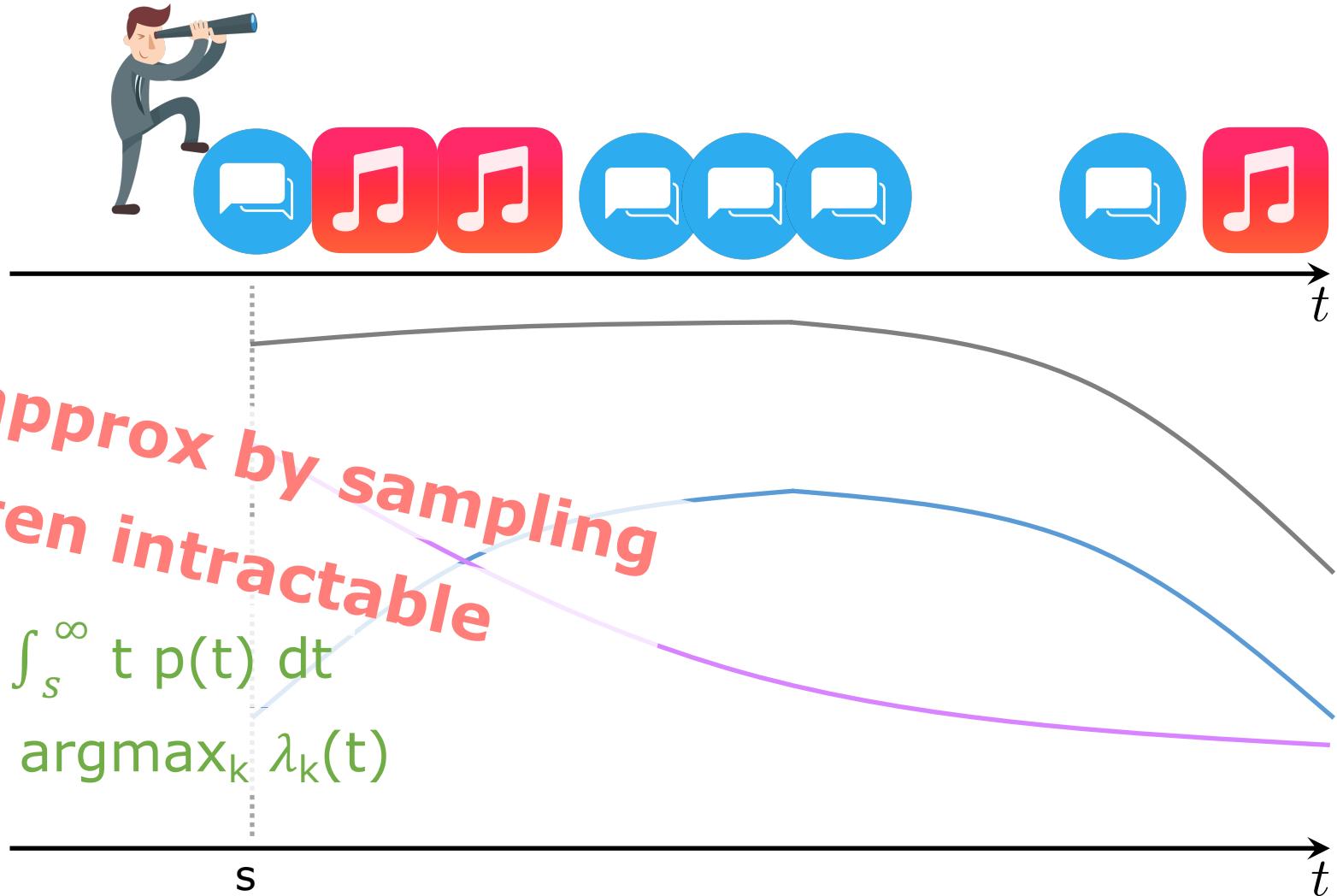
Predict future by sampling



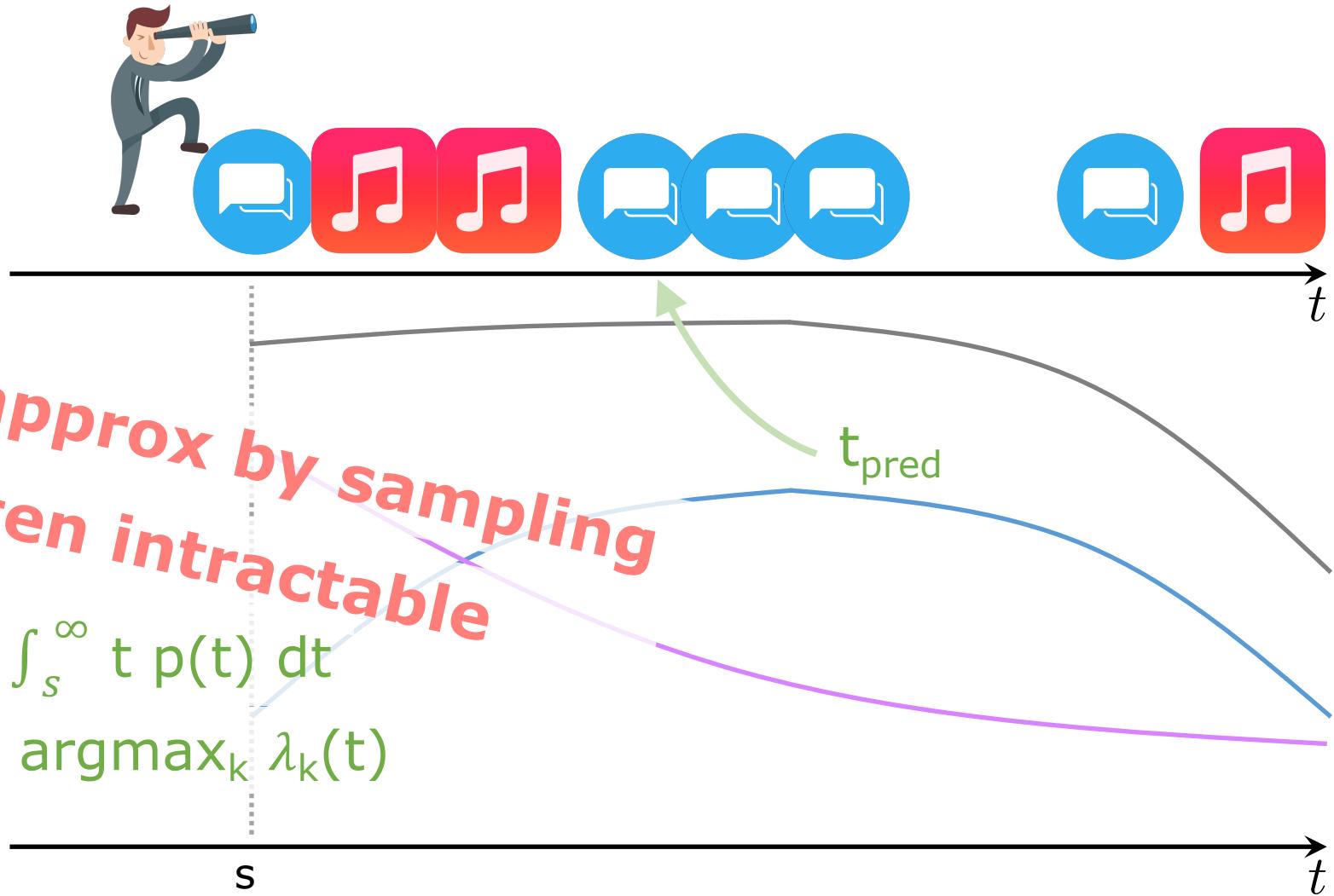
Predict future by sampling



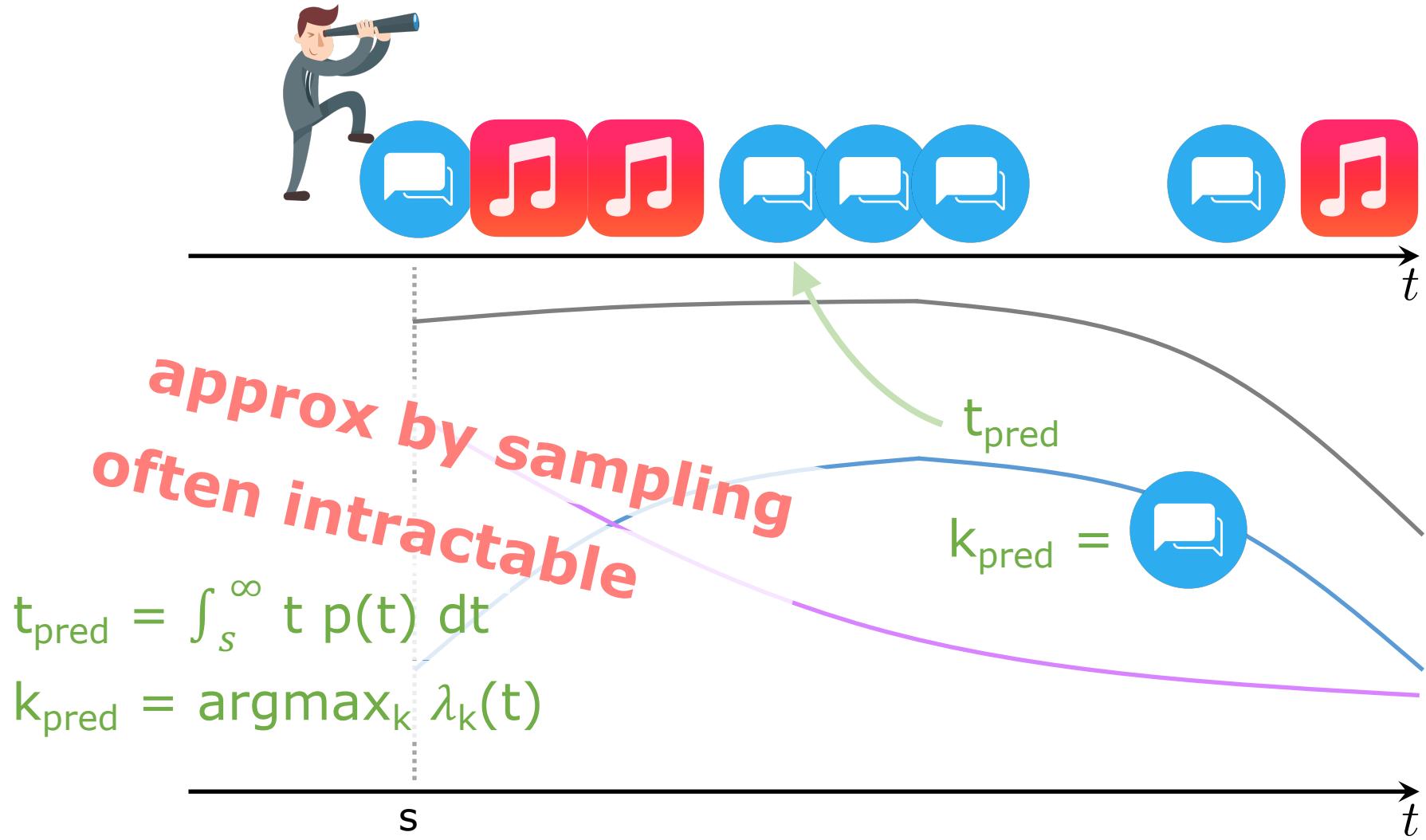
Predict future by sampling



Predict future by sampling



Predict future by sampling



<http://bburl/tpp-slides-p2>
<http://bburl/tpp-lab-p2>

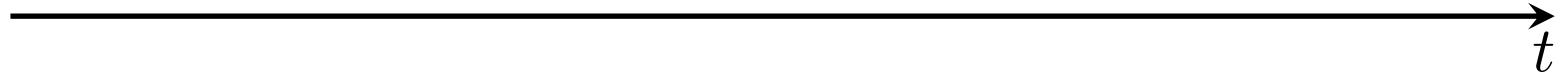
Any Questions?

<http://bburl/tpp-slides-p2>

Hawkes process

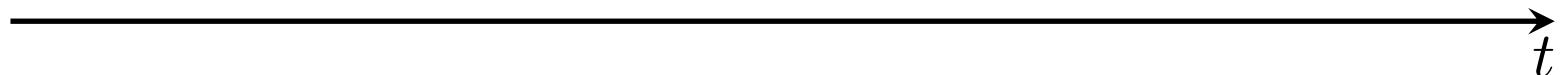
Music

Social



λ_{music}

λ_{social}



Hawkes process

Music

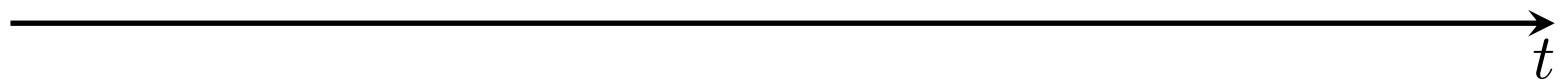


Social



λ_{music}

λ_{social}

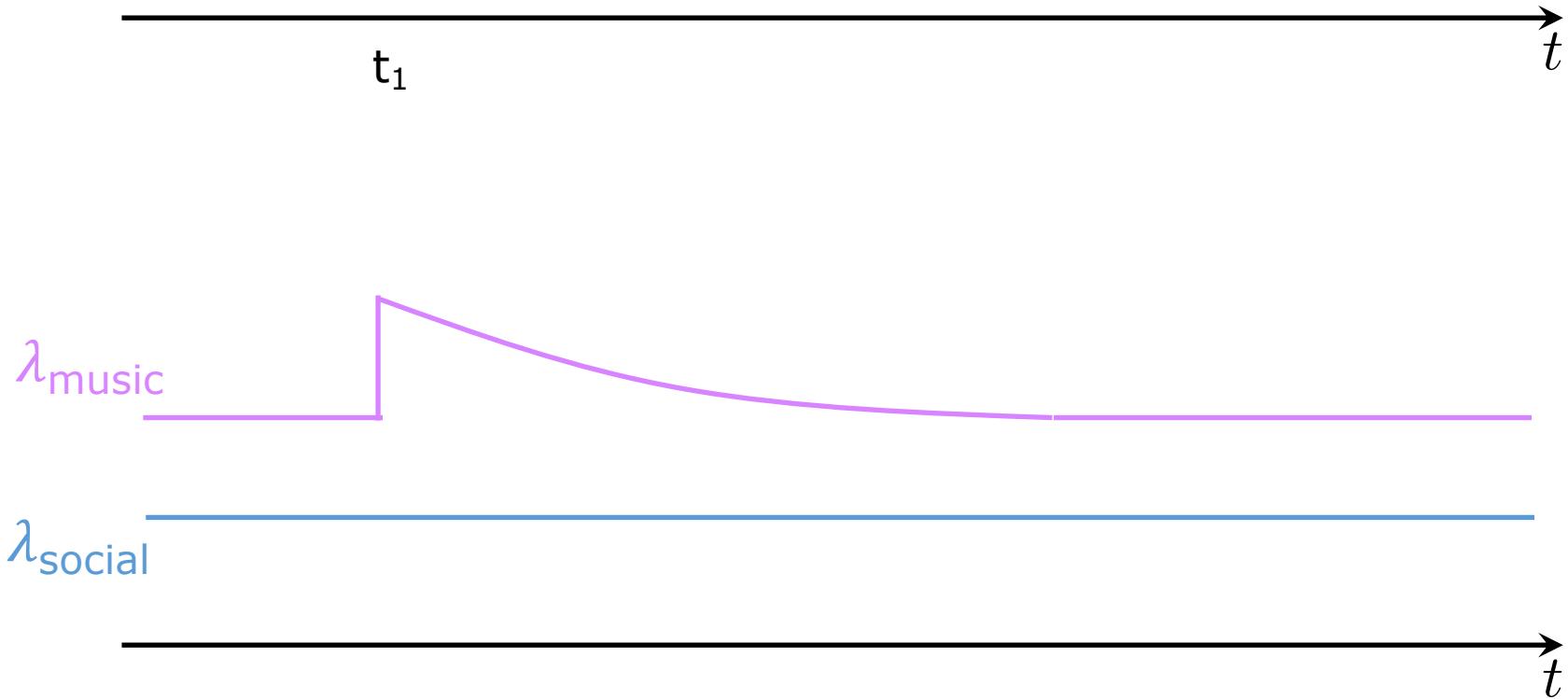


Hawkes process

Music



Social

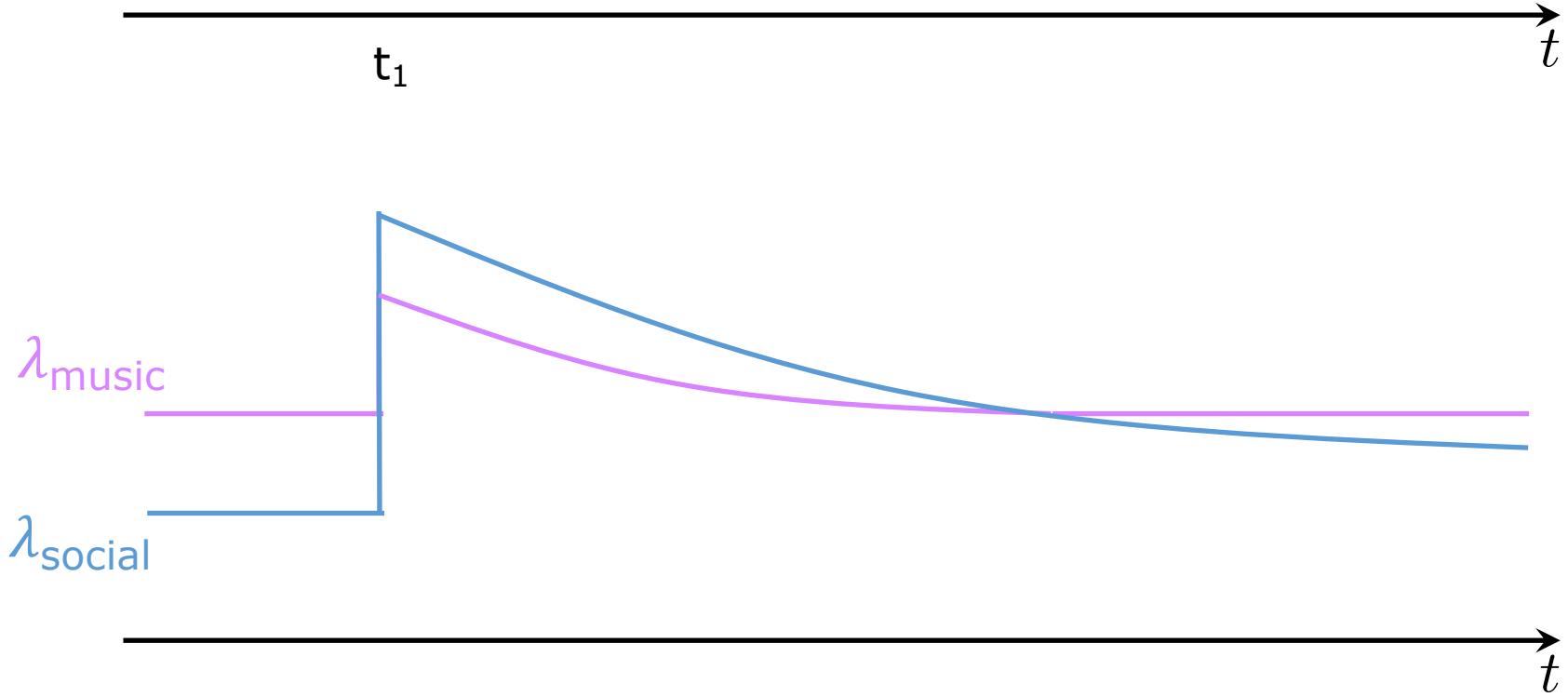


Hawkes process

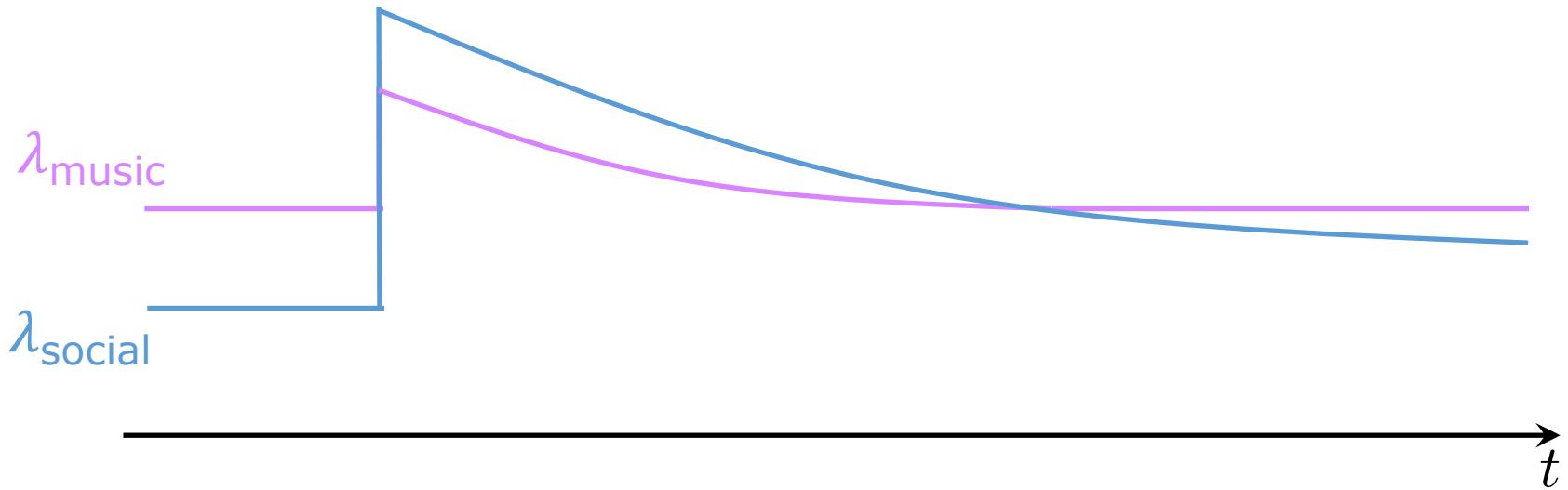
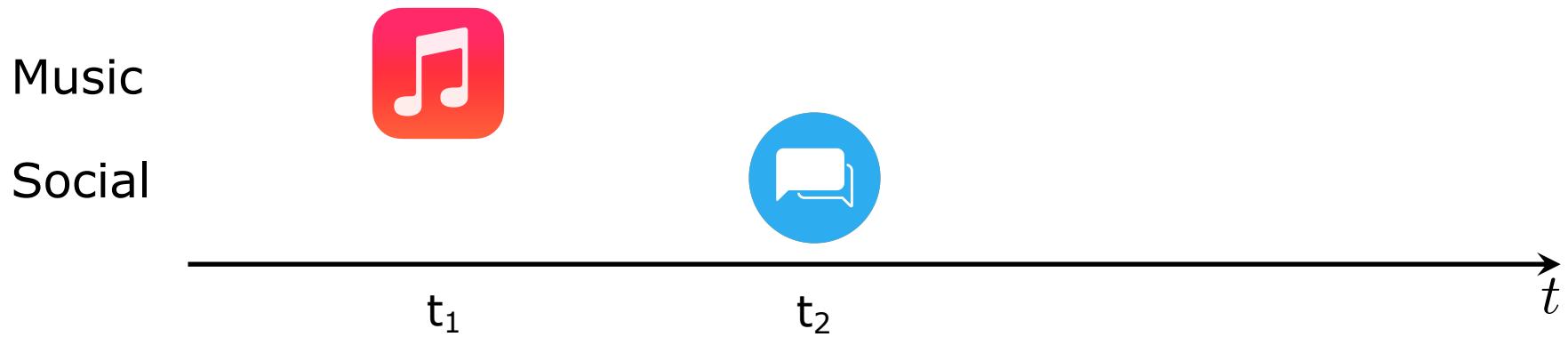
Music



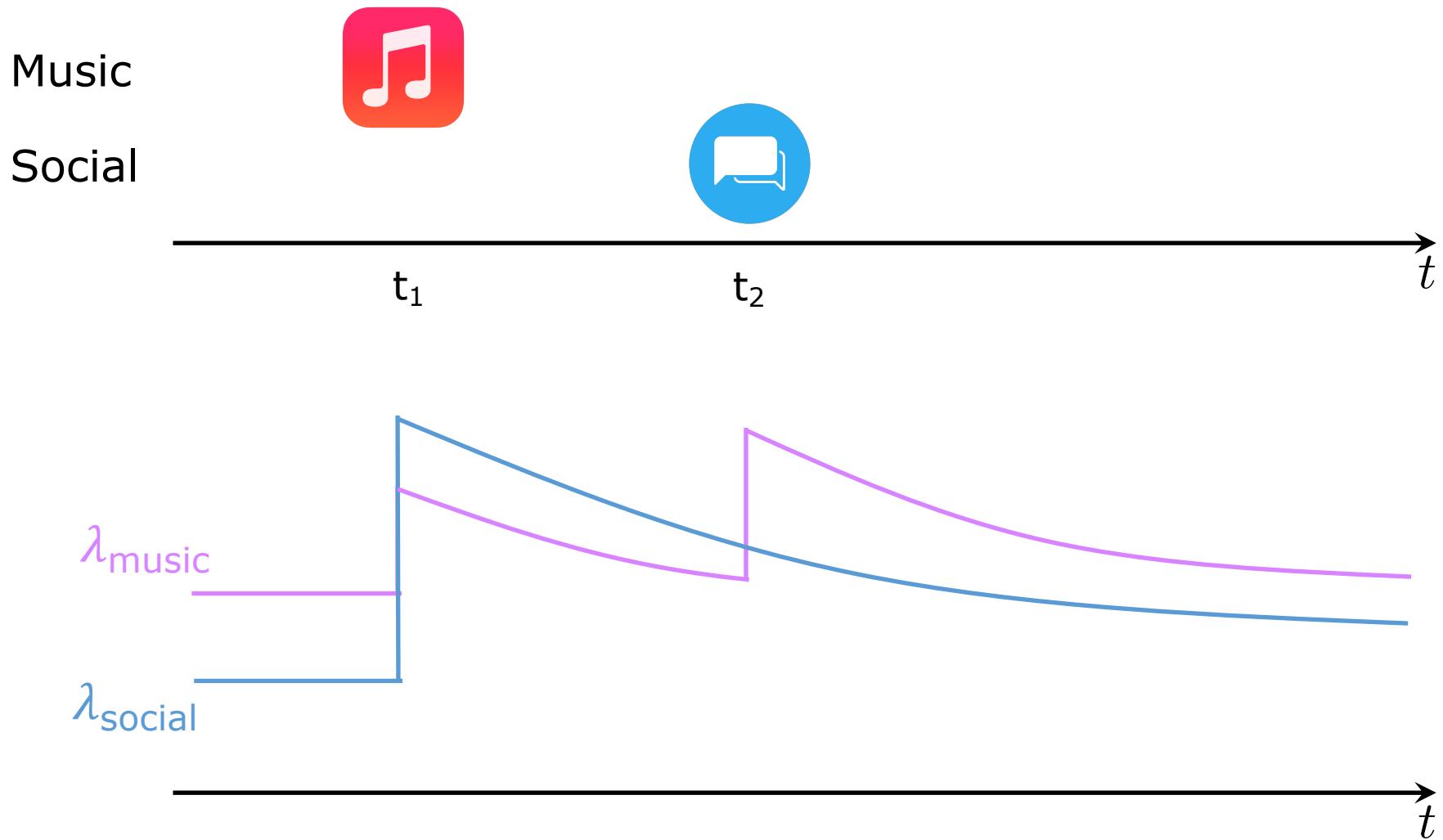
Social



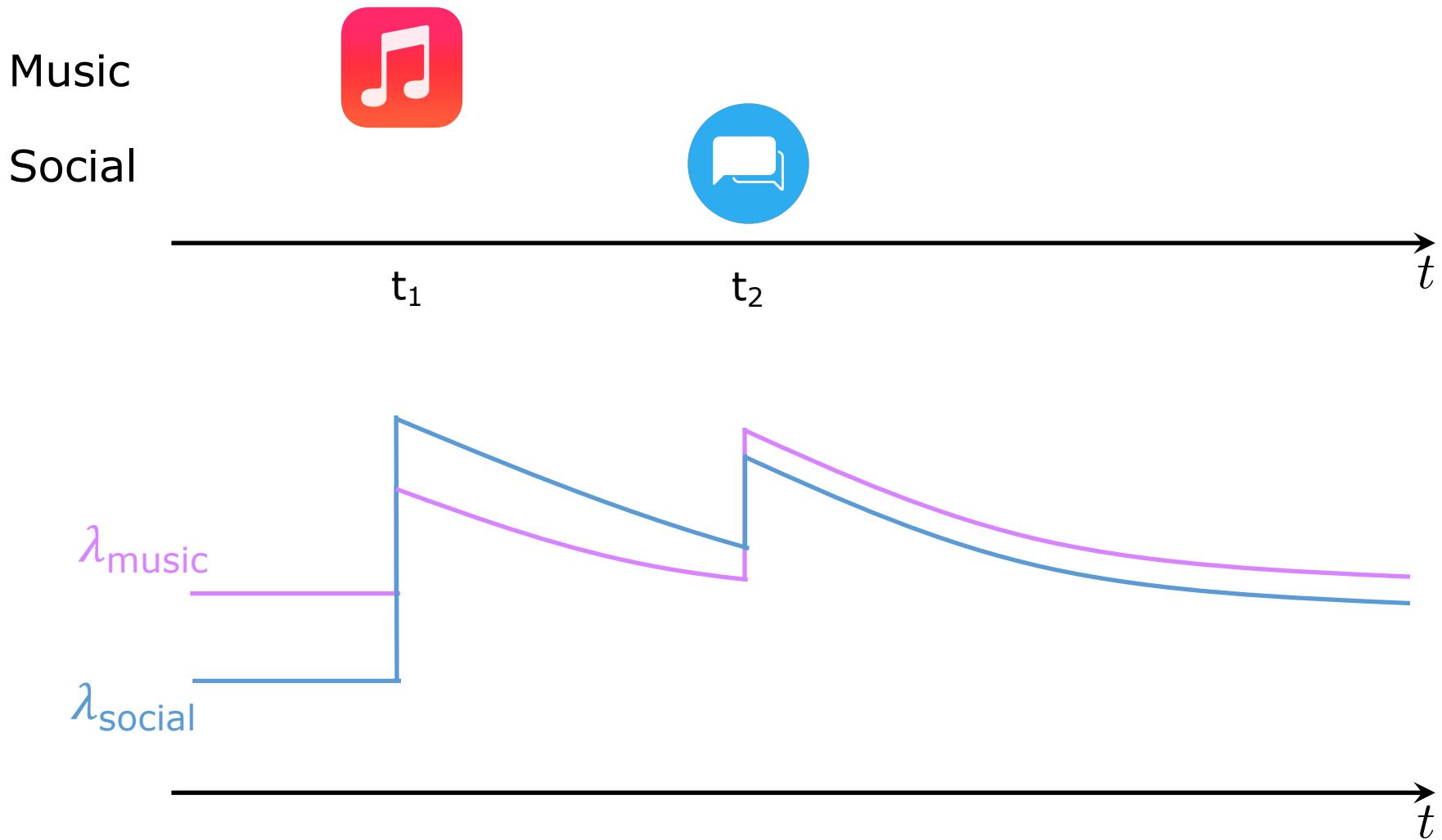
Hawkes process



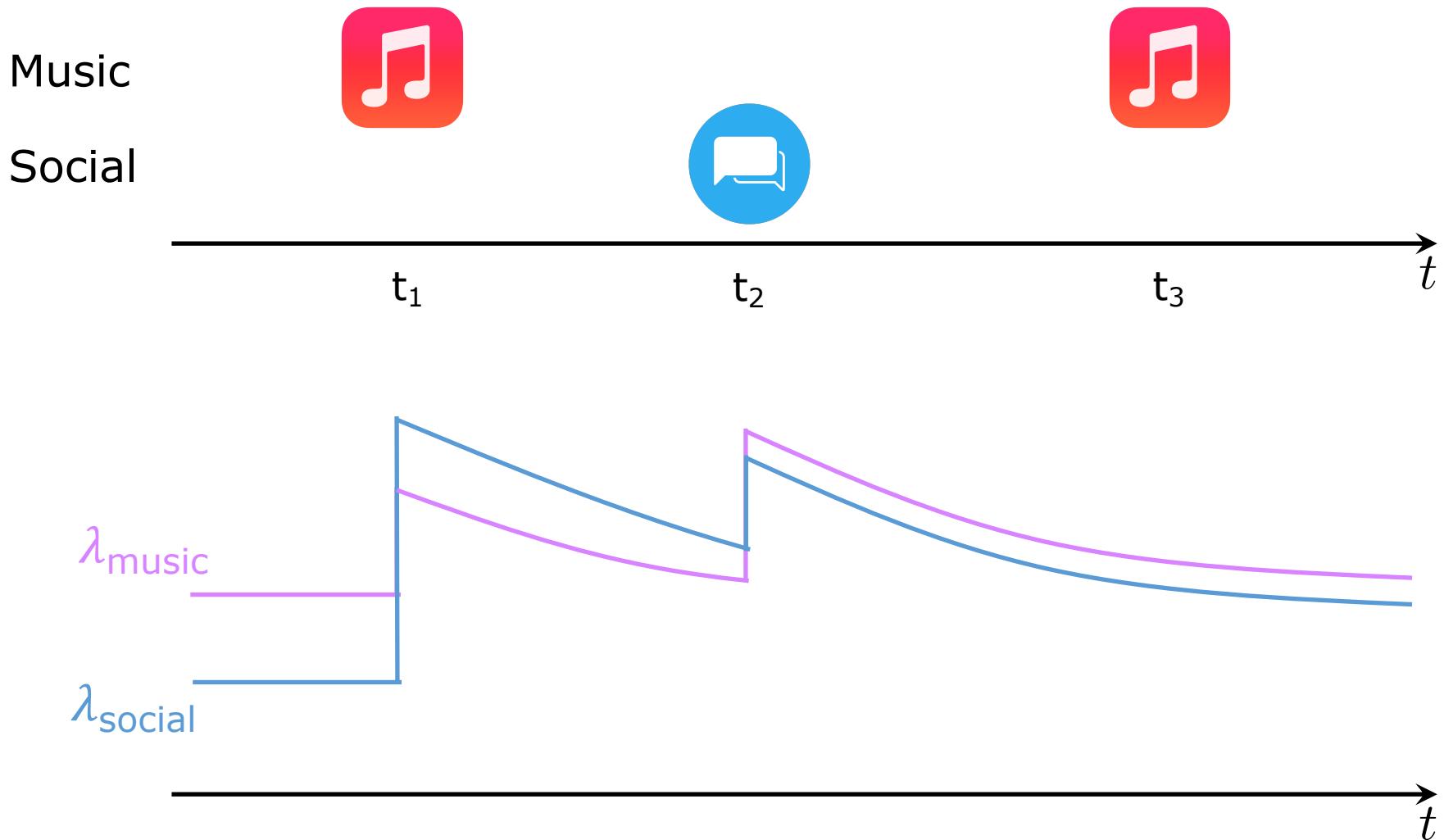
Hawkes process



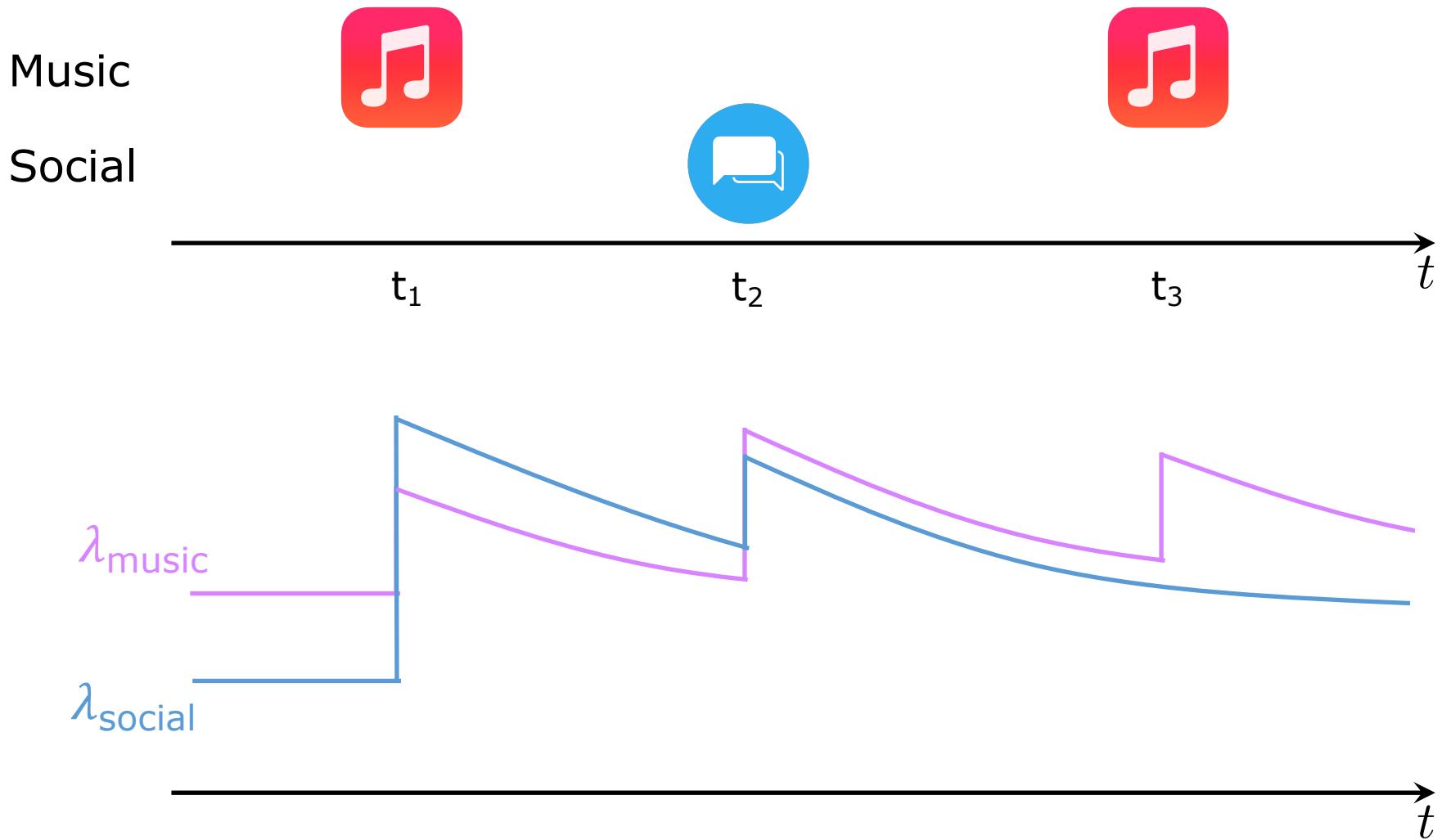
Hawkes process



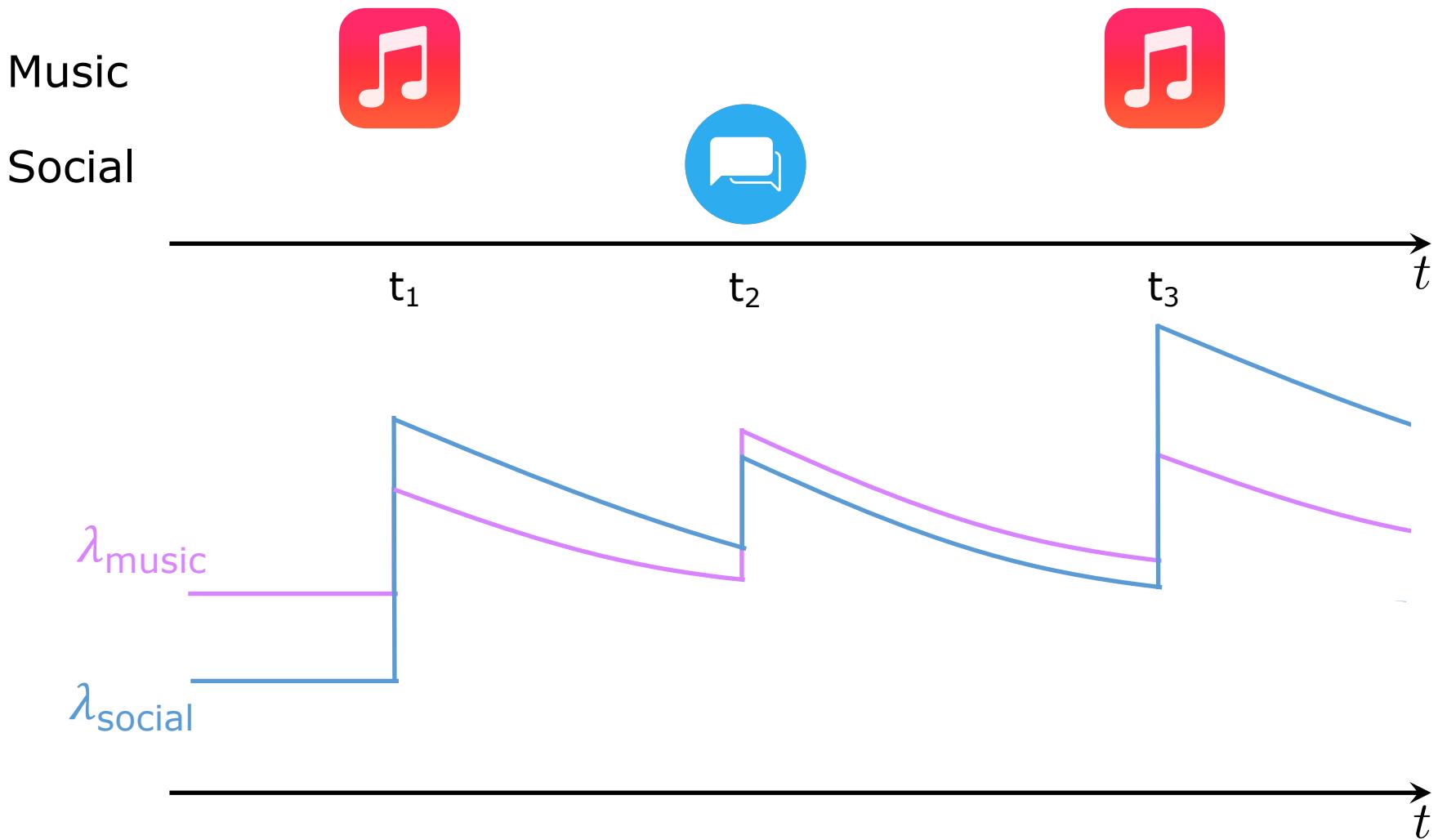
Hawkes process



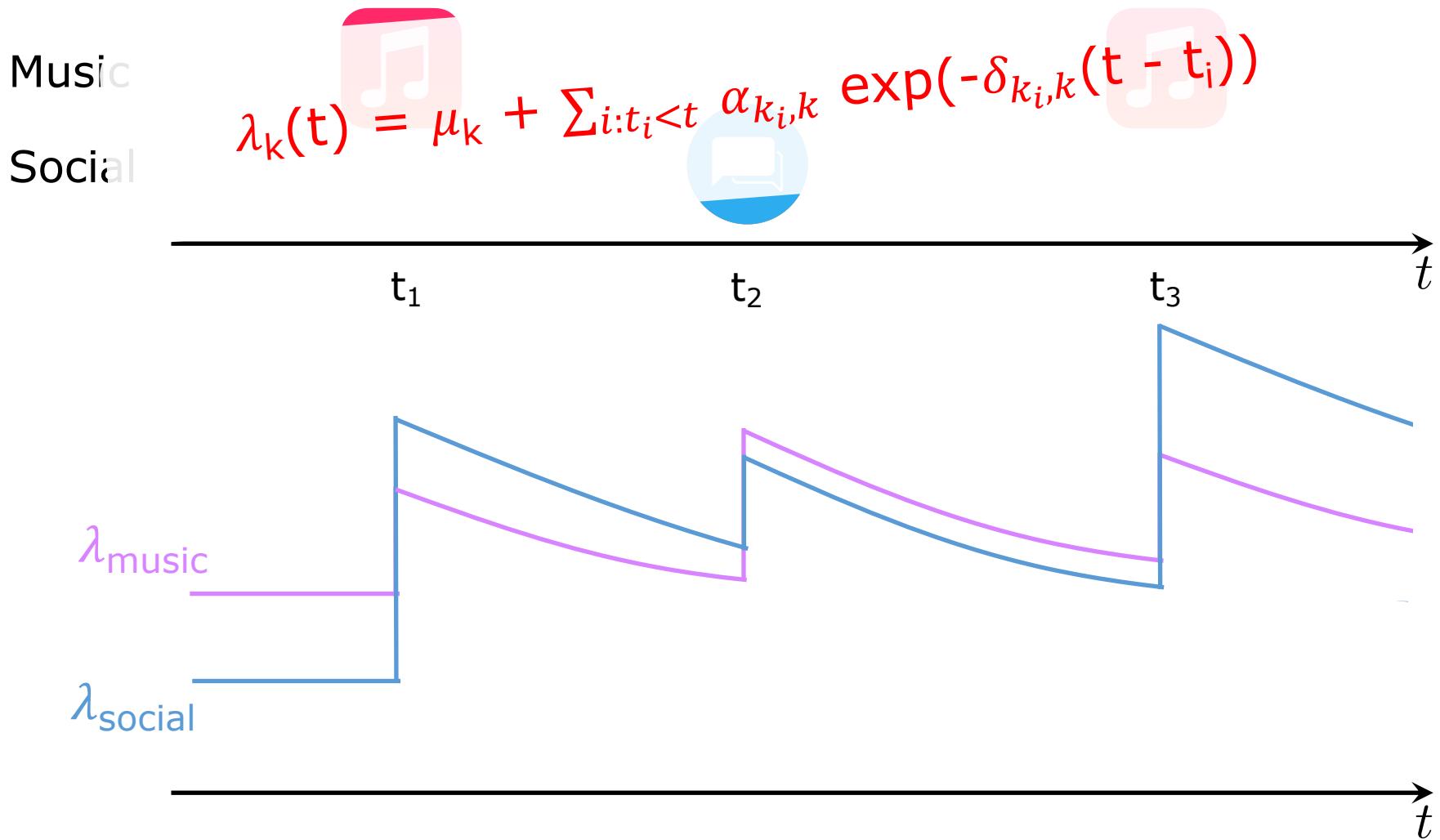
Hawkes process



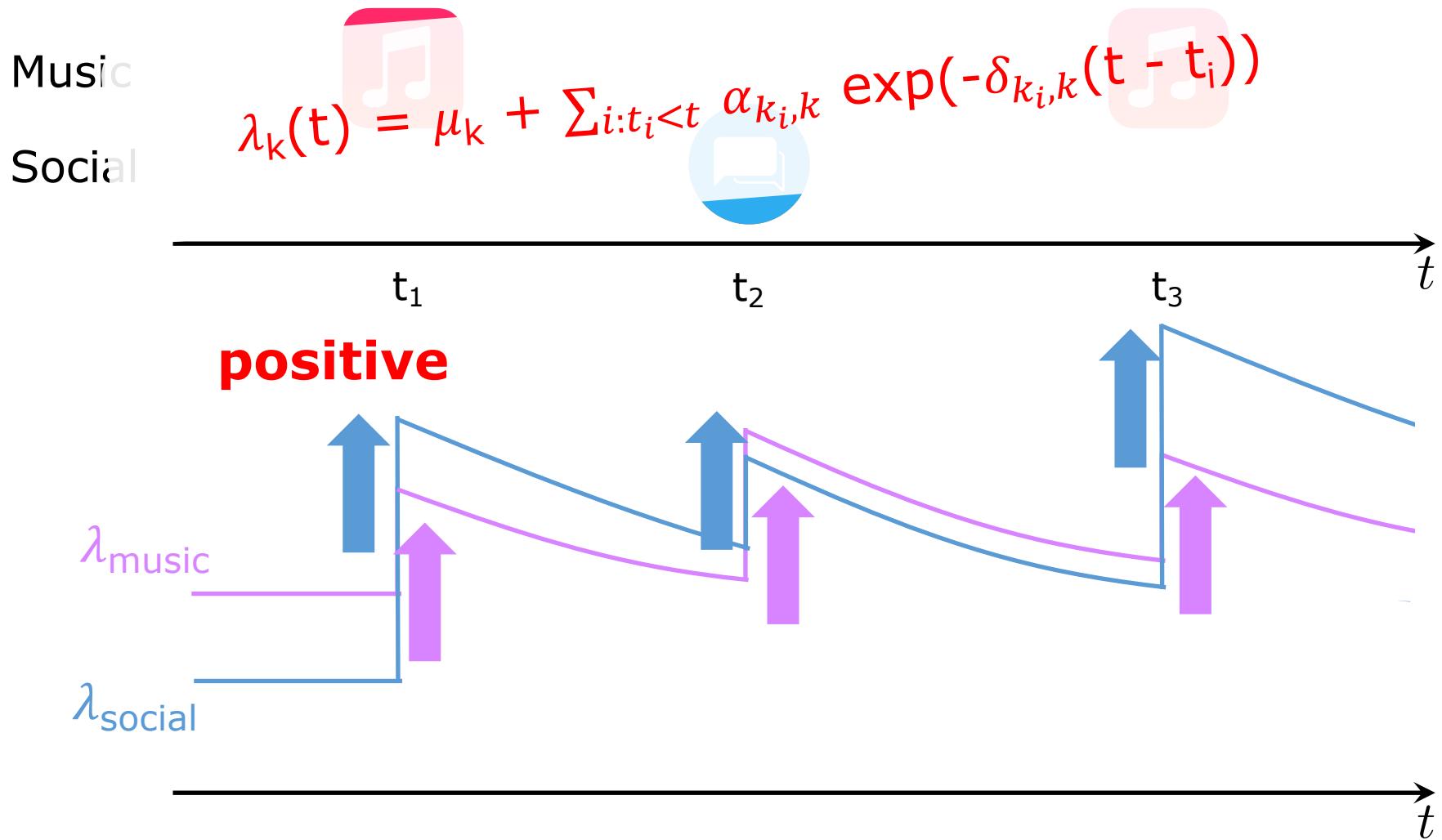
Hawkes process



Hawkes process



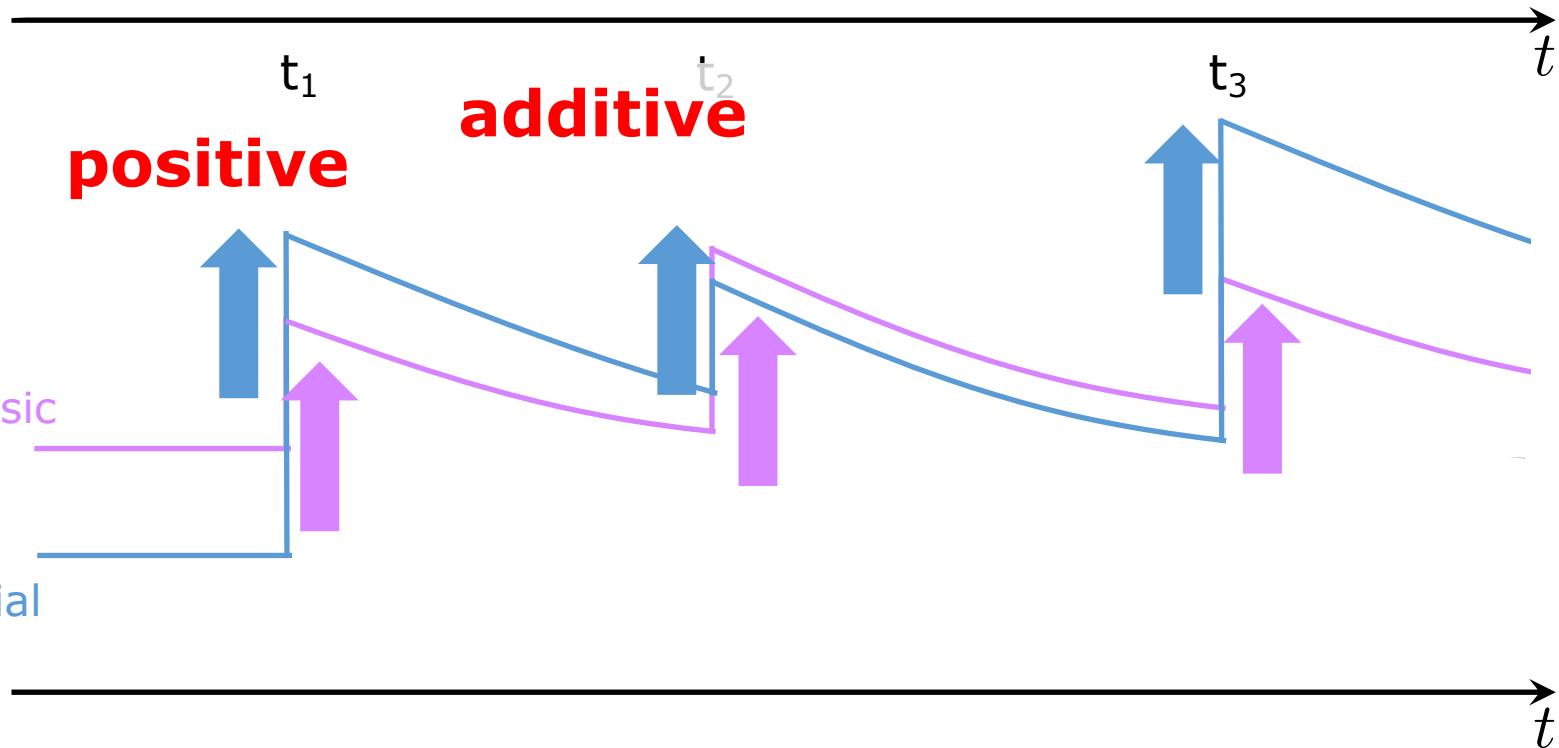
Hawkes process



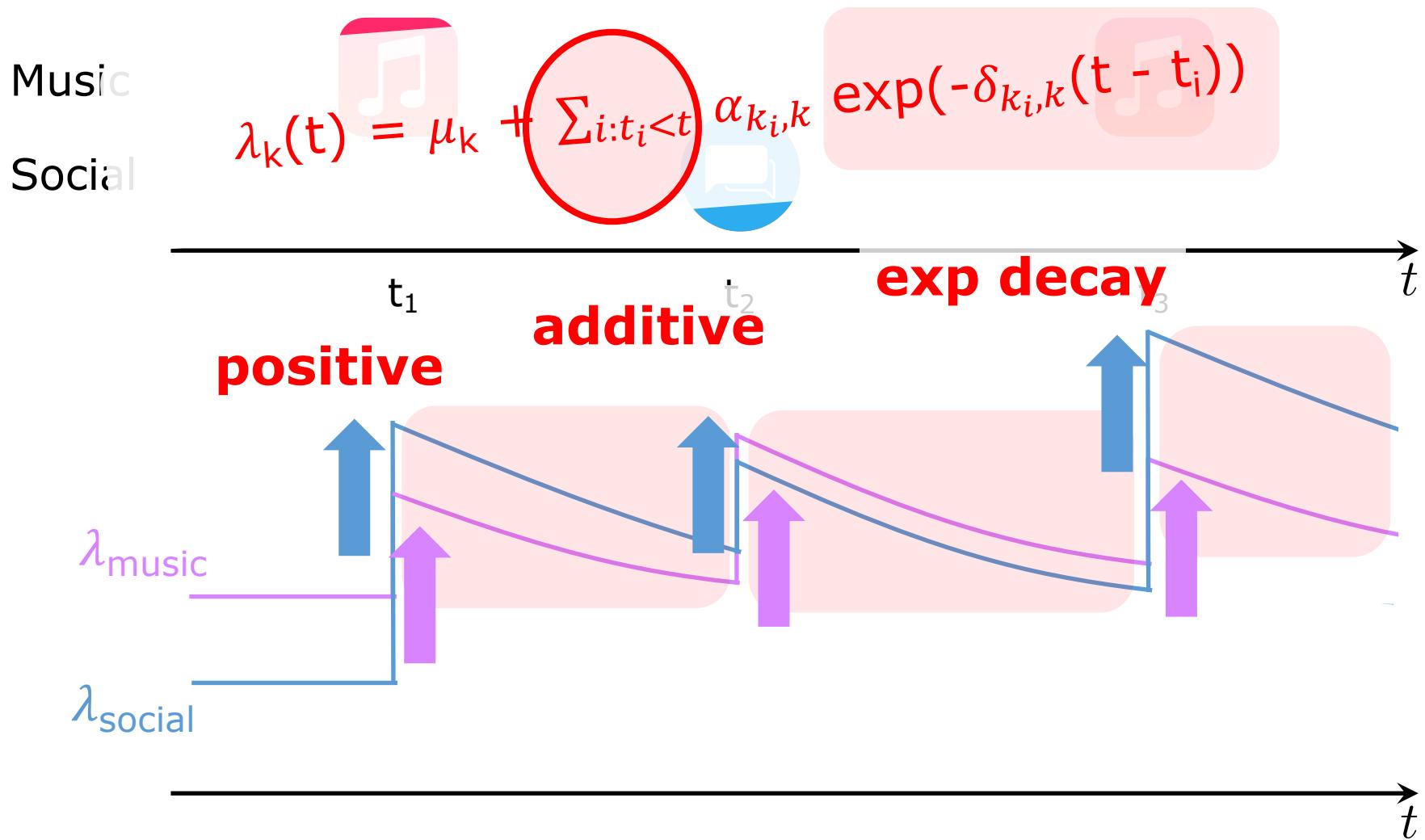
Hawkes process

Music
Social

$$\lambda_k(t) = \mu_k + \sum_{i:t_i < t} \alpha_{k_i, k} \exp(-\delta_{k_i, k}(t - t_i))$$



Hawkes process



Recap of language models



Recap of language models



I like working at Bloomberg

Recap of language models



$P(\text{working} \mid \text{I like}) = 0.01$

$P(\text{at} \mid \text{like working}) = 0.01$

$P(\text{Bloomberg} \mid \text{working at}) = 0.12$

I like working at Bloomberg

Recap of language models



$P(\text{working} \mid \text{I like}) = 0.01$

$P(\text{at} \mid \text{like working}) = 0.01$

$P(\text{Bloomberg} \mid \text{working at}) = 0.12$



Recap of language models

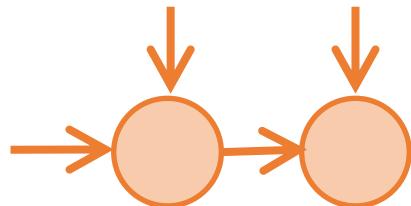


$P(\text{working} \mid \text{I like}) = 0.01$

$P(\text{at} \mid \text{like working}) = 0.01$

$P(\text{Bloomberg} \mid \text{working at}) = 0.12$

I like working at Bloomberg



Recap of language models



$$P(\text{working} \mid \text{I like}) = 0.01$$

$$P(\text{at} \mid \text{like working}) = 0.01$$

$$P(\text{Bloomberg} \mid \text{working at}) = 0.12$$



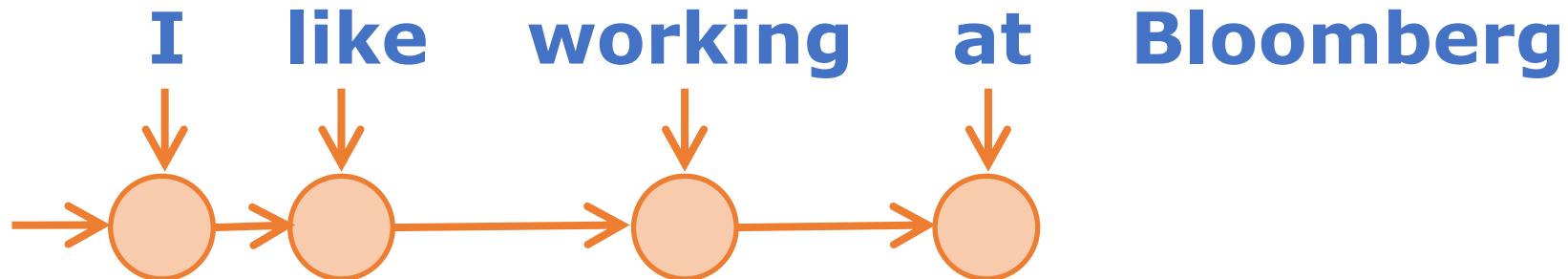
Recap of language models



$$P(\text{working} \mid \text{I like}) = 0.01$$

$$P(\text{at} \mid \text{like working}) = 0.01$$

$$P(\text{Bloomberg} \mid \text{working at}) = 0.12$$



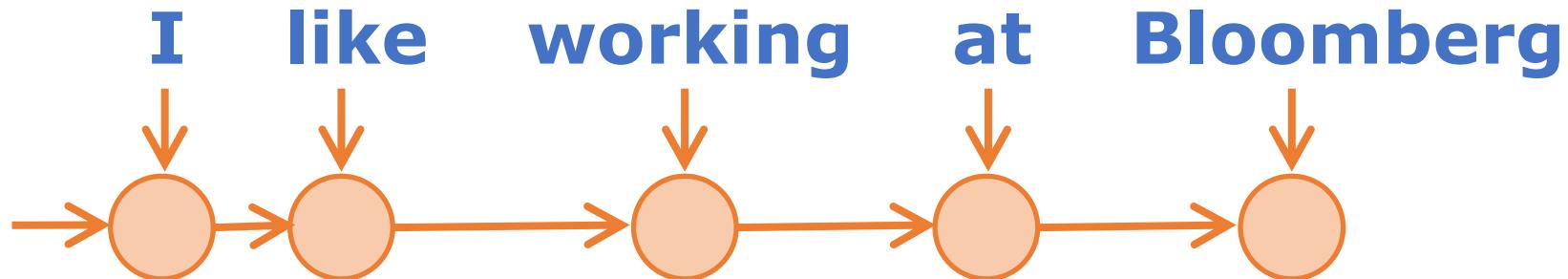
Recap of language models



$$P(\text{working} \mid \text{I like}) = 0.01$$

$$P(\text{at} \mid \text{like working}) = 0.01$$

$$P(\text{Bloomberg} \mid \text{working at}) = 0.12$$



Use recurrent neural net (RNN)



Use recurrent neural net (RNN)



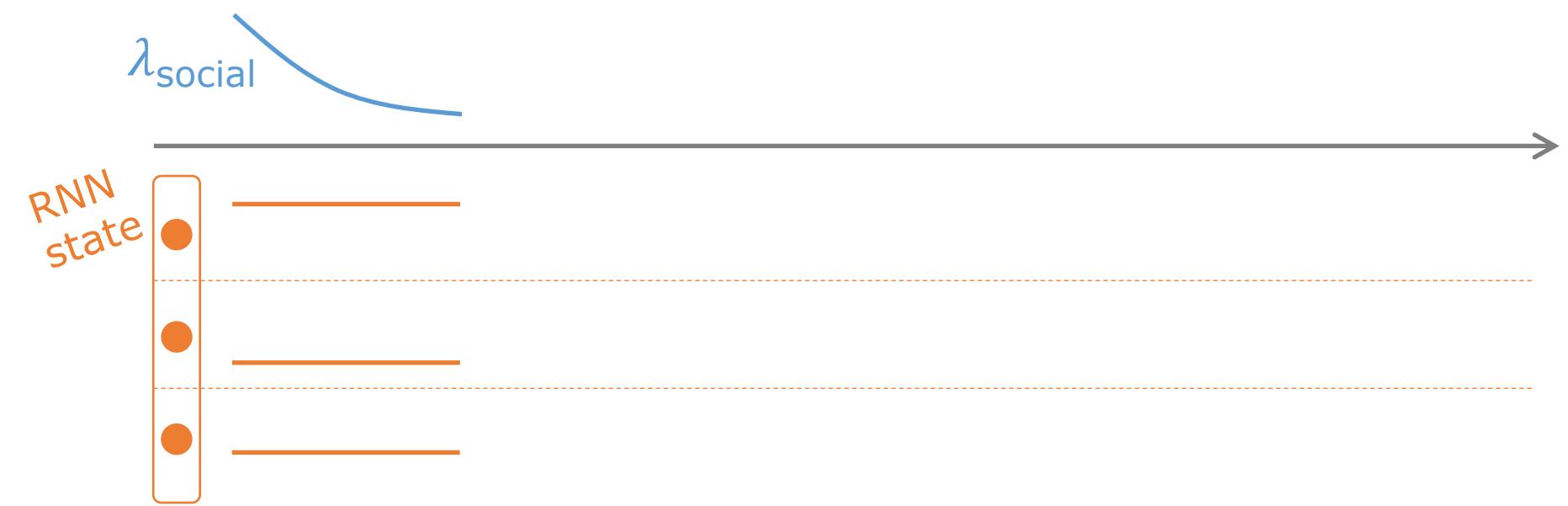
Use recurrent neural net (RNN)



Use recurrent neural net (RNN)



Use recurrent neural net (RNN)



Use recurrent neural net (RNN)



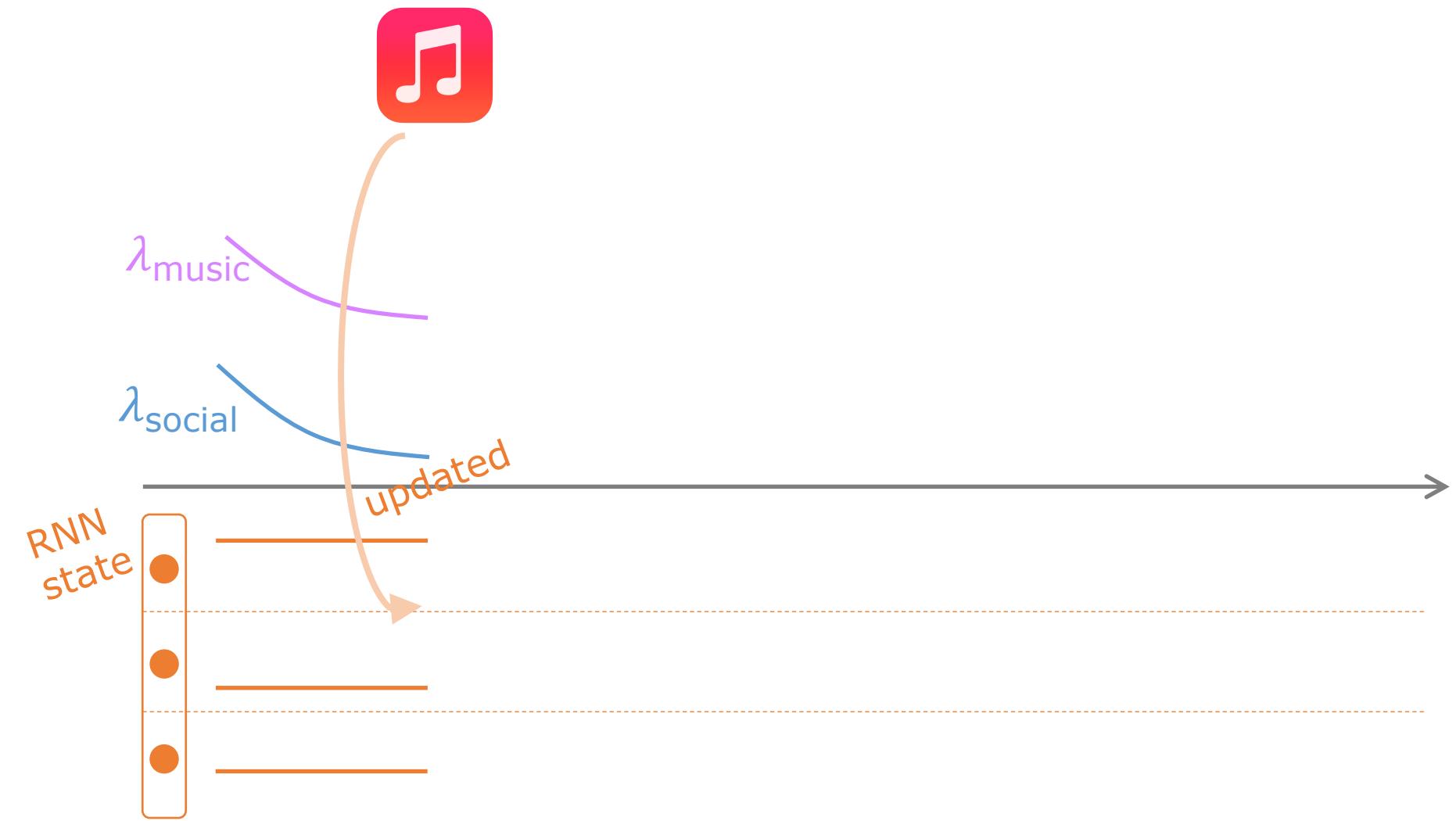
Use recurrent neural net (RNN)



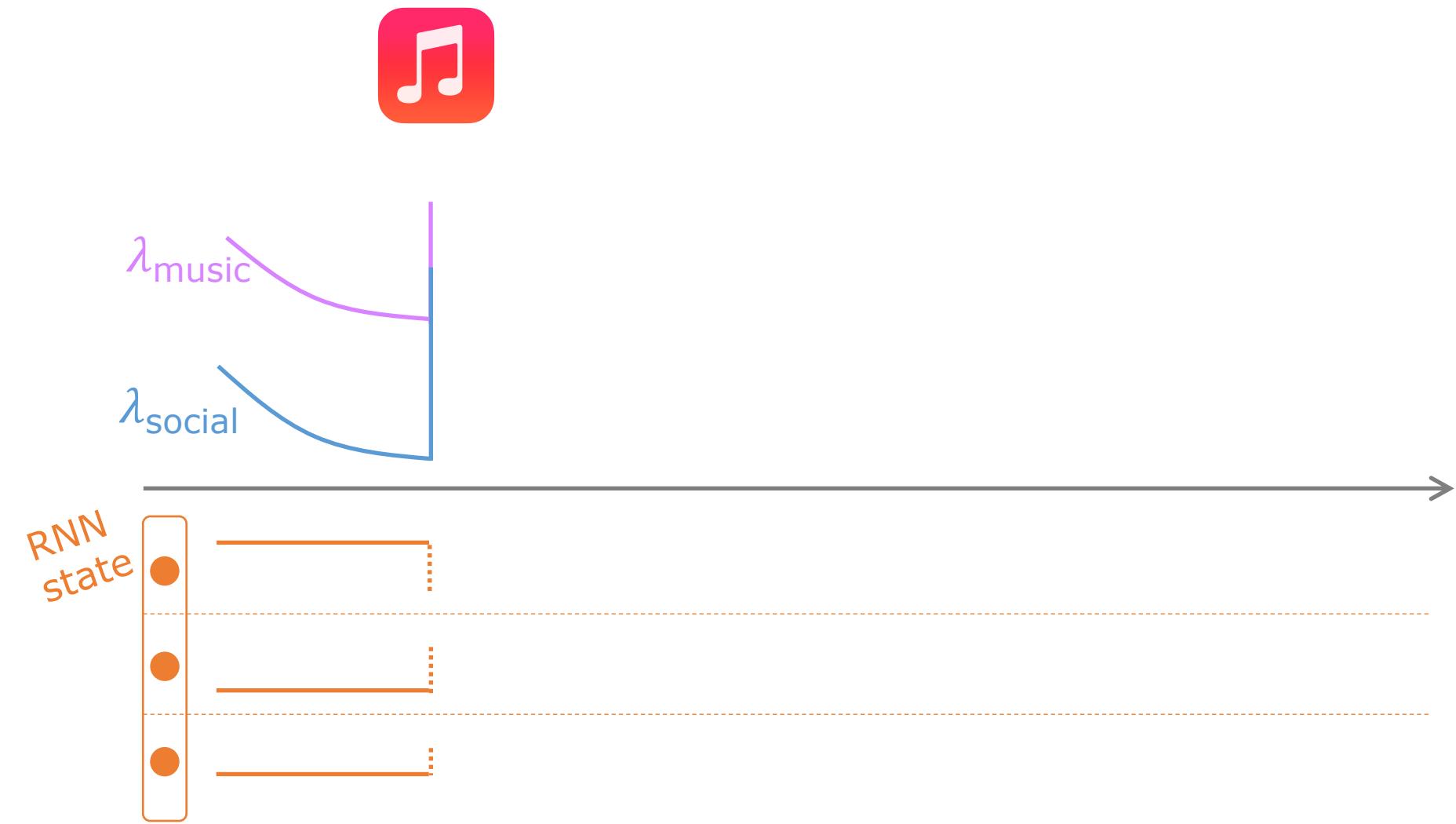
Use recurrent neural net (RNN)



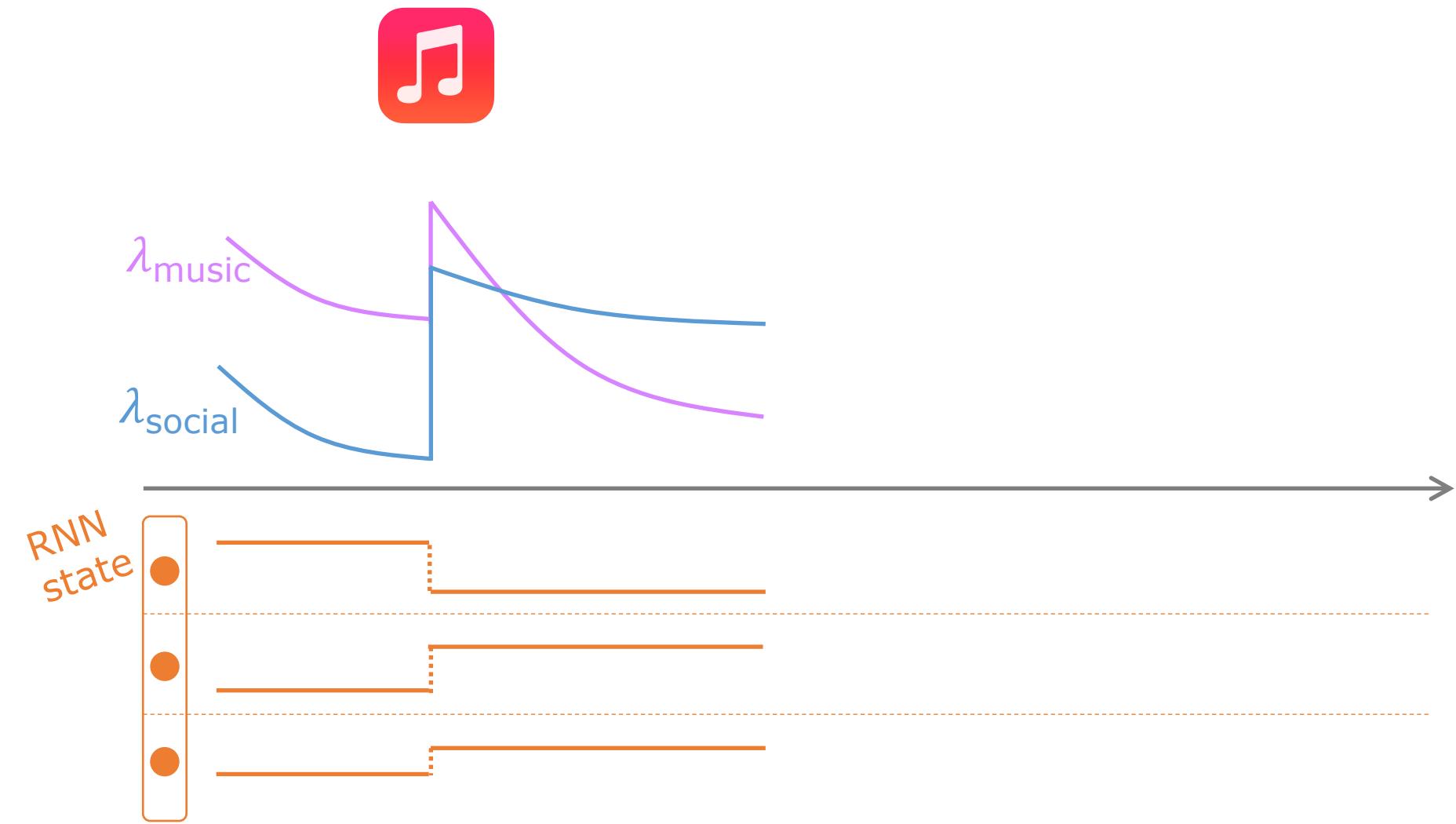
Use recurrent neural net (RNN)



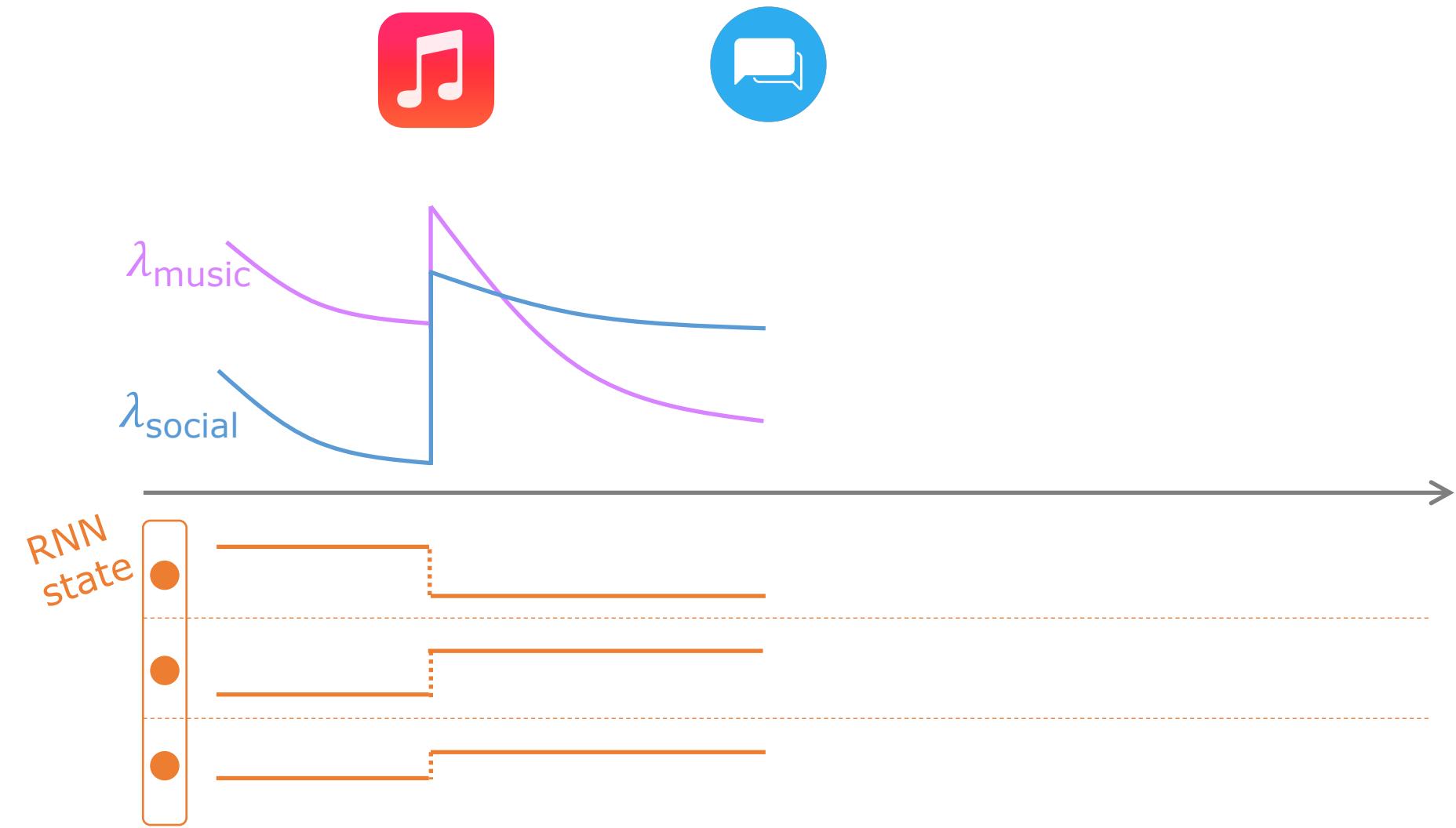
Use recurrent neural net (RNN)



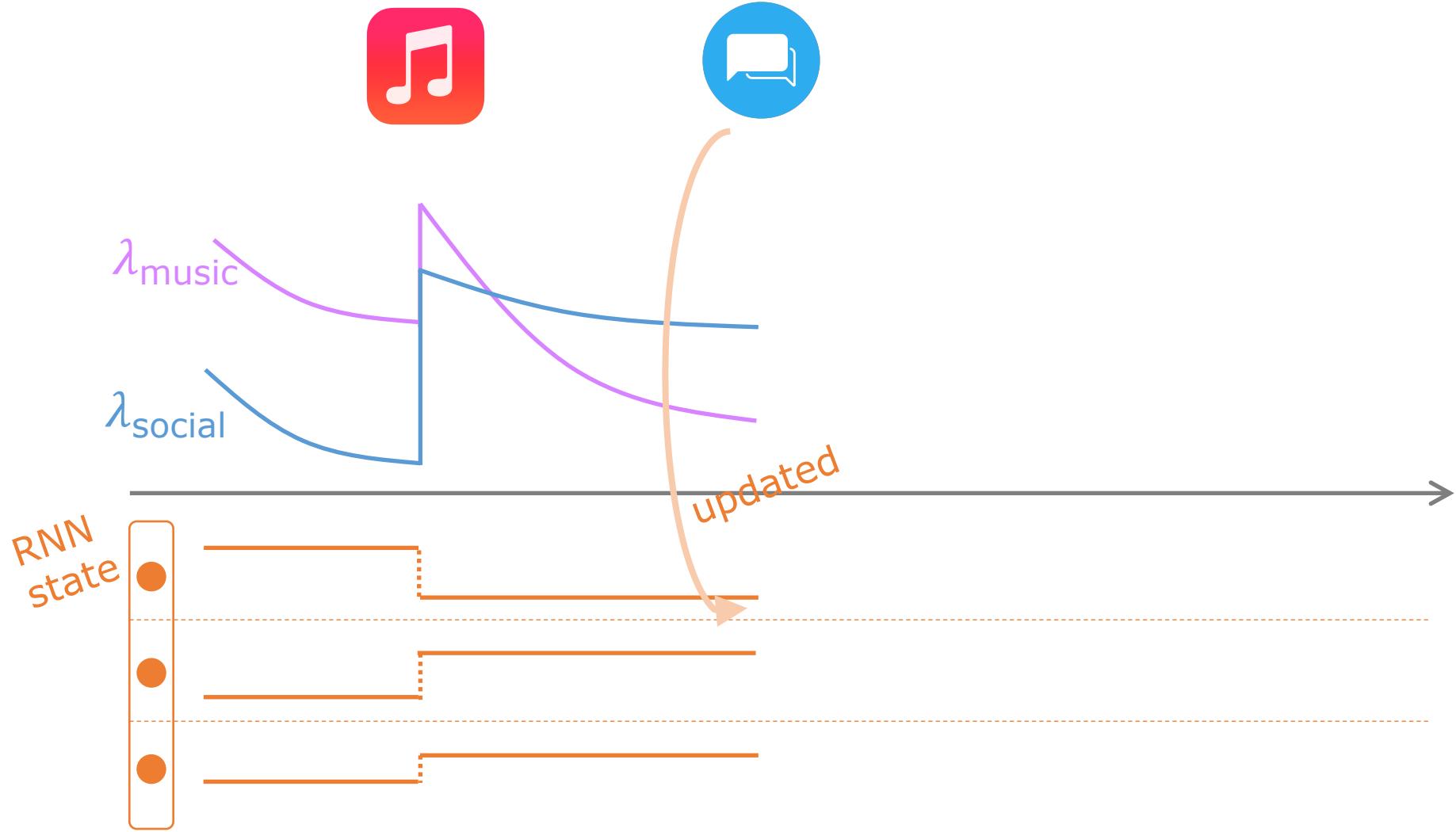
Use recurrent neural net (RNN)



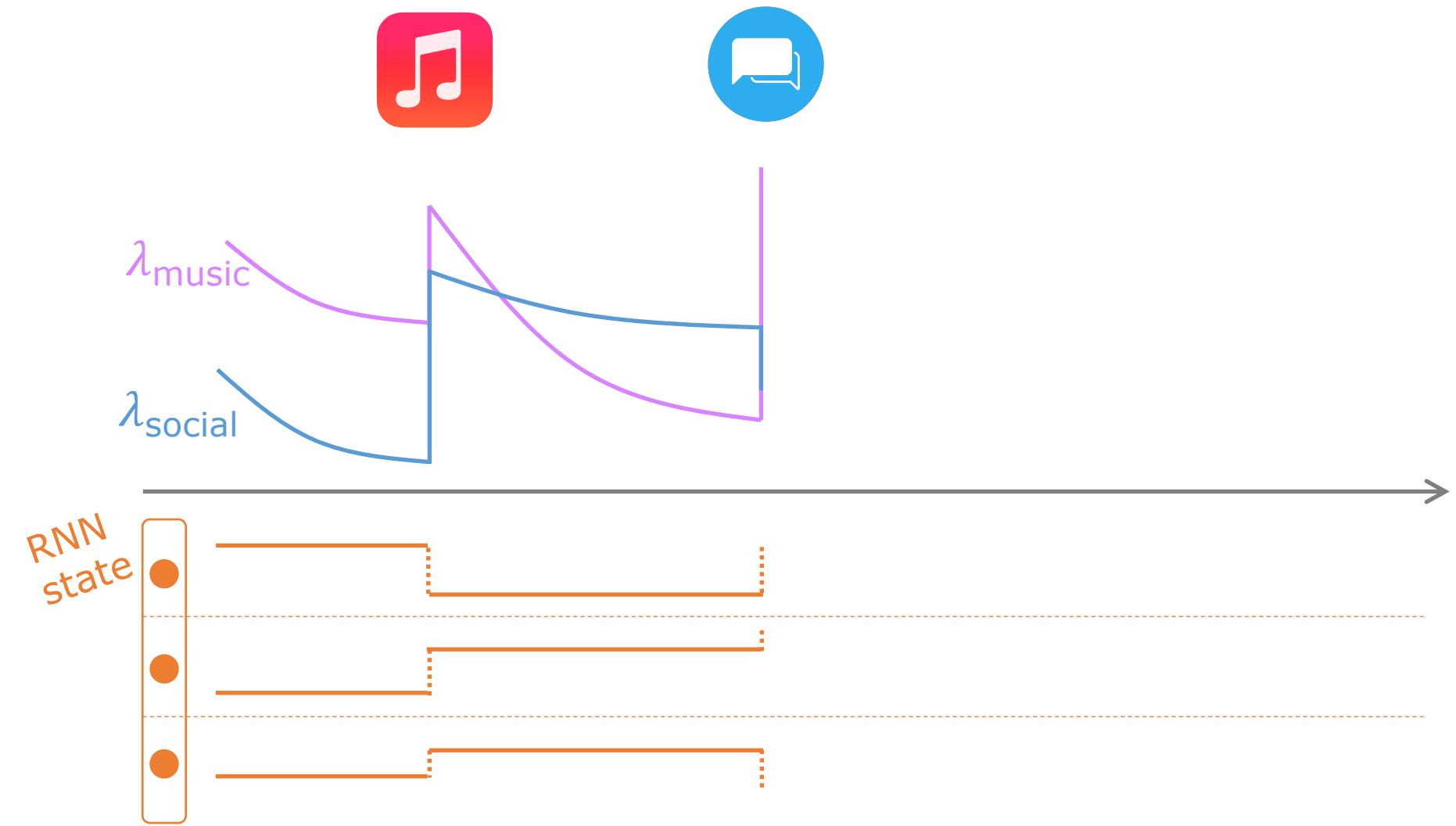
Use recurrent neural net (RNN)



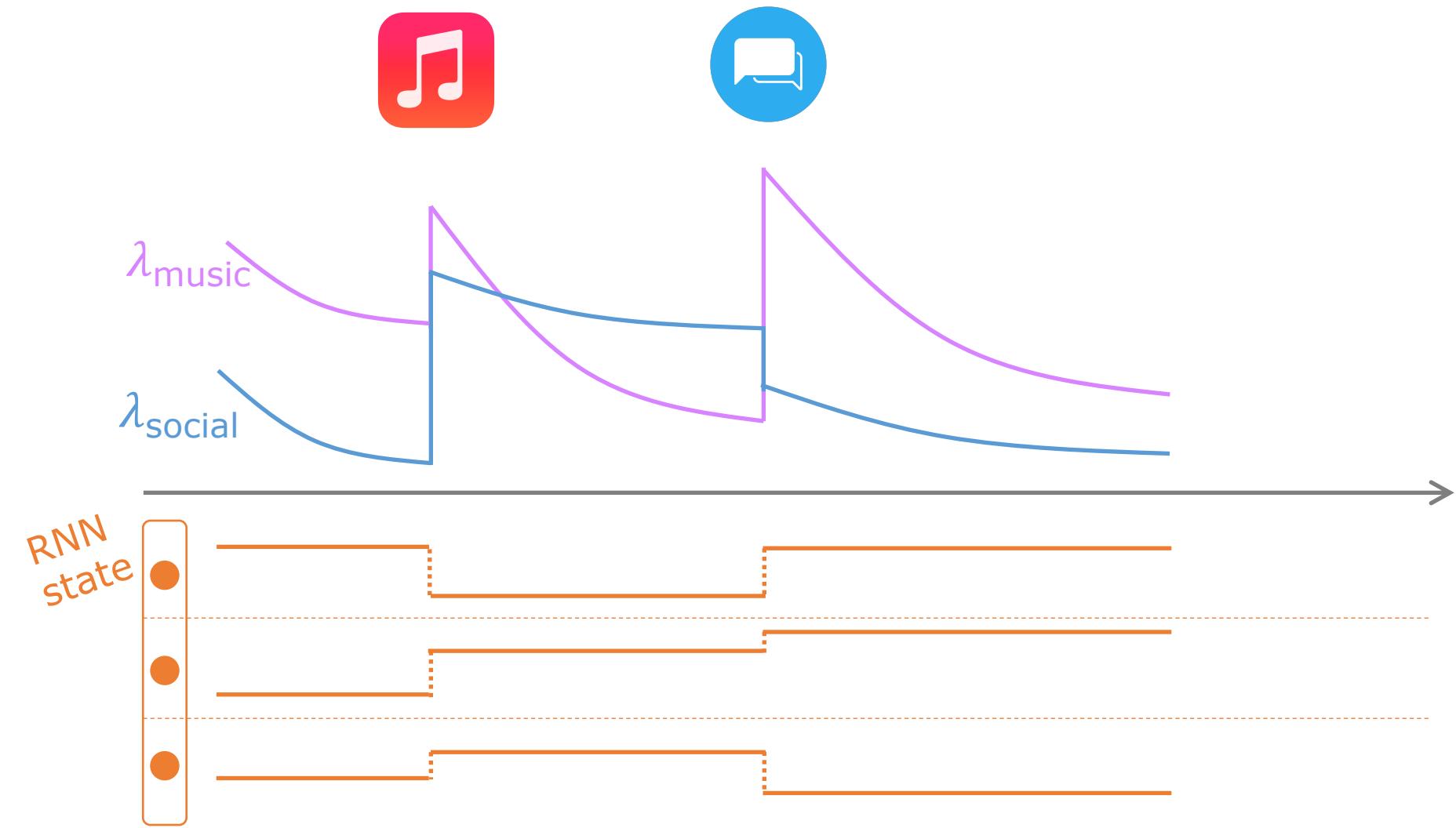
Use recurrent neural net (RNN)



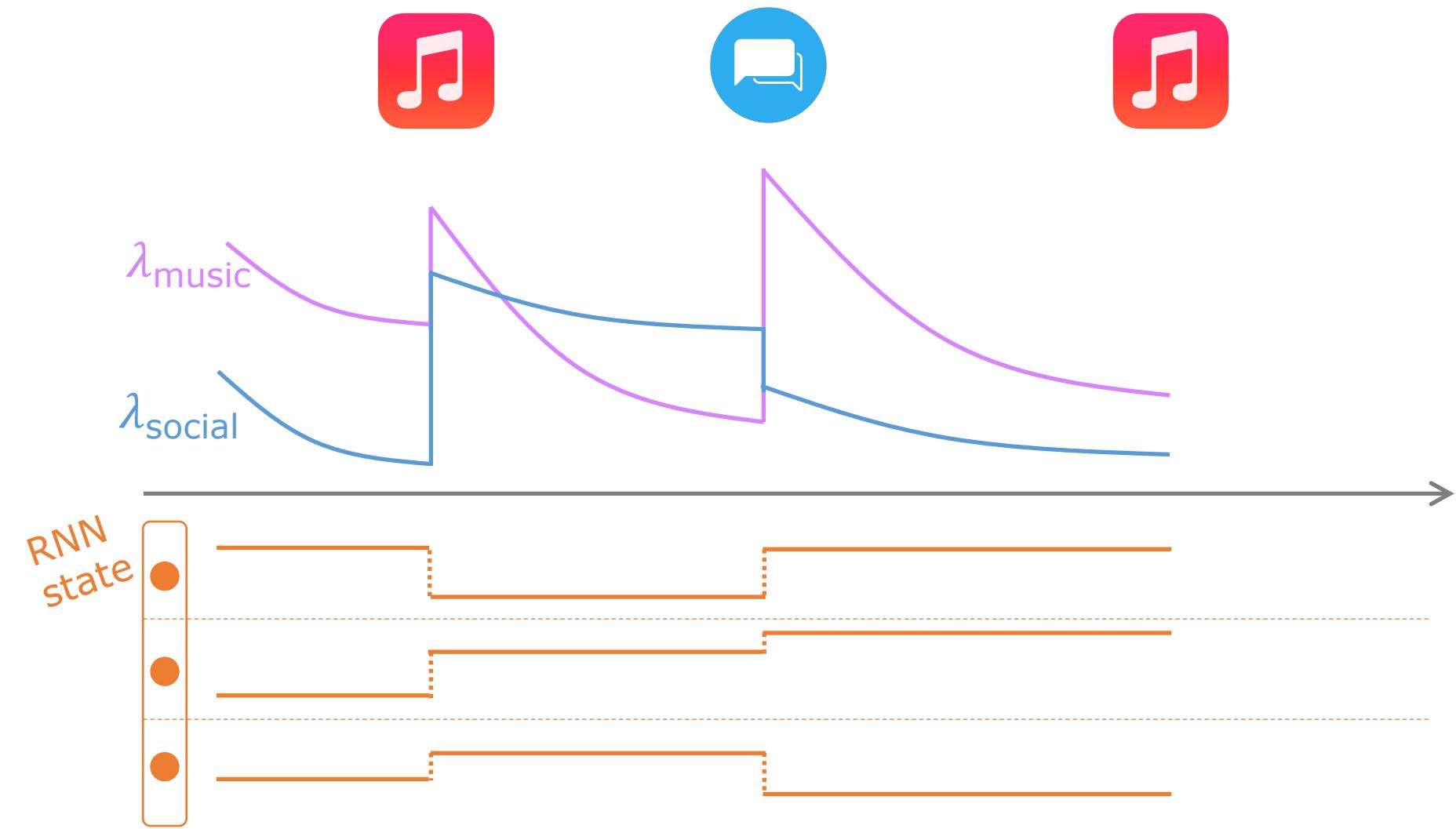
Use recurrent neural net (RNN)



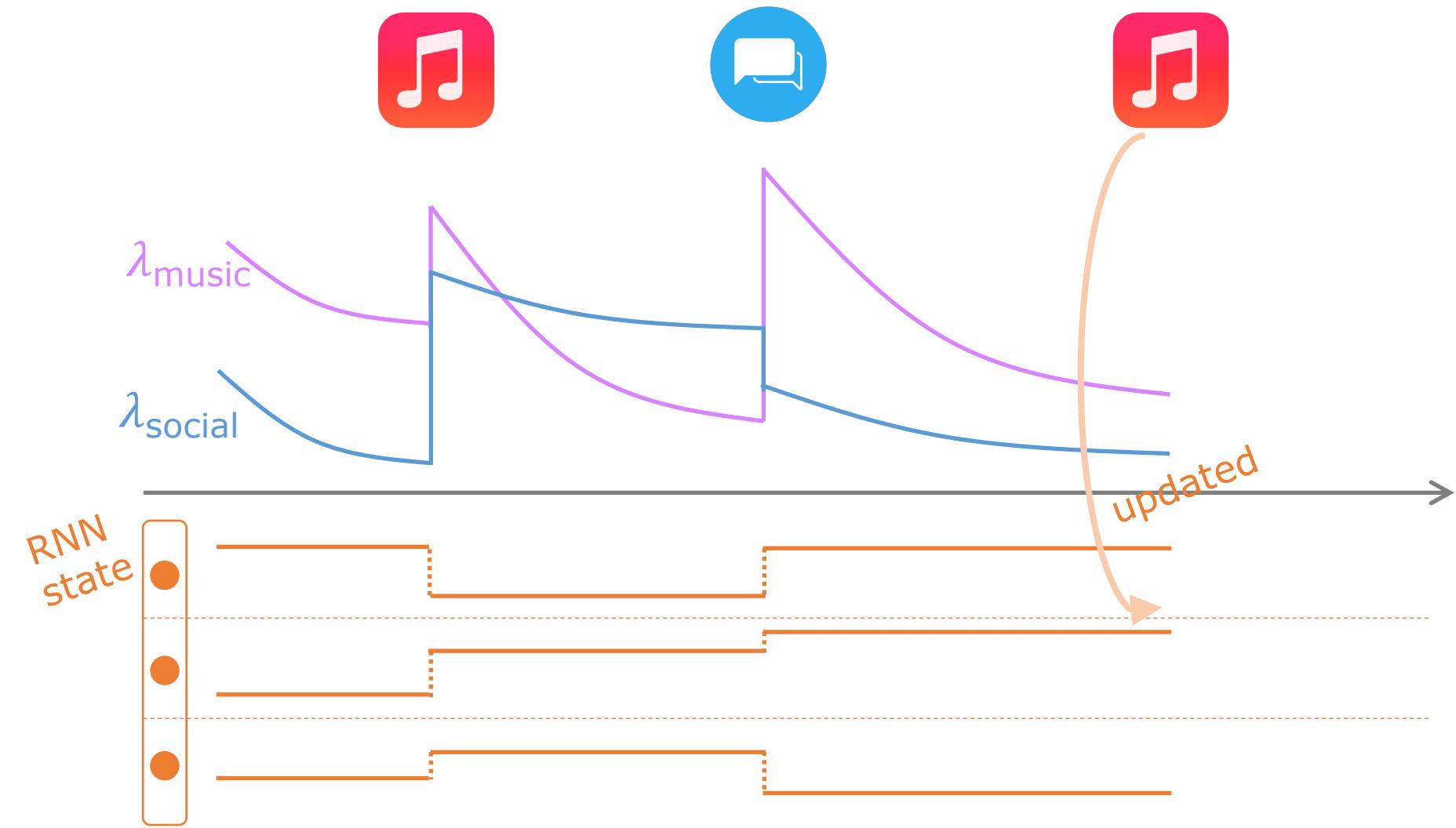
Use recurrent neural net (RNN)



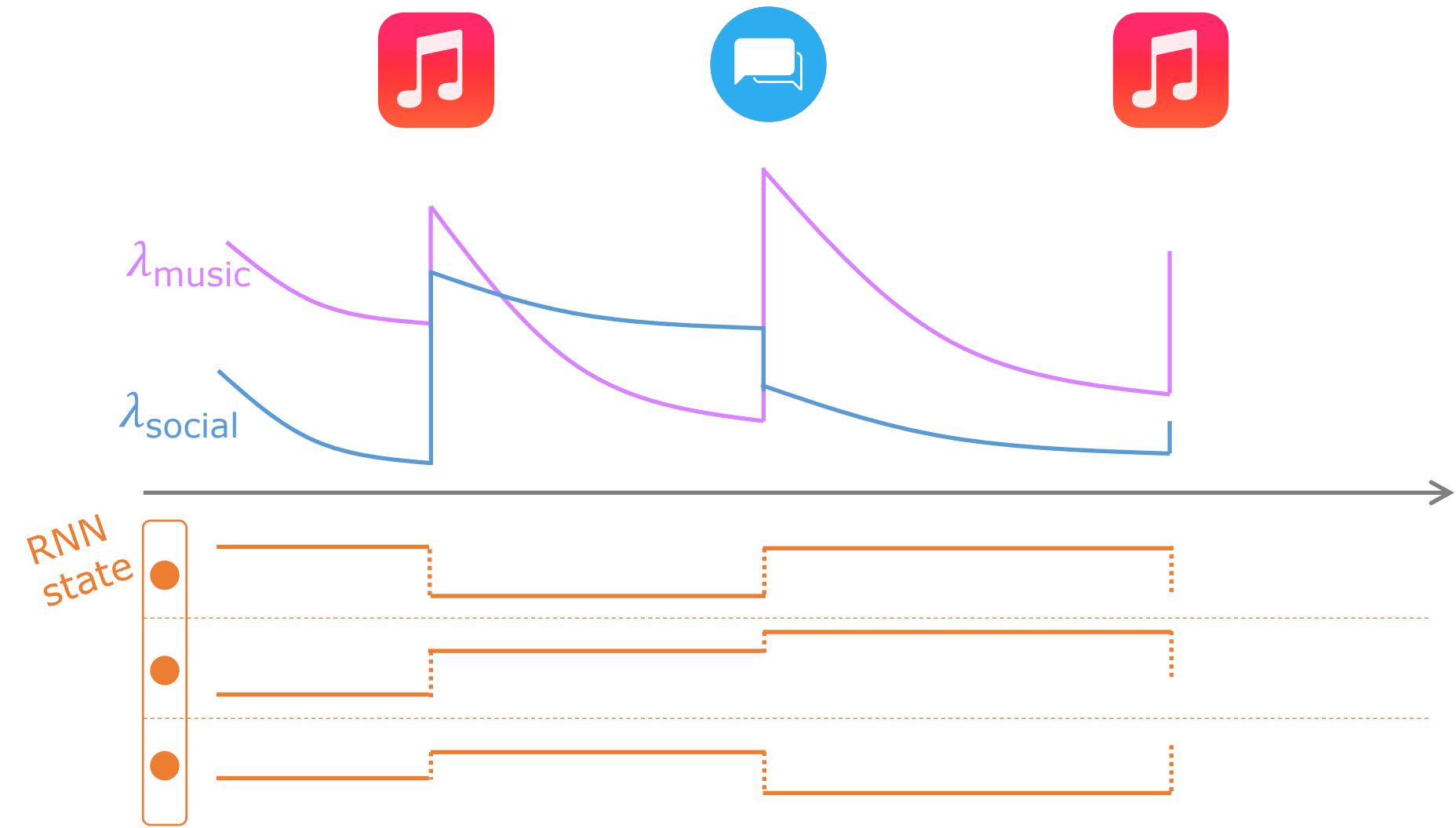
Use recurrent neural net (RNN)



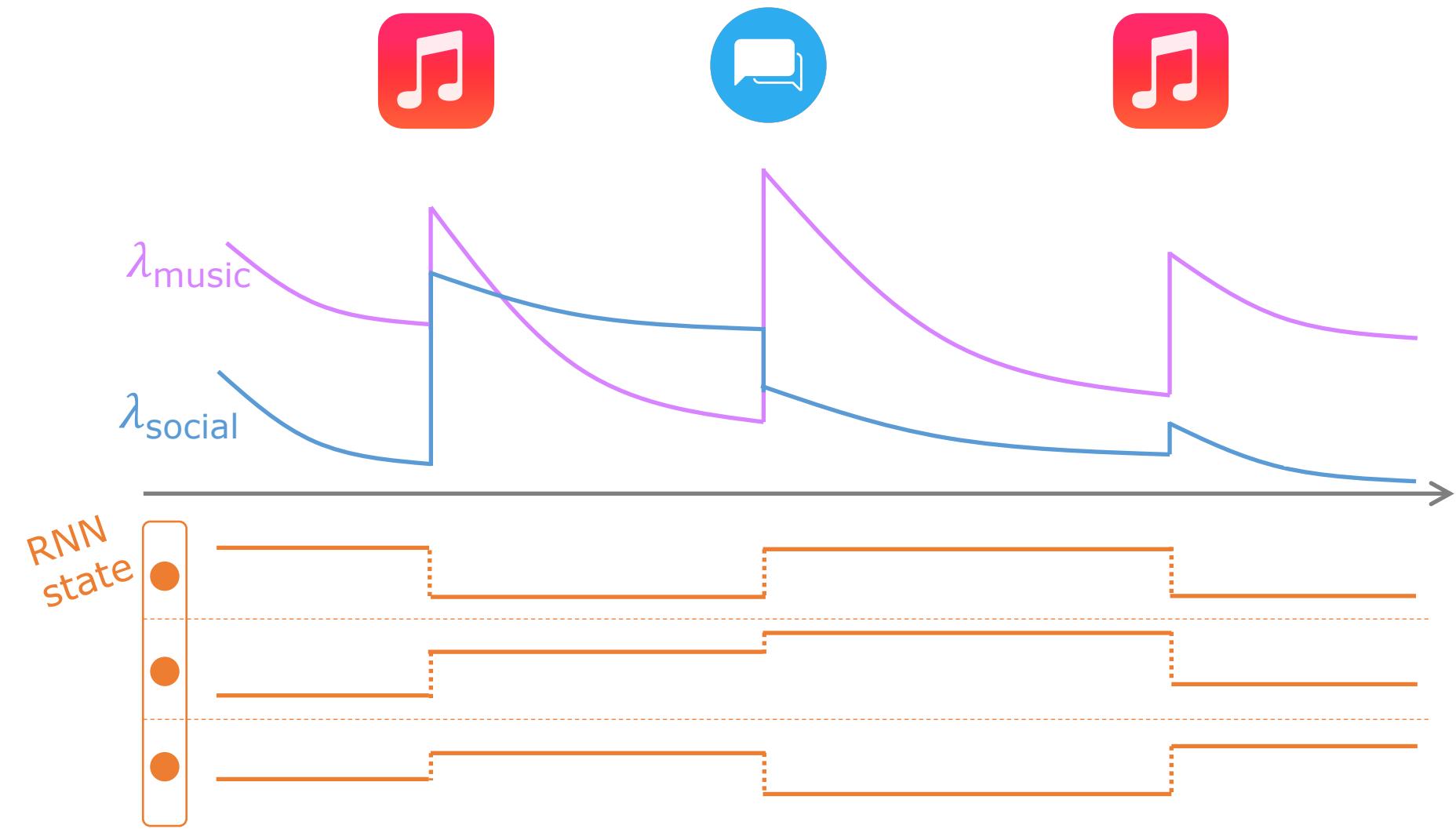
Use recurrent neural net (RNN)



Use recurrent neural net (RNN)

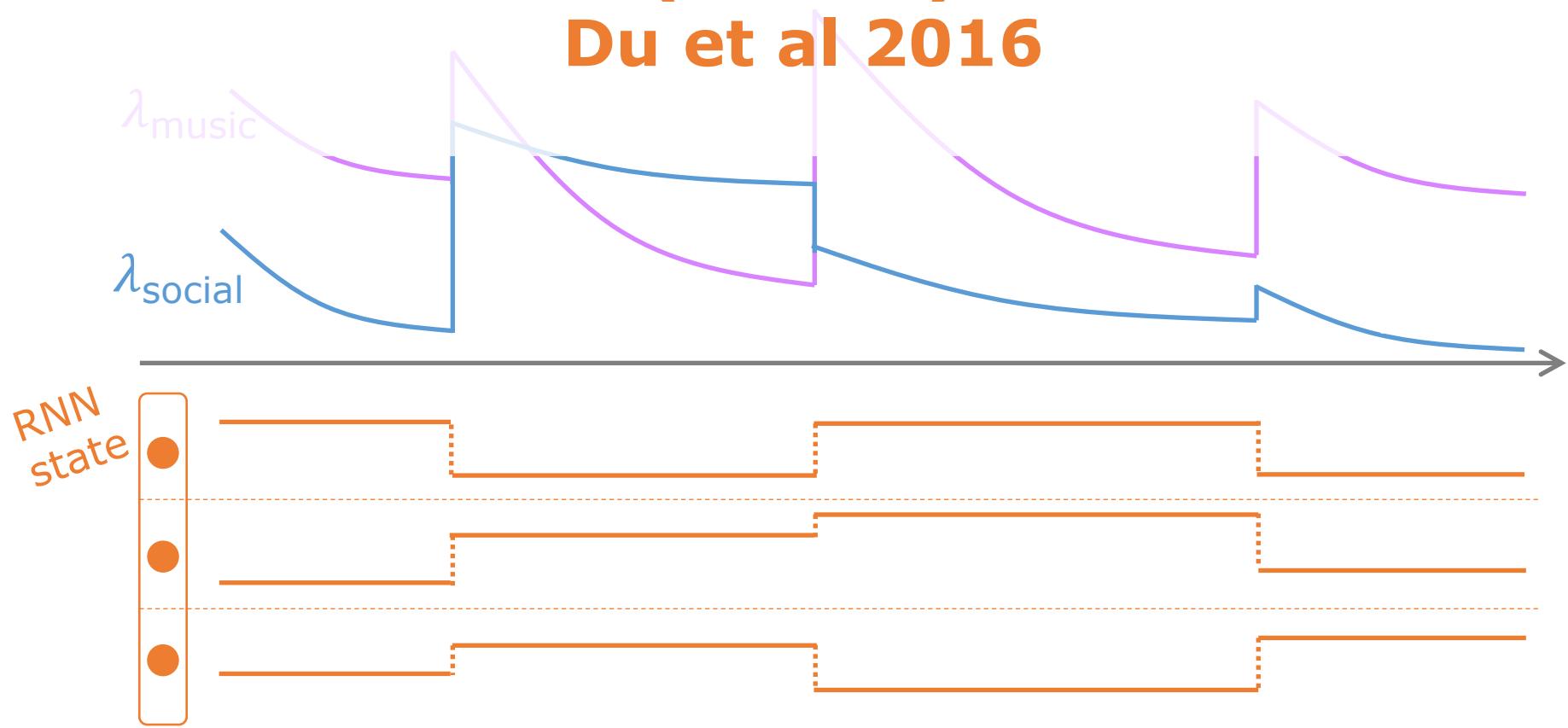


Use recurrent neural net (RNN)



Use recurrent neural net (RNN)

Recurrent Marked Temporal Point Process (RMTPP) Du et al 2016



Let LSTM state *evolve* with time



Let LSTM state evolve with time



Let LSTM state evolve with time



Let LSTM state evolve with time



Let LSTM state evolve with time



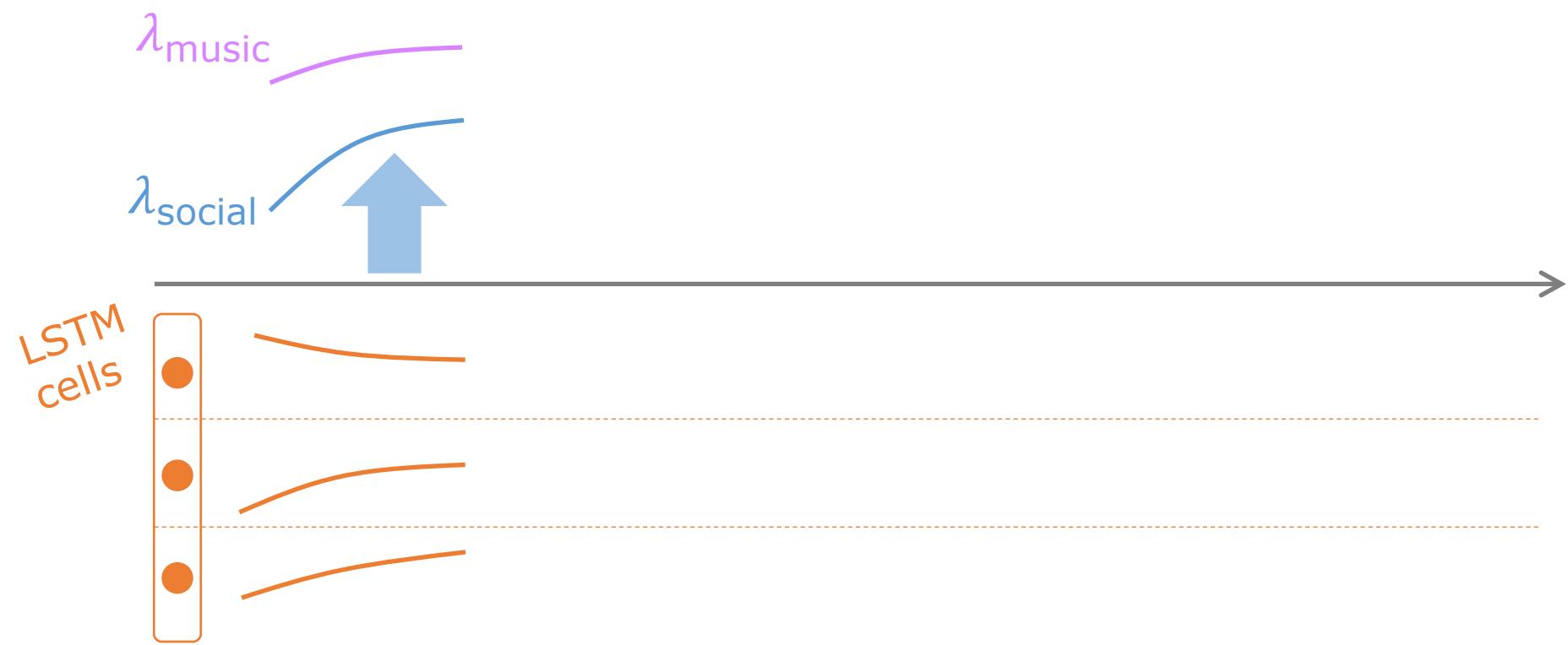
Let LSTM state evolve with time



Let LSTM state evolve with time



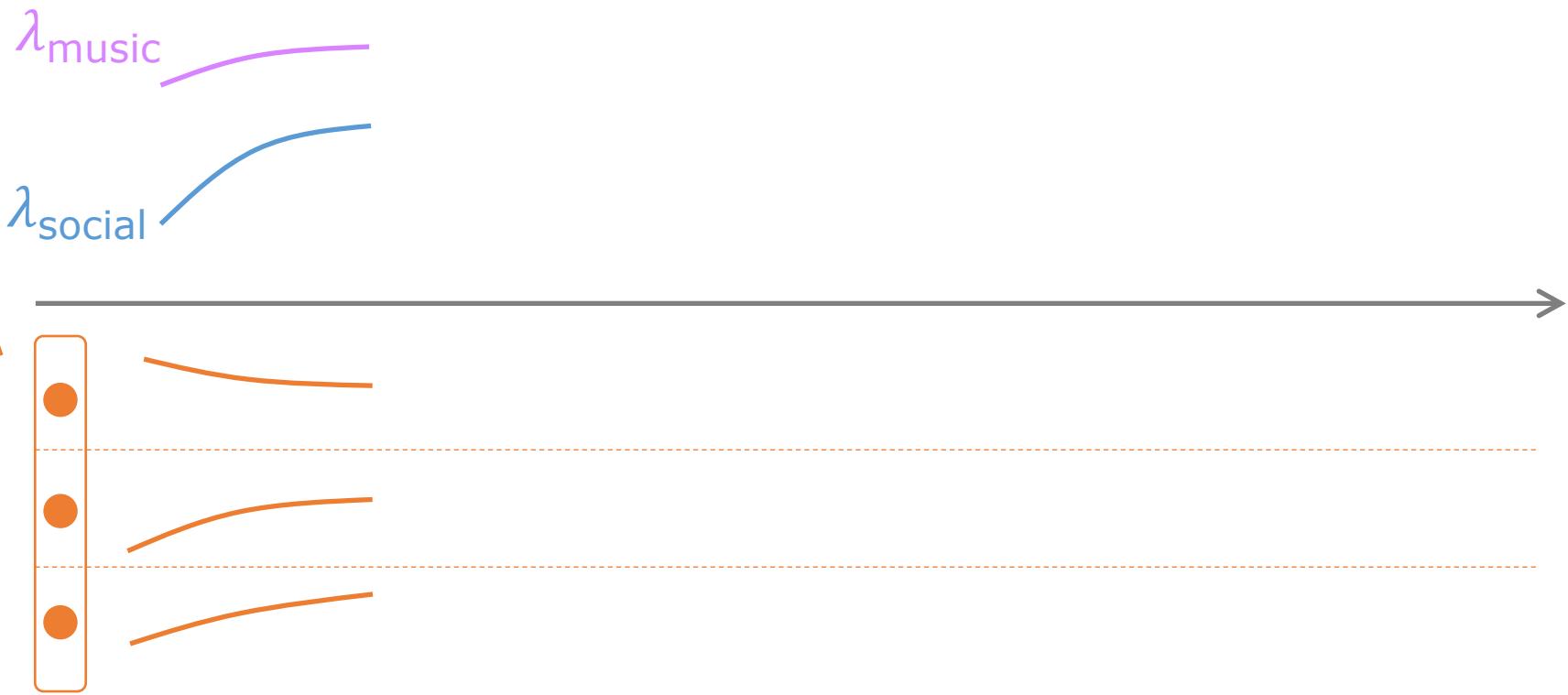
Let LSTM state evolve with time



Let LSTM state evolve with time



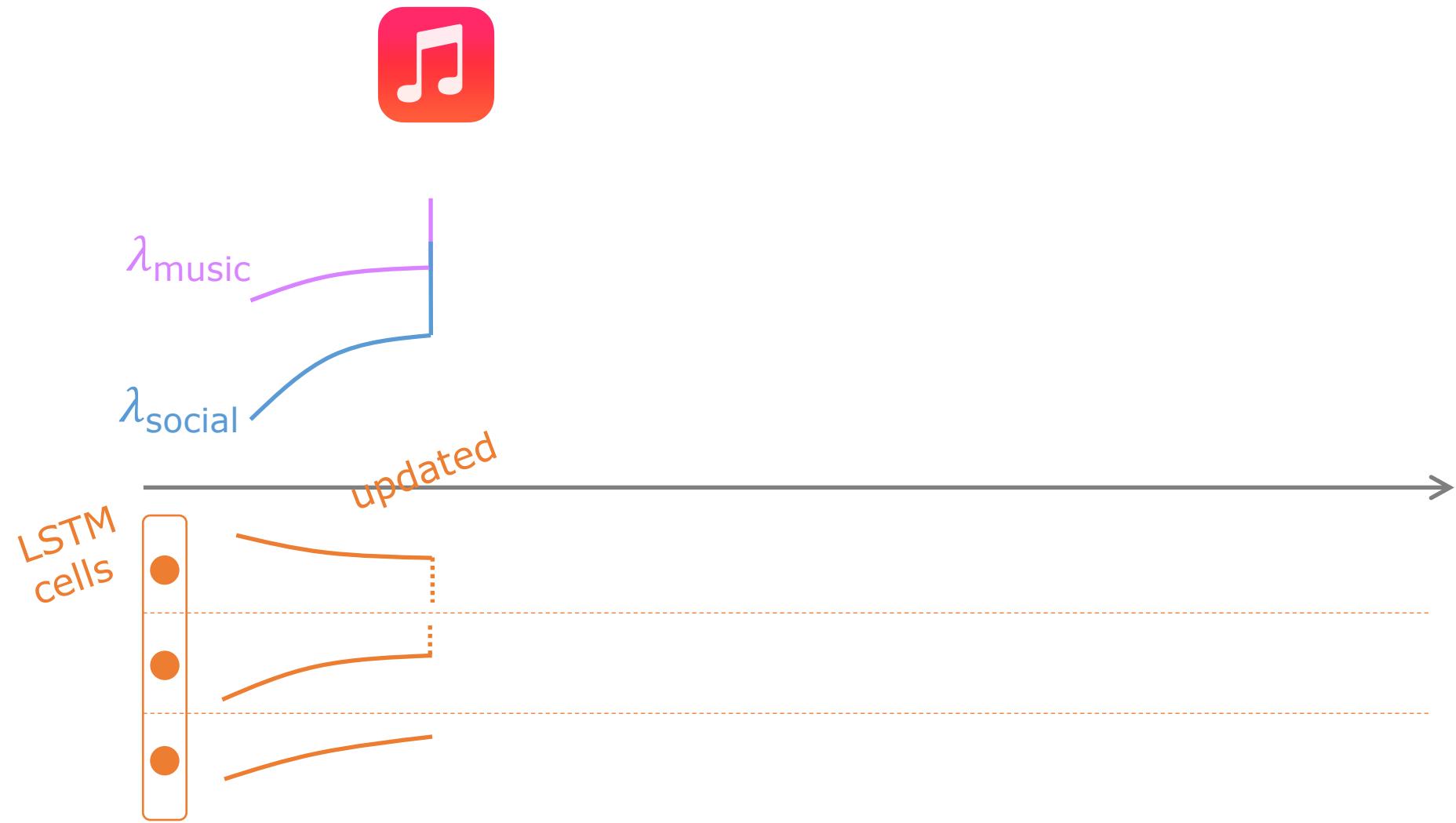
Let LSTM state evolve with time



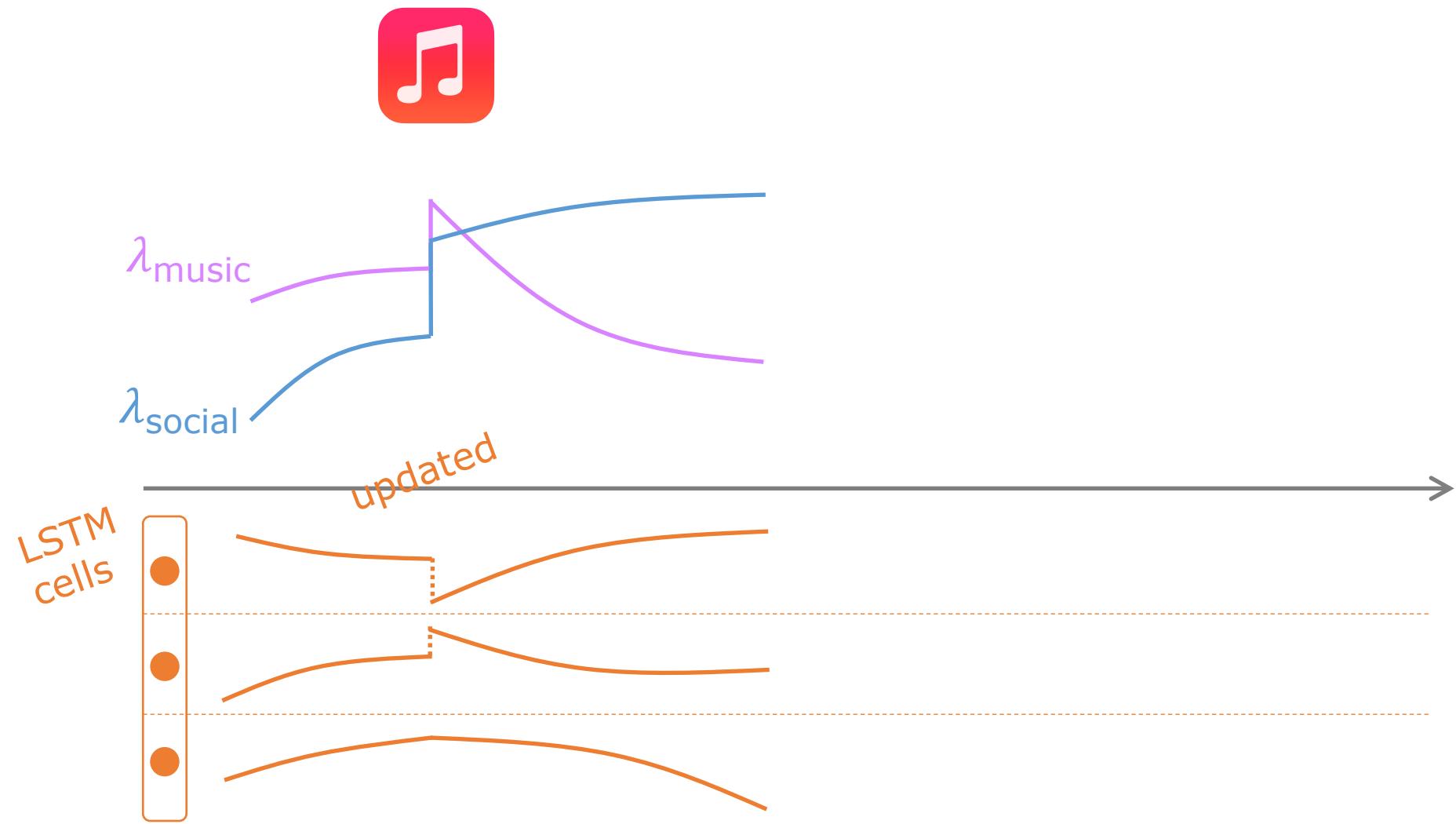
Let LSTM state evolve with time



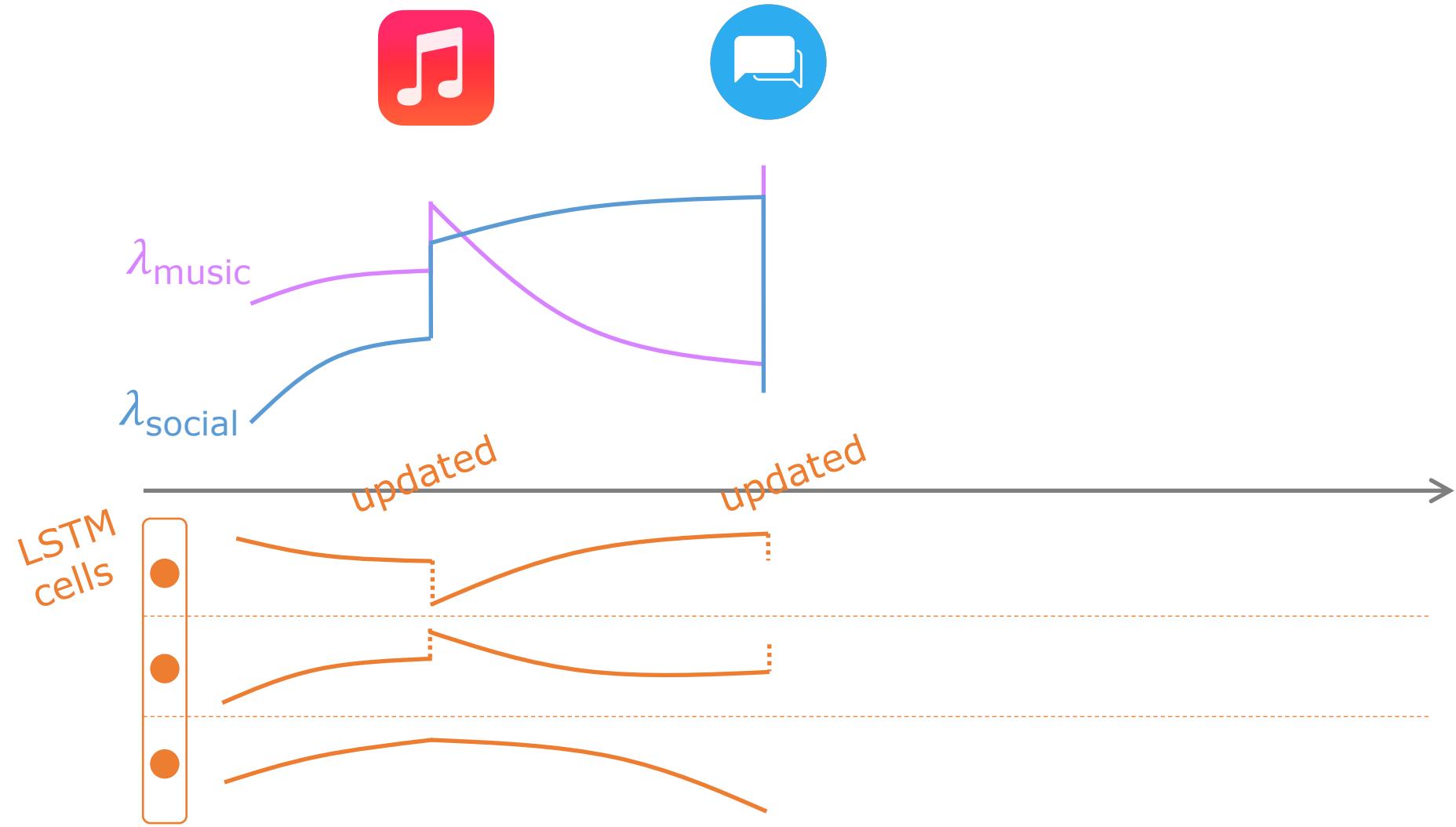
Let LSTM state evolve with time



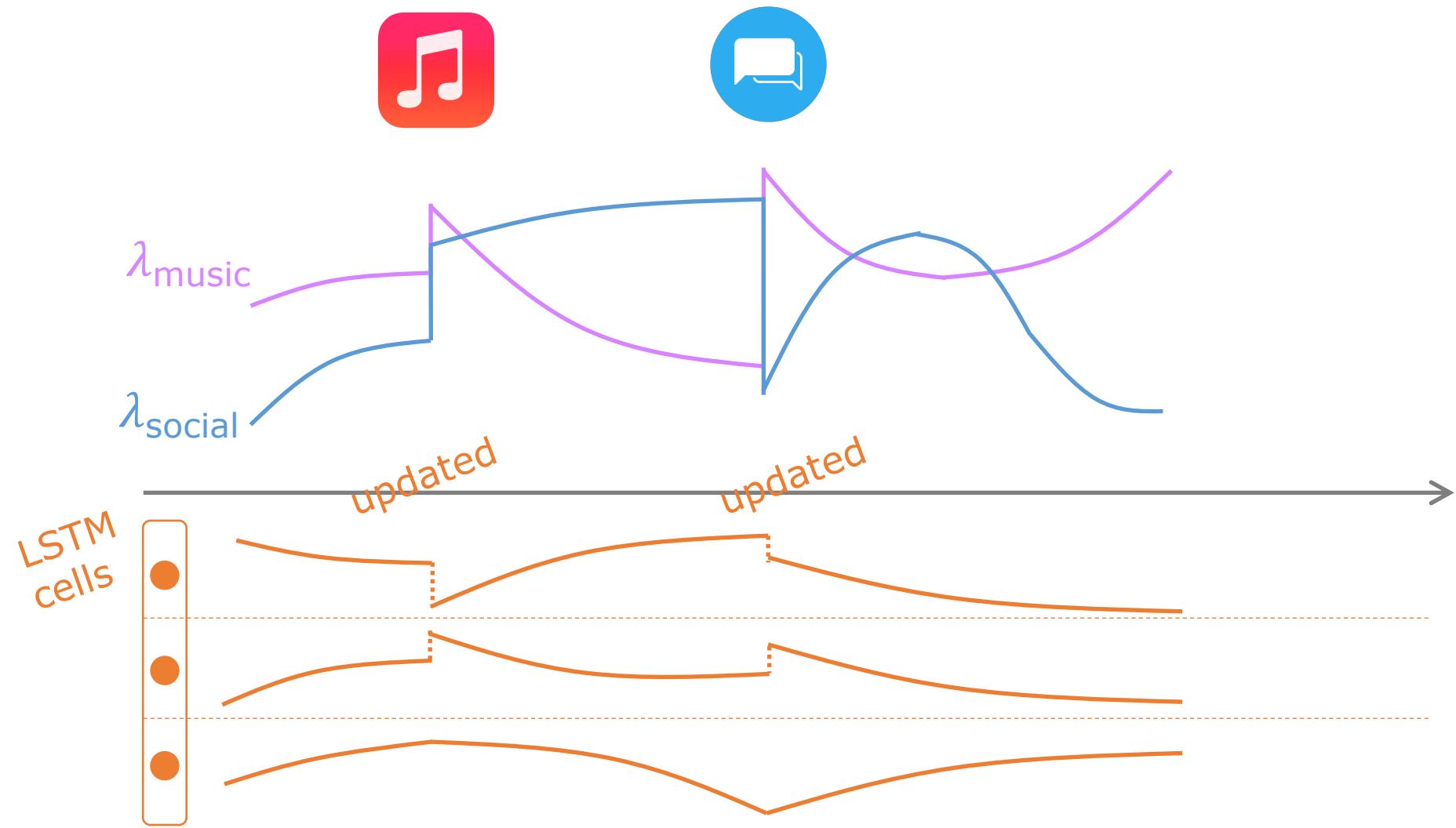
Let LSTM state evolve with time



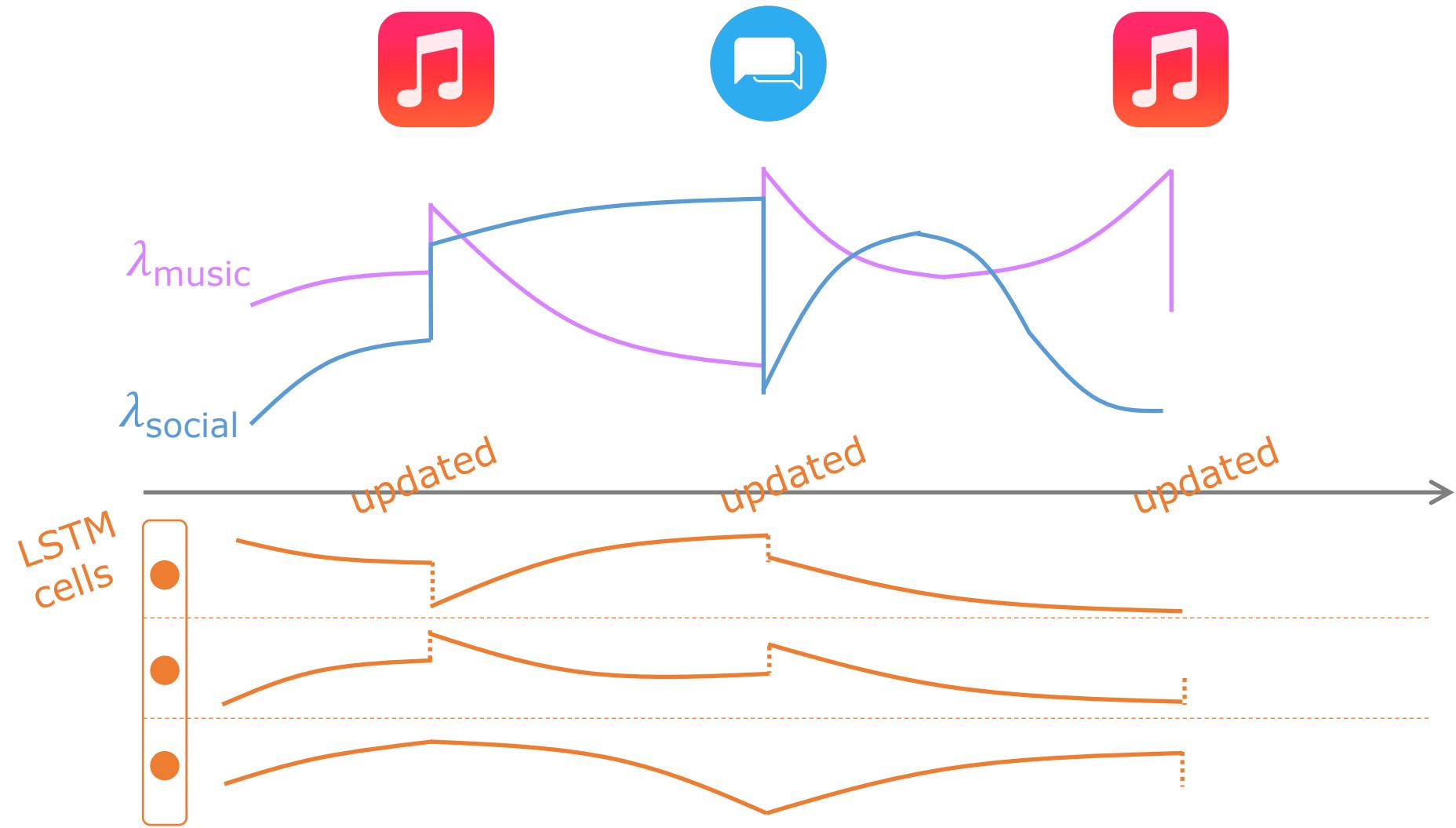
Let LSTM state evolve with time



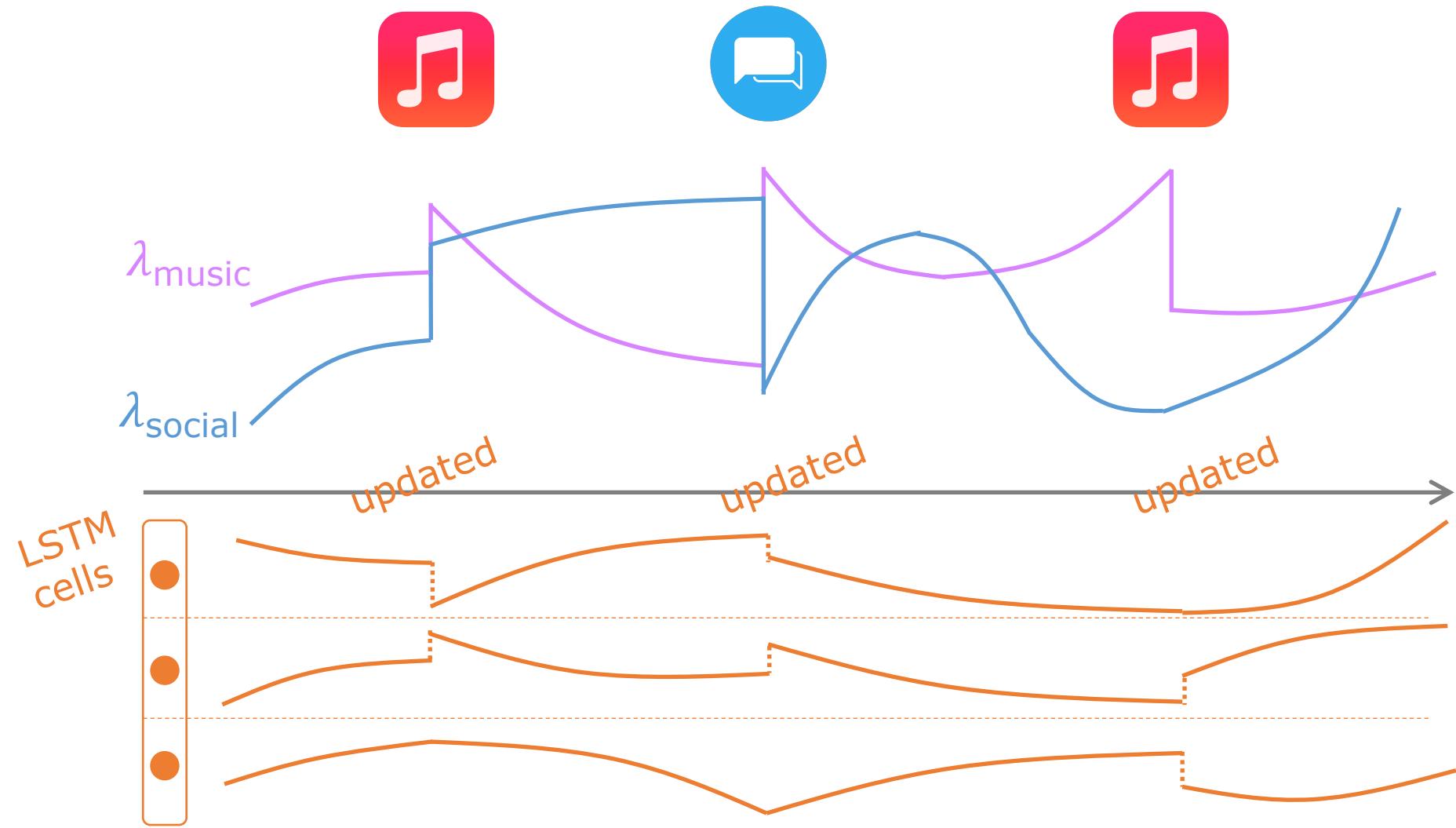
Let LSTM state evolve with time



Let LSTM state evolve with time

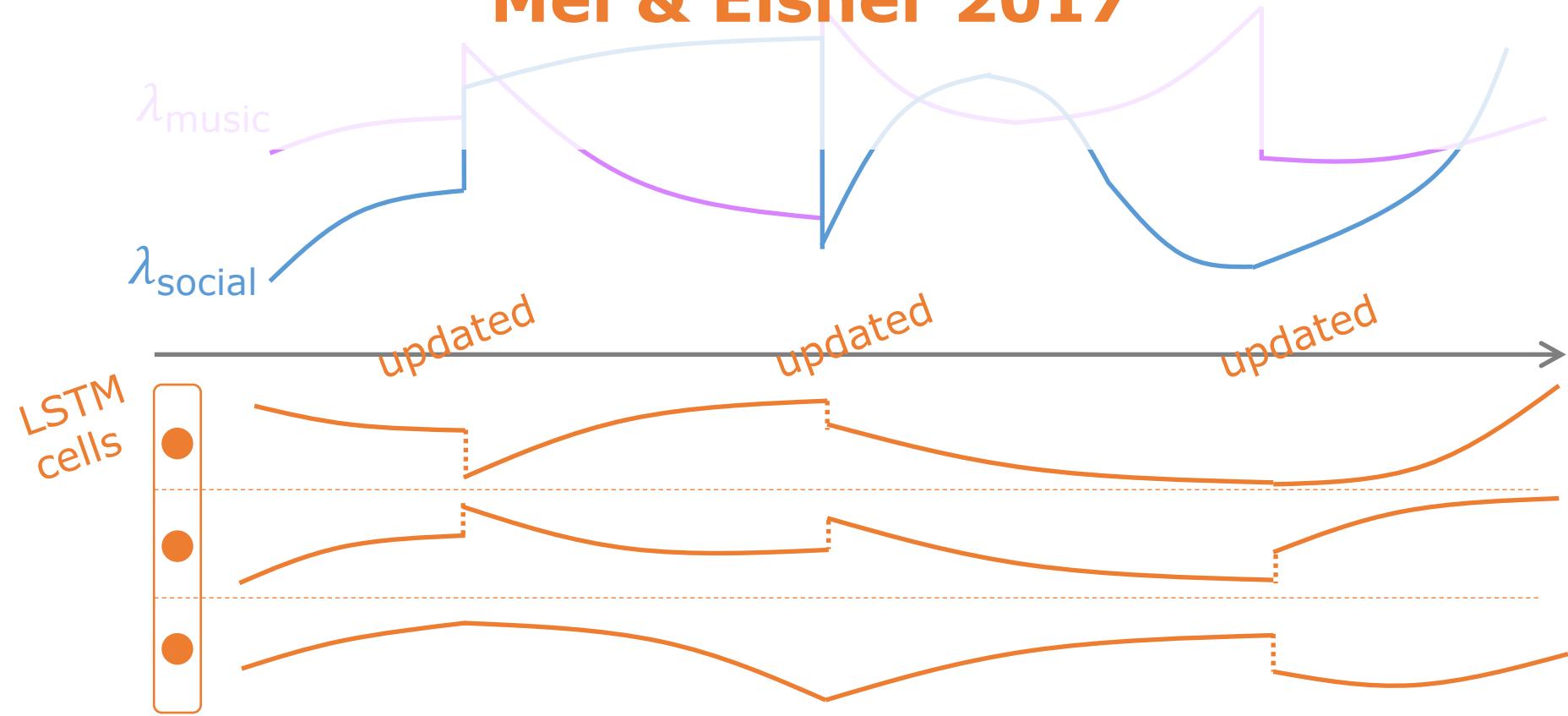


Let LSTM state evolve with time

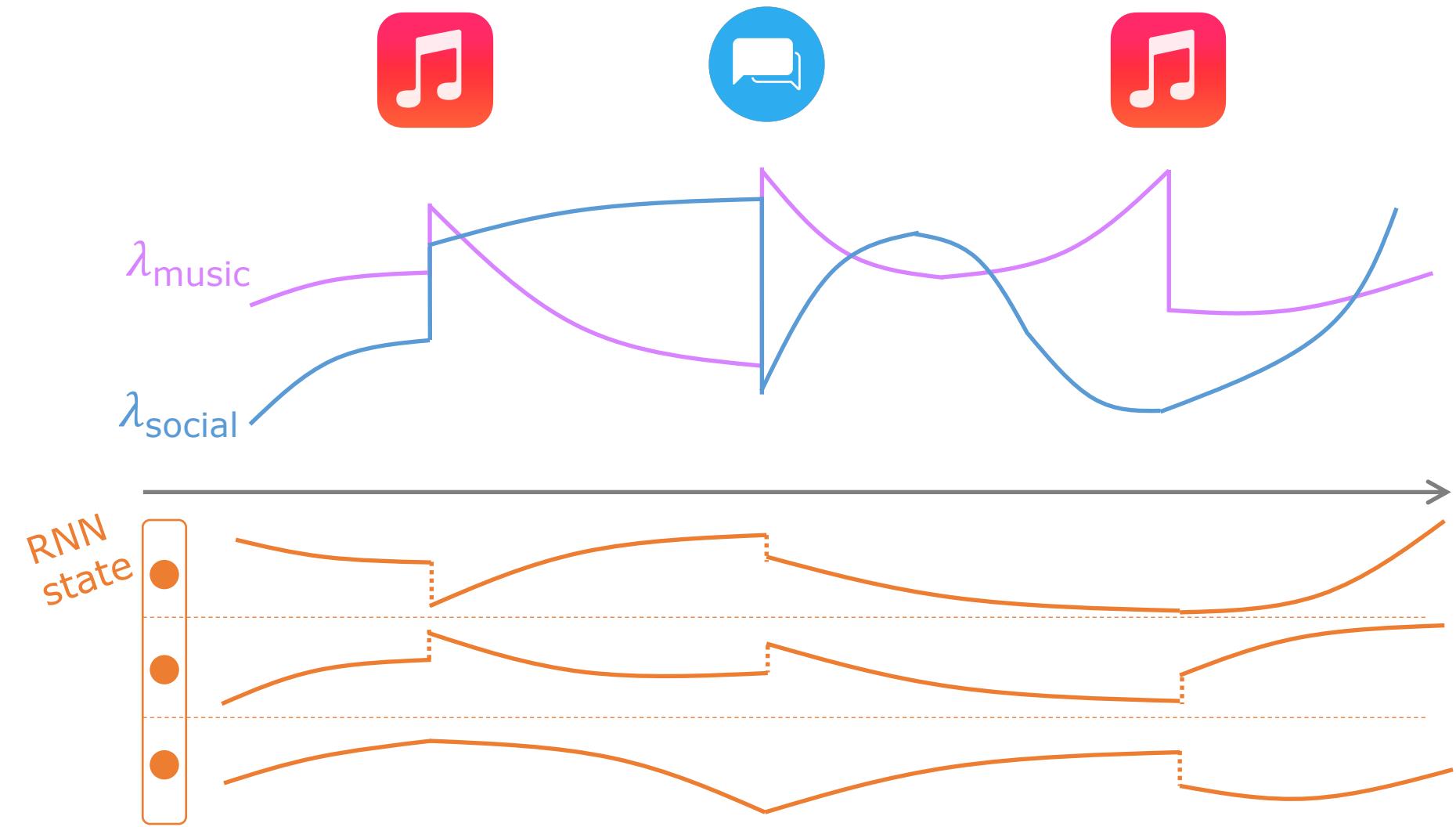


Let LSTM state evolve with time

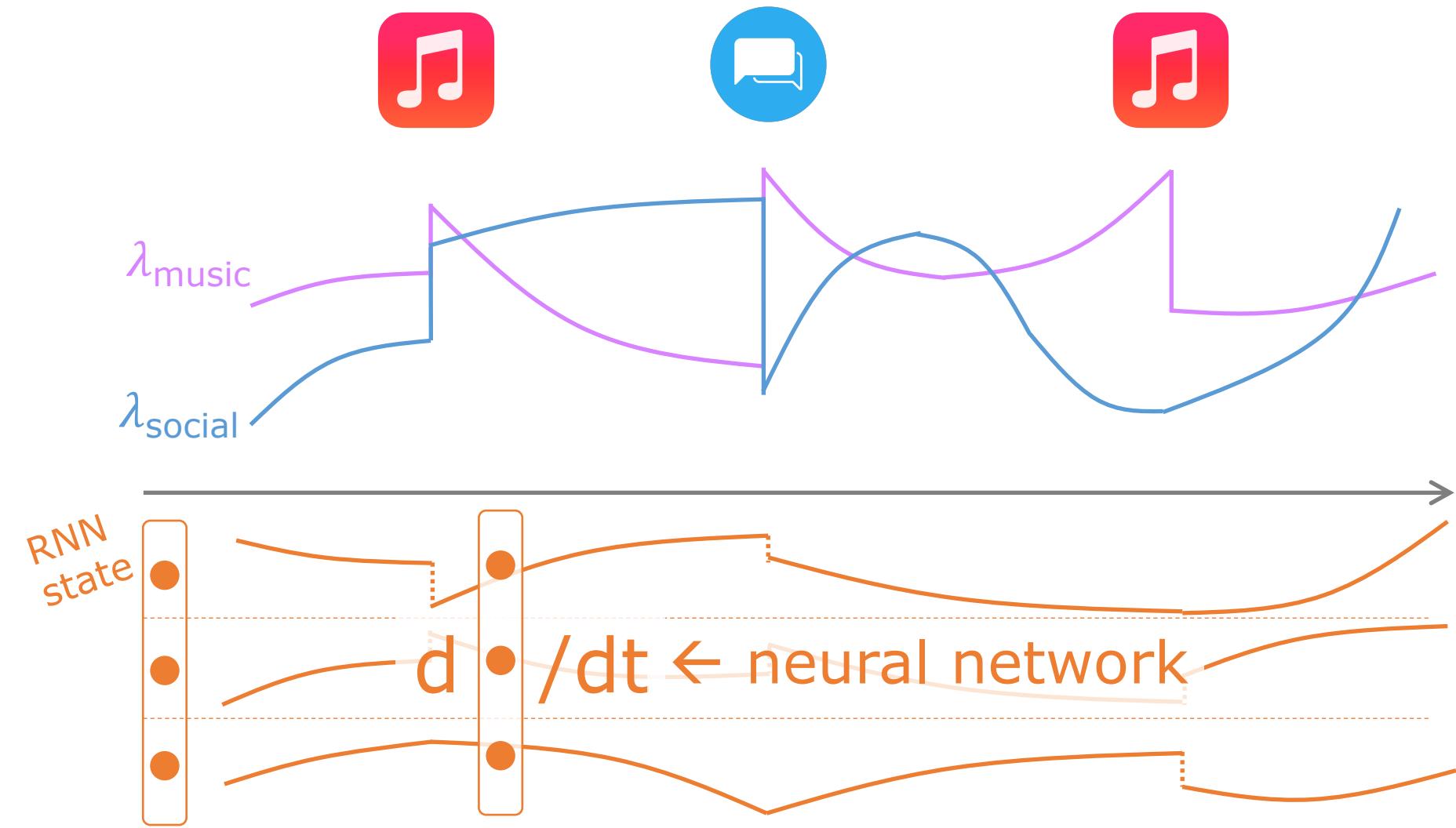
Neural Hawkes Process (NHP) Mei & Eisner 2017



Modulate the *derivative* of RNN state

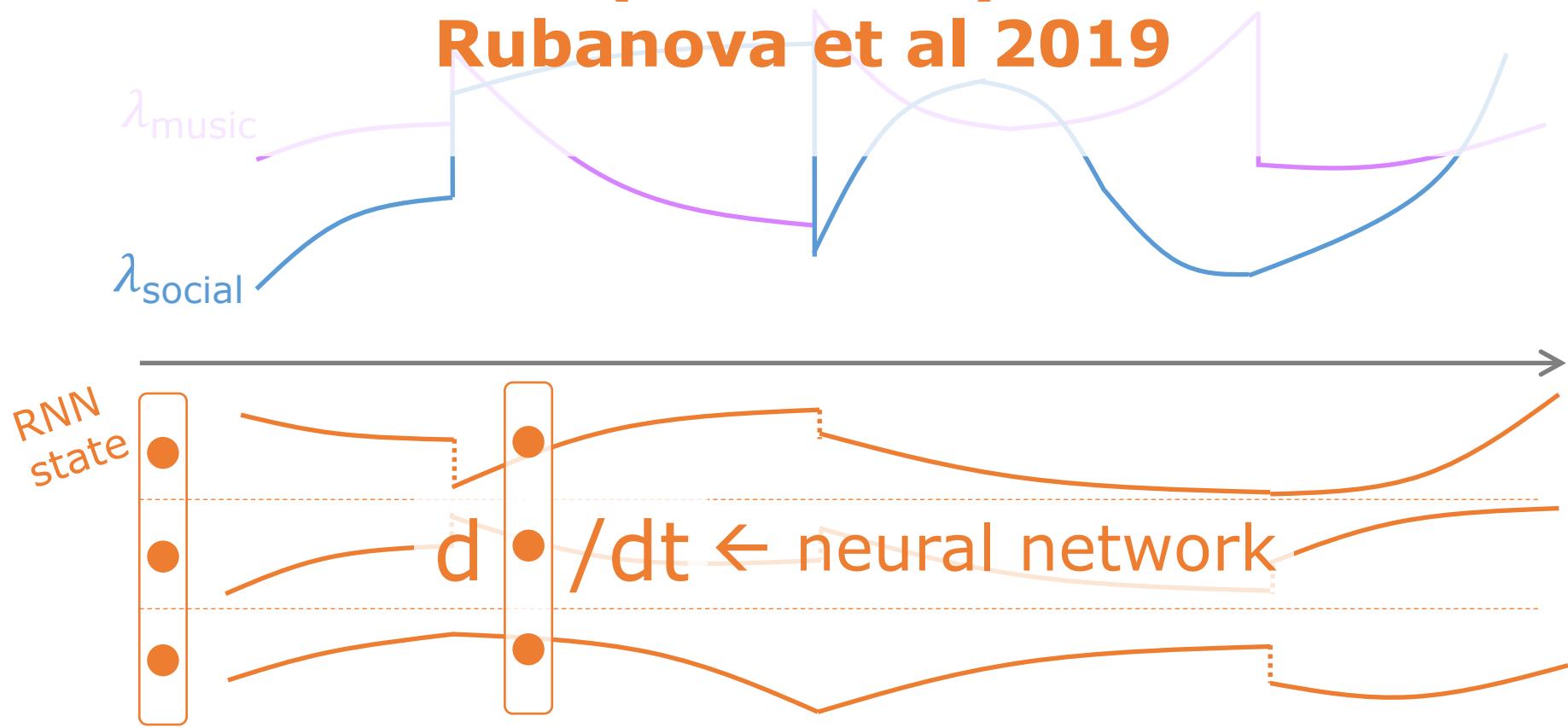


Modulate the *derivative* of RNN state



Modulate the *derivative* of RNN state

RNN with dynamics defined by ODE
(ODE-RNN)
Rubanova et al 2019



~~RNN/LSTM~~ Transformer



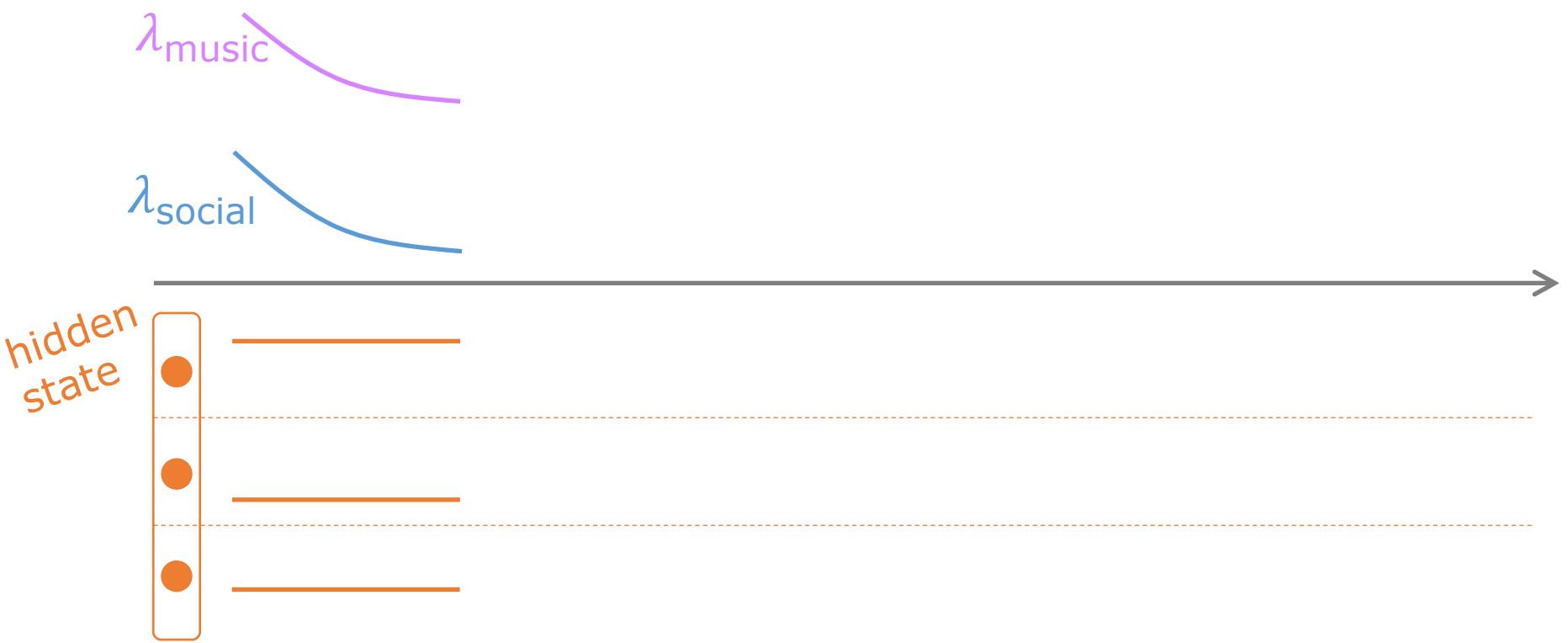
~~RNN/LSTM~~ Transformer



~~RNN/LSTM~~ Transformer



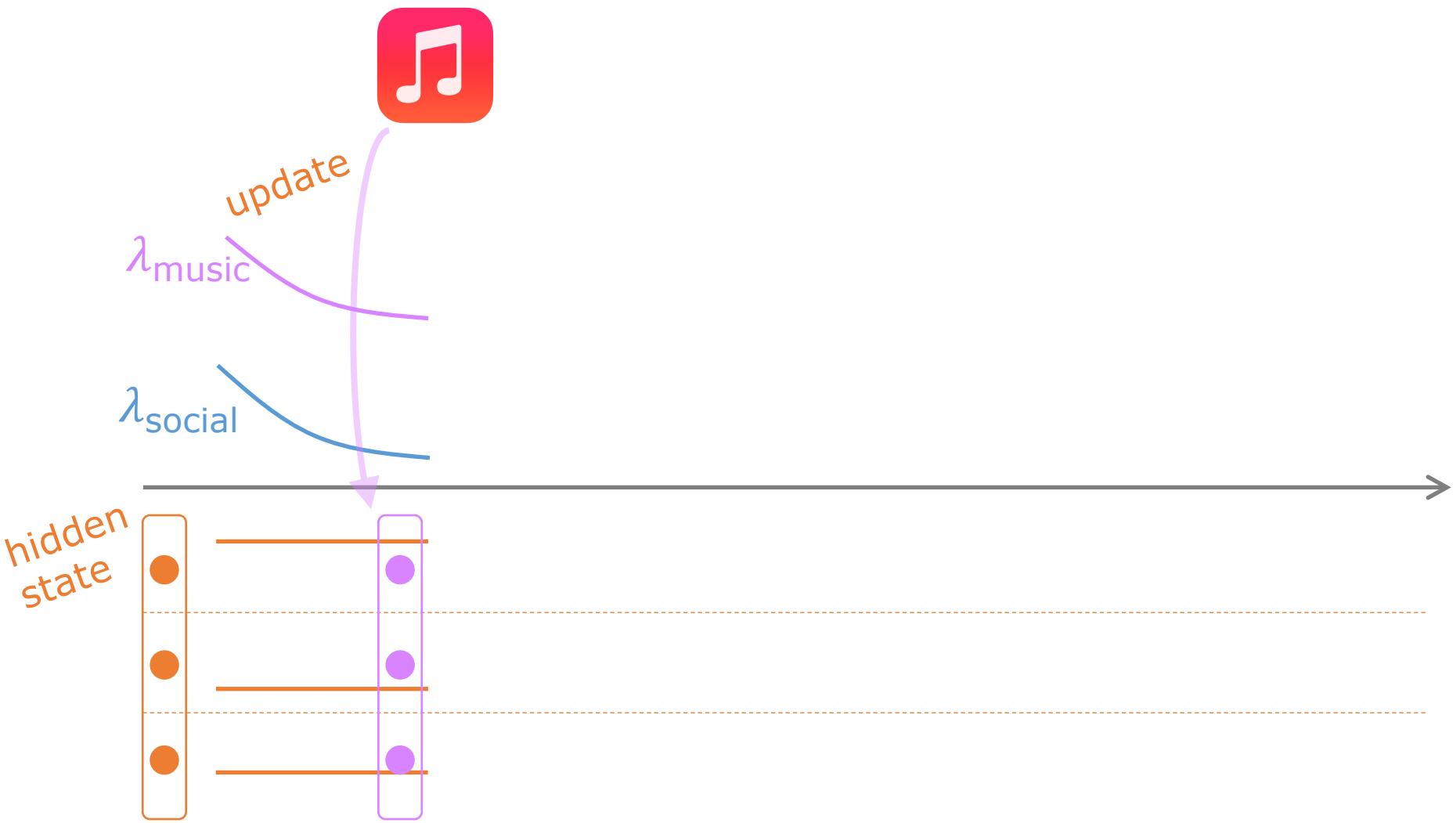
RNN/LSTM Transformer



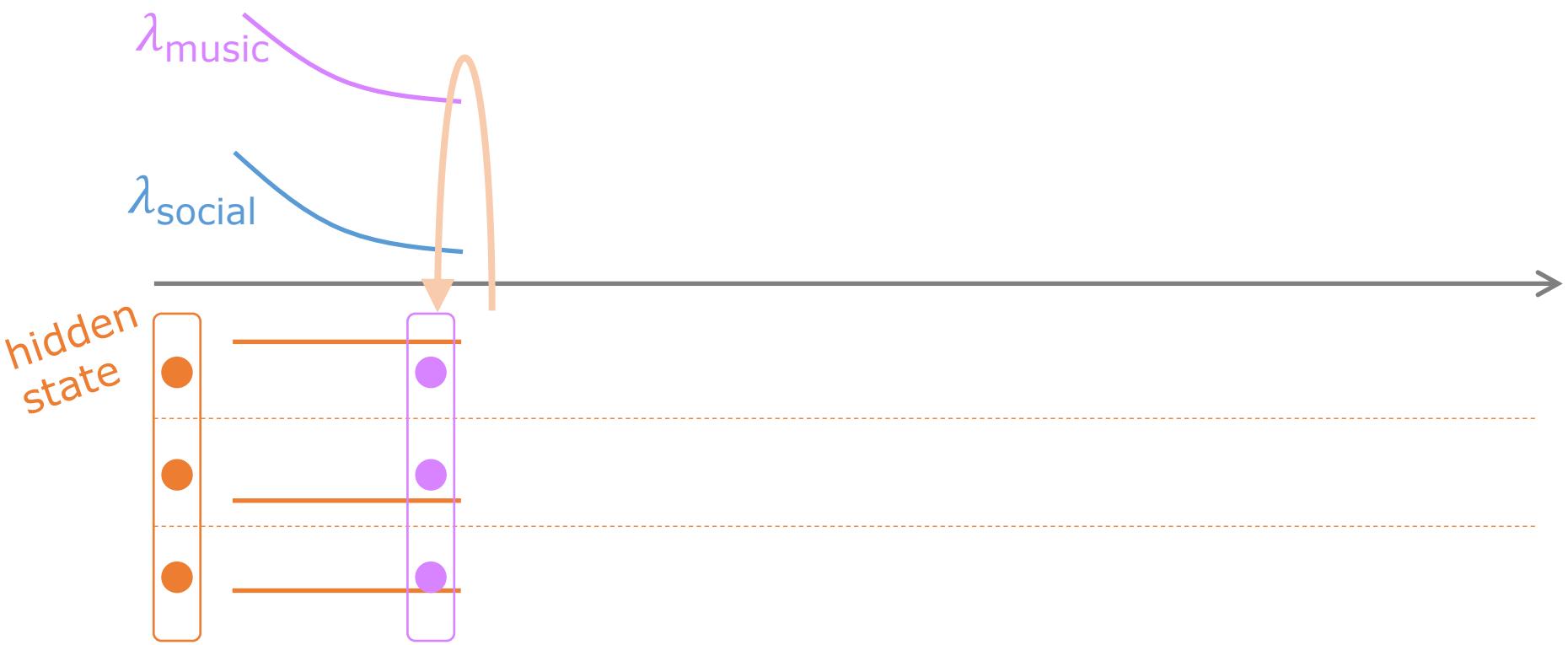
~~RNN/LSTM~~ Transformer



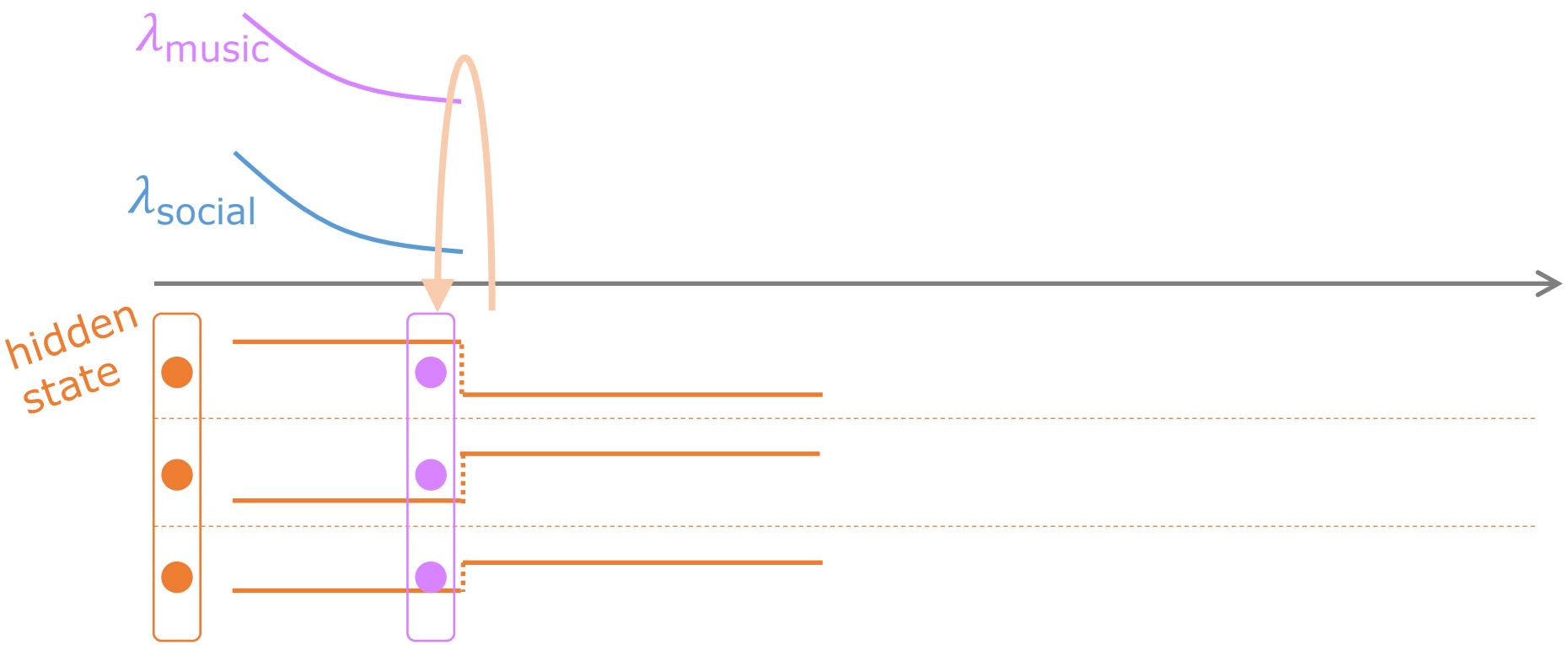
~~RNN/LSTM~~ Transformer



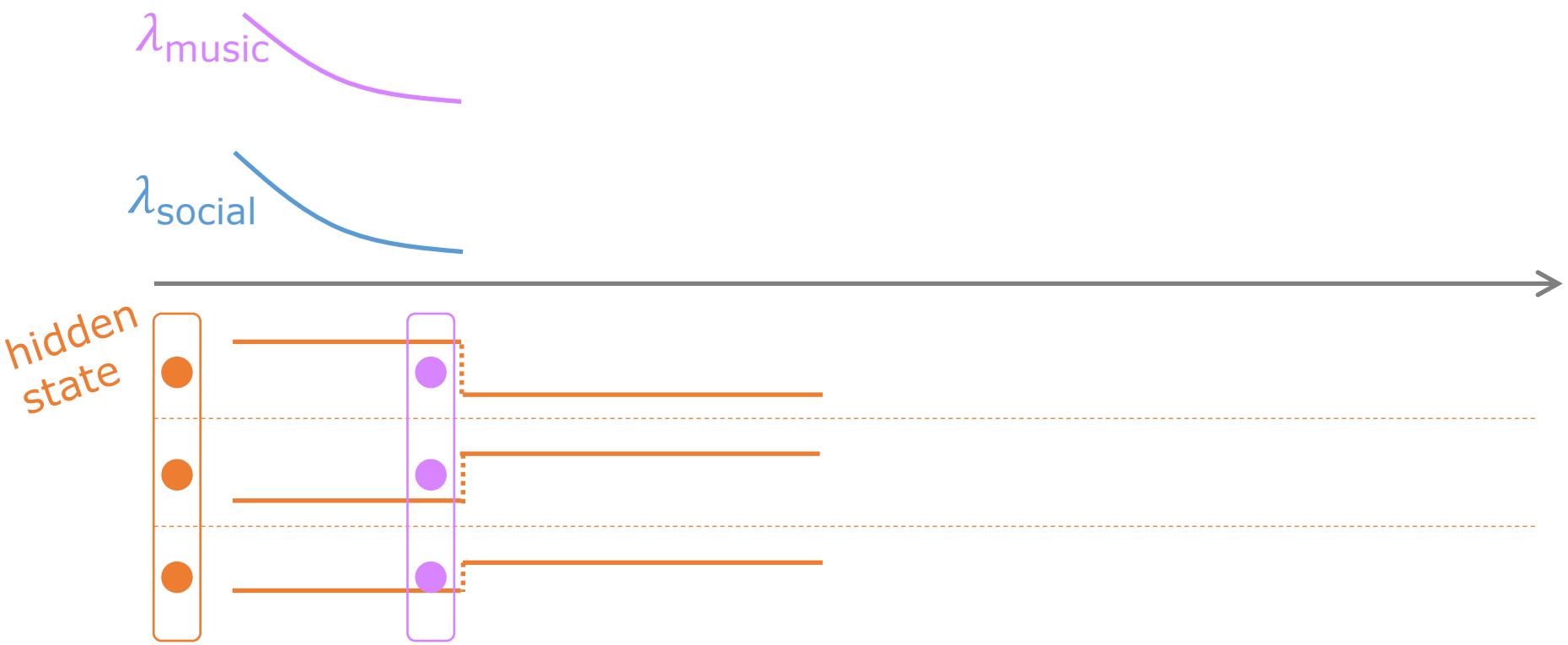
~~RNN/LSTM~~ Transformer



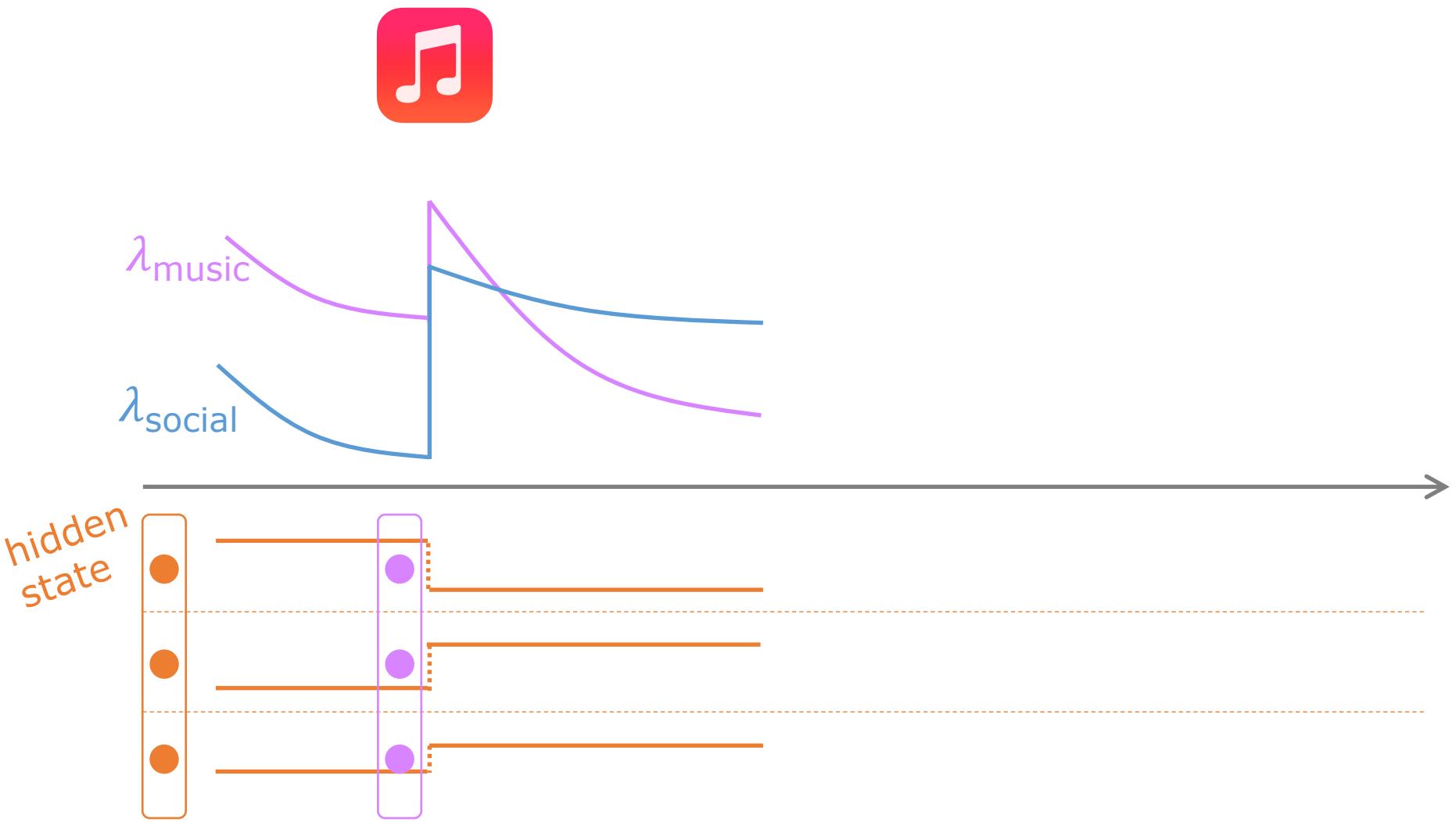
~~RNN/LSTM~~ Transformer



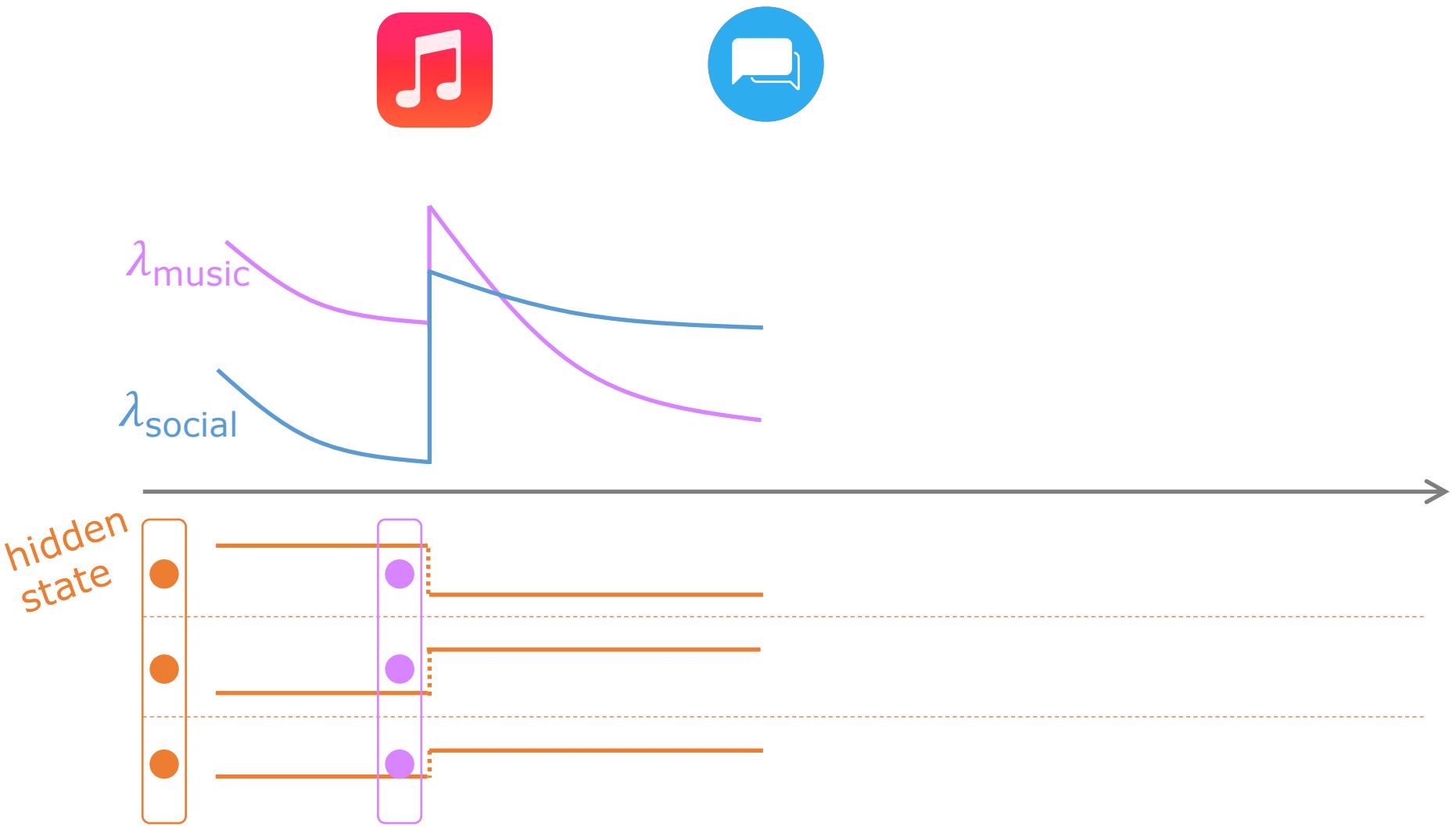
~~RNN/LSTM~~ Transformer



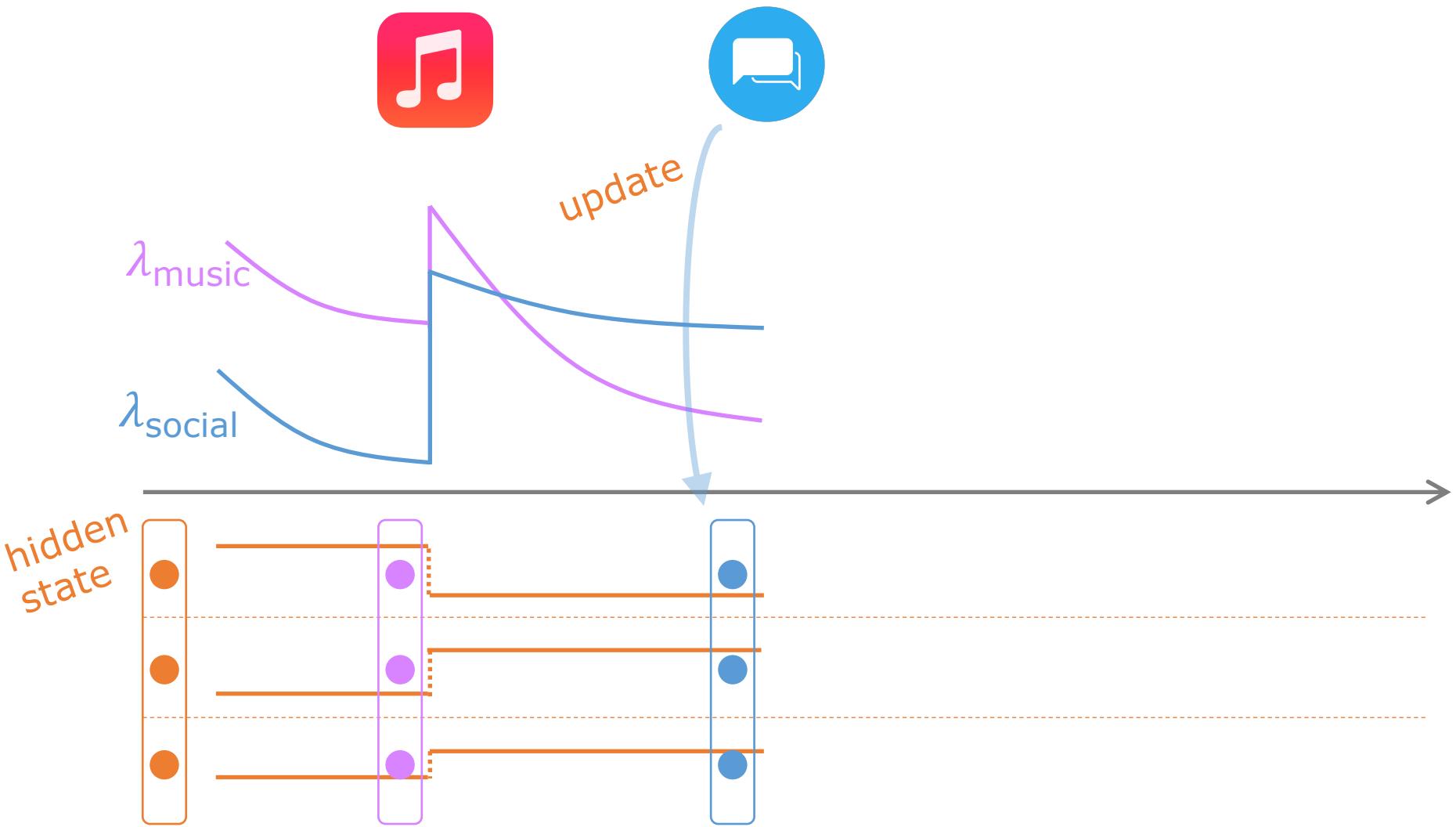
RNN/LSTM Transformer



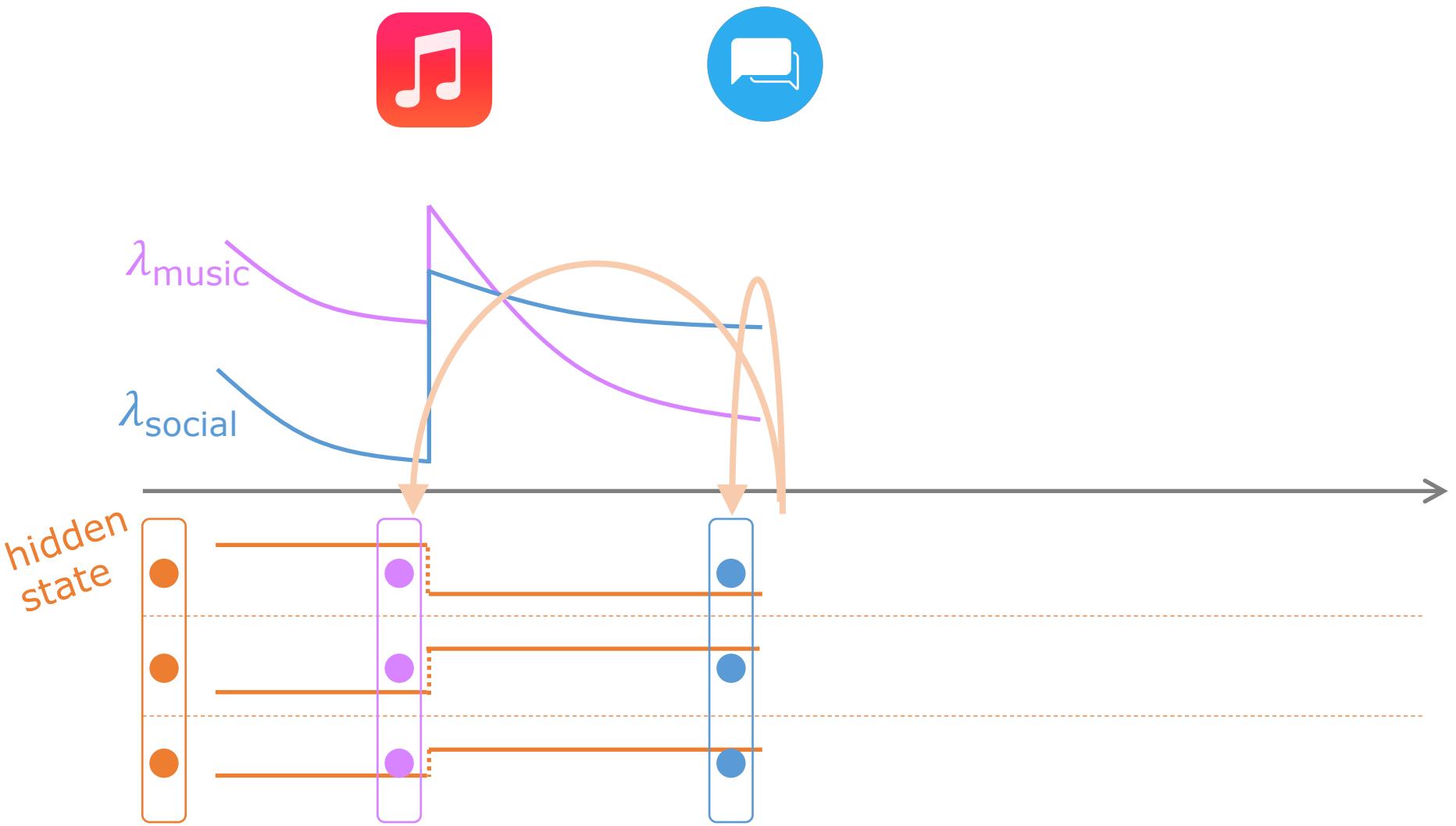
~~RNN/LSTM~~ Transformer



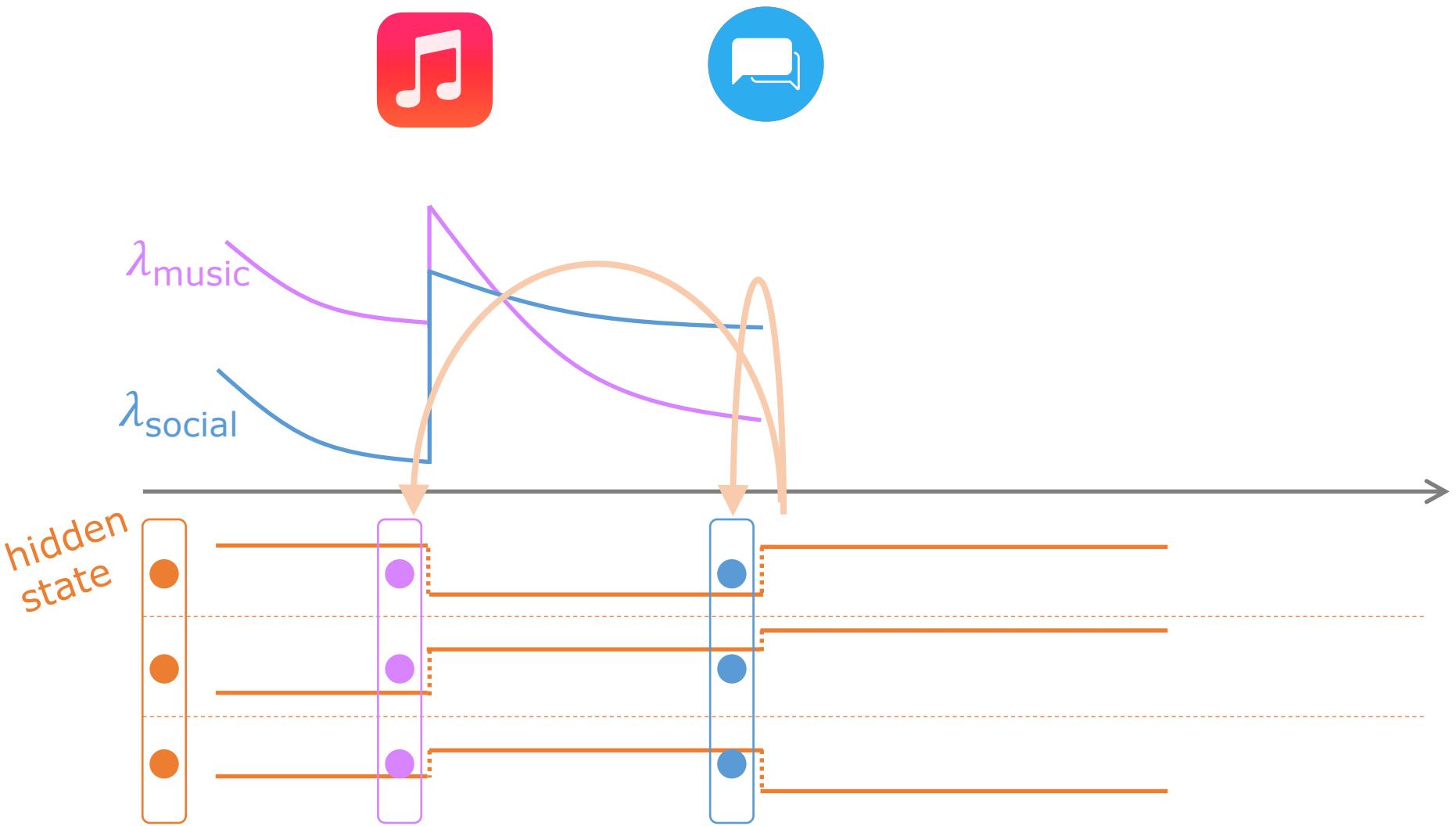
~~RNN/LSTM~~ Transformer



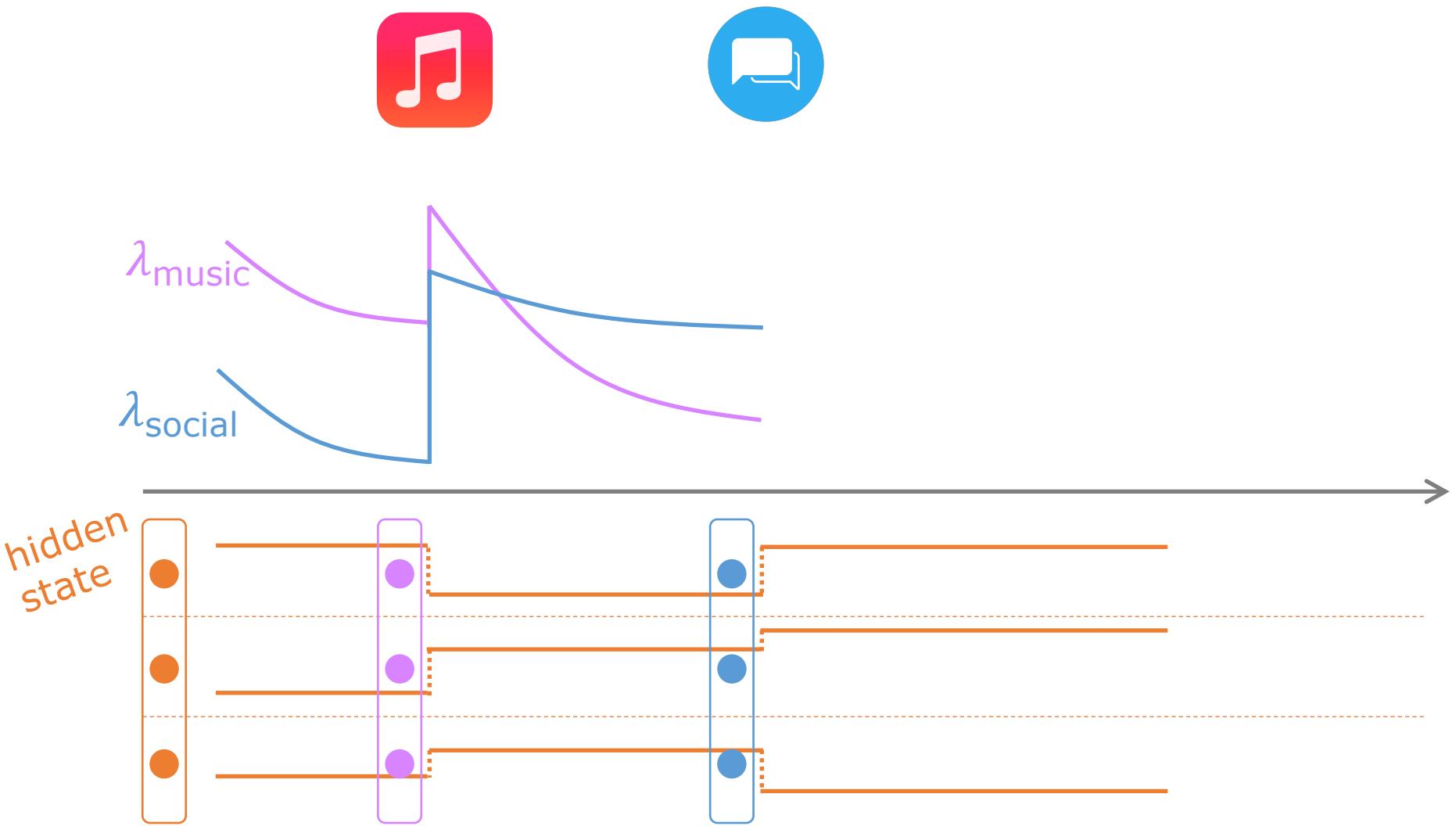
~~RNN/LSTM~~ Transformer



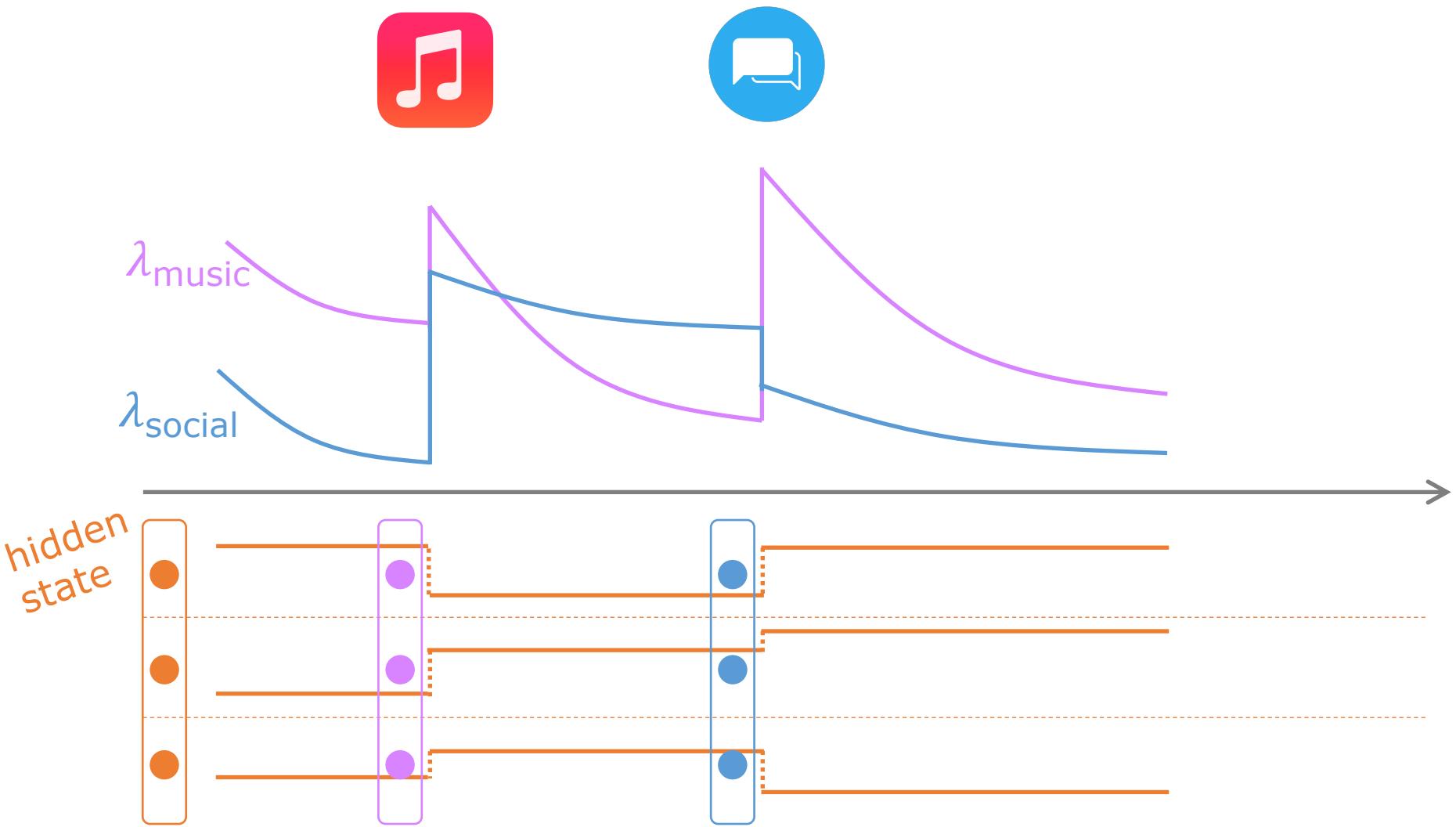
RNN/LSTM Transformer



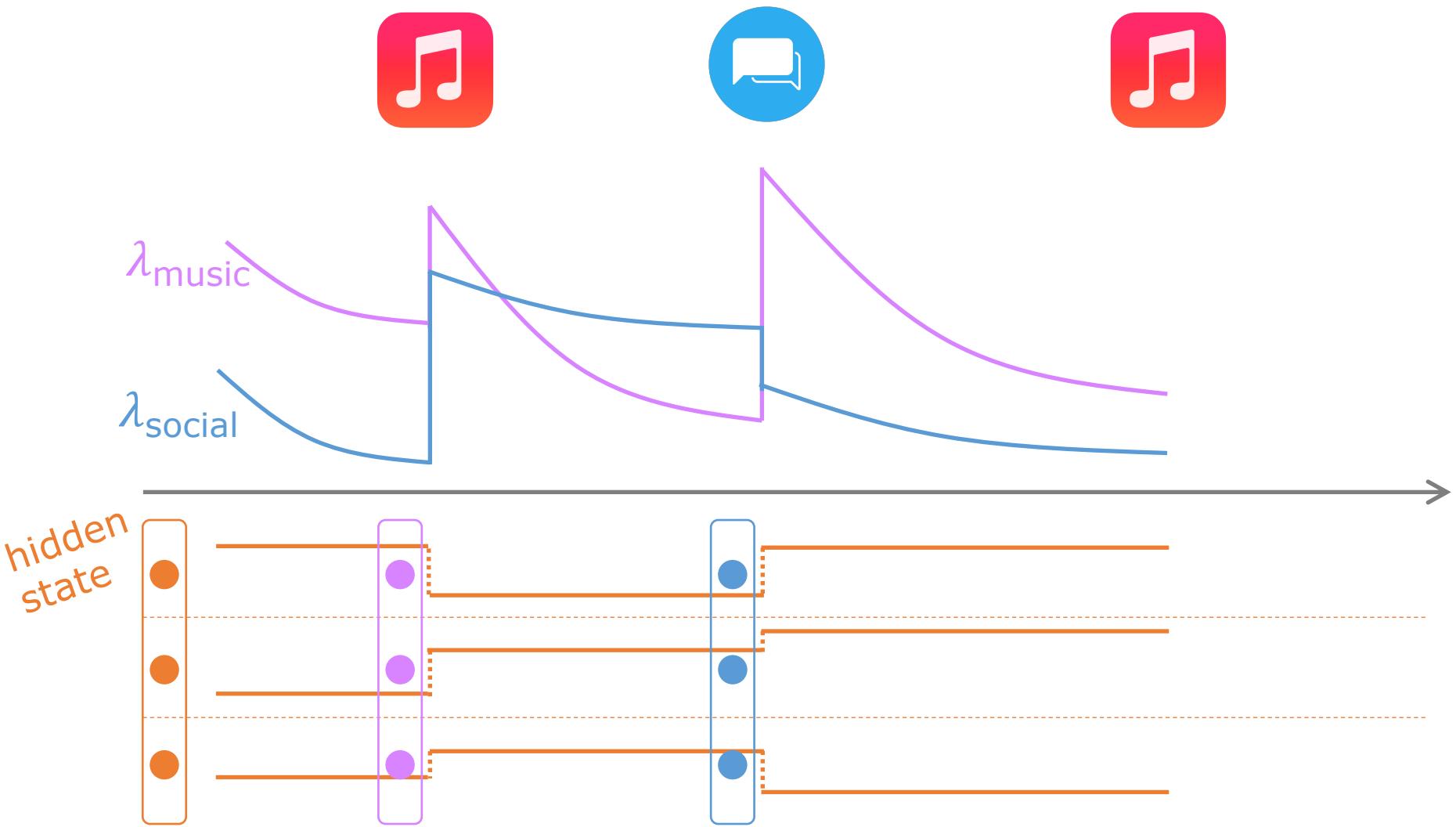
RNN/LSTM Transformer



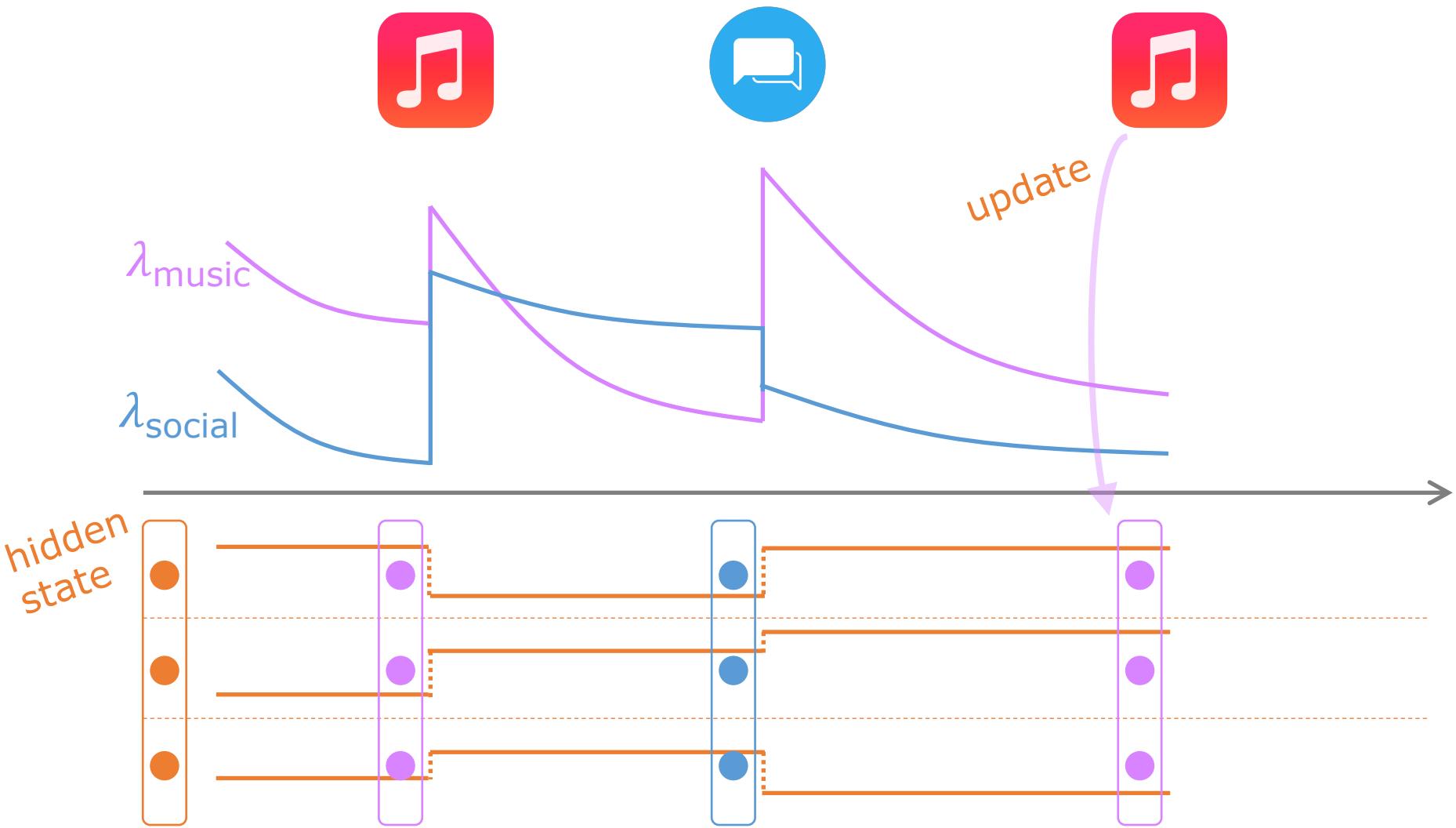
RNN/LSTM Transformer



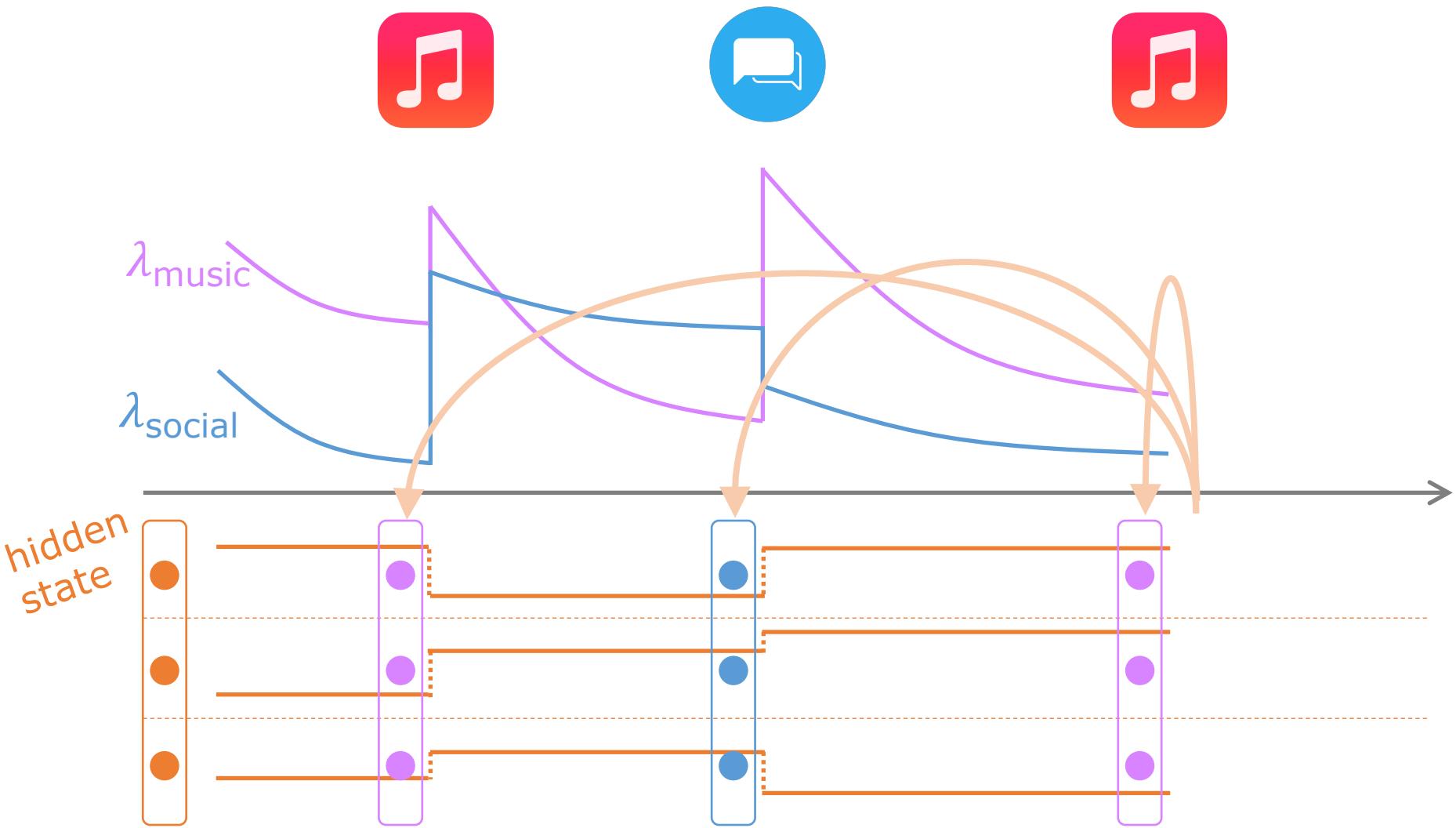
RNN/LSTM Transformer



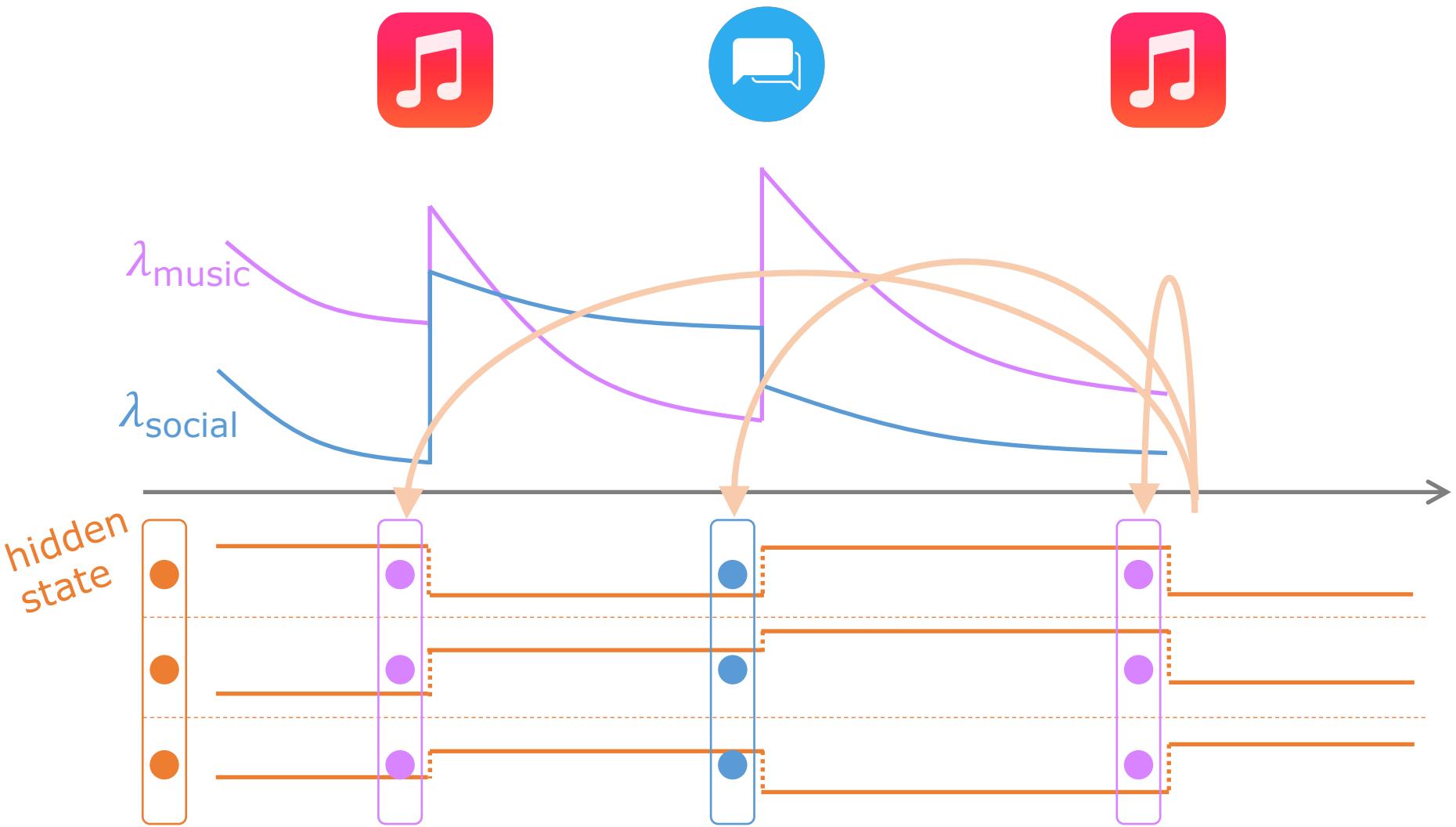
~~RNN/LSTM~~ Transformer



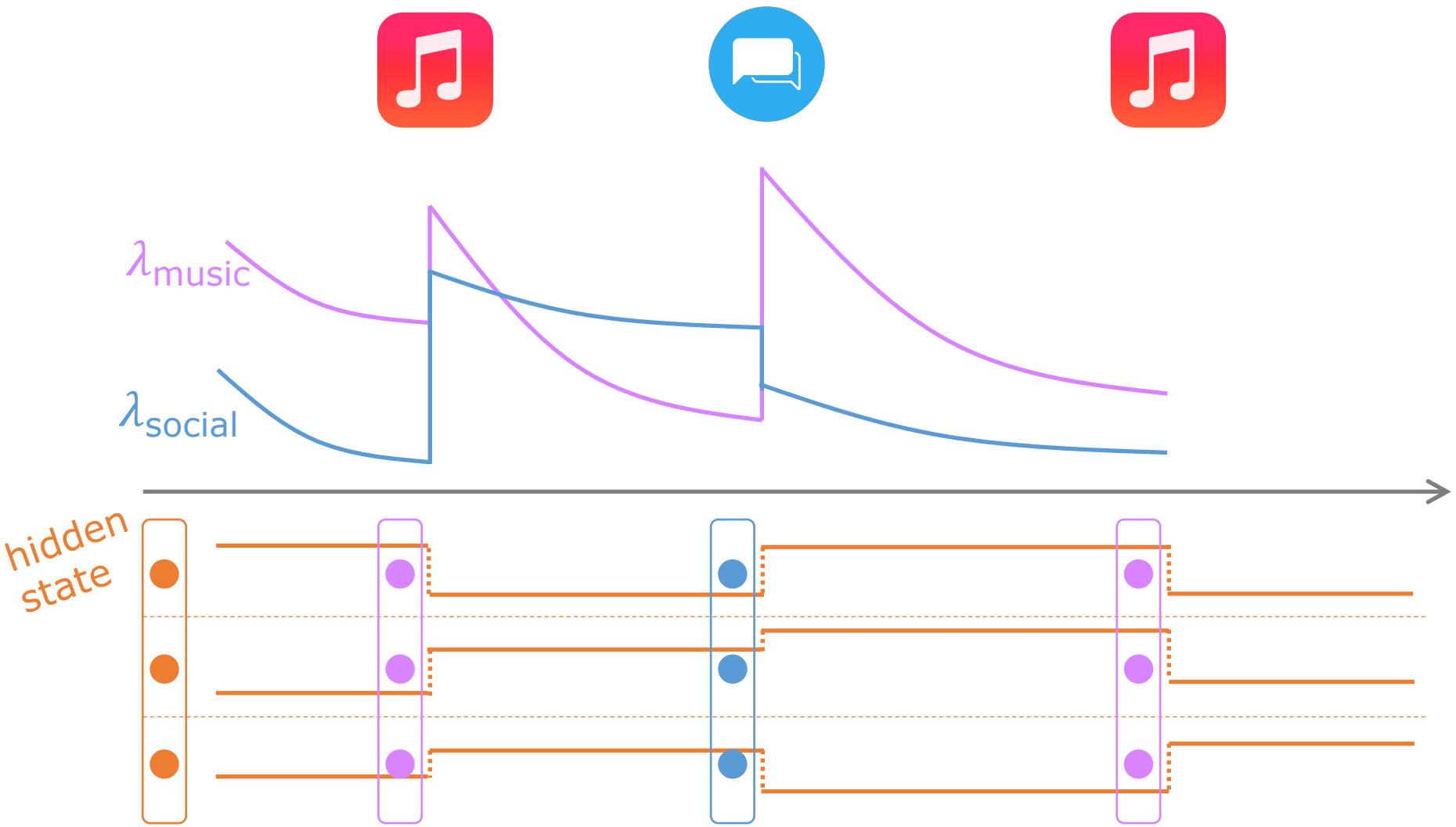
RNN/LSTM Transformer



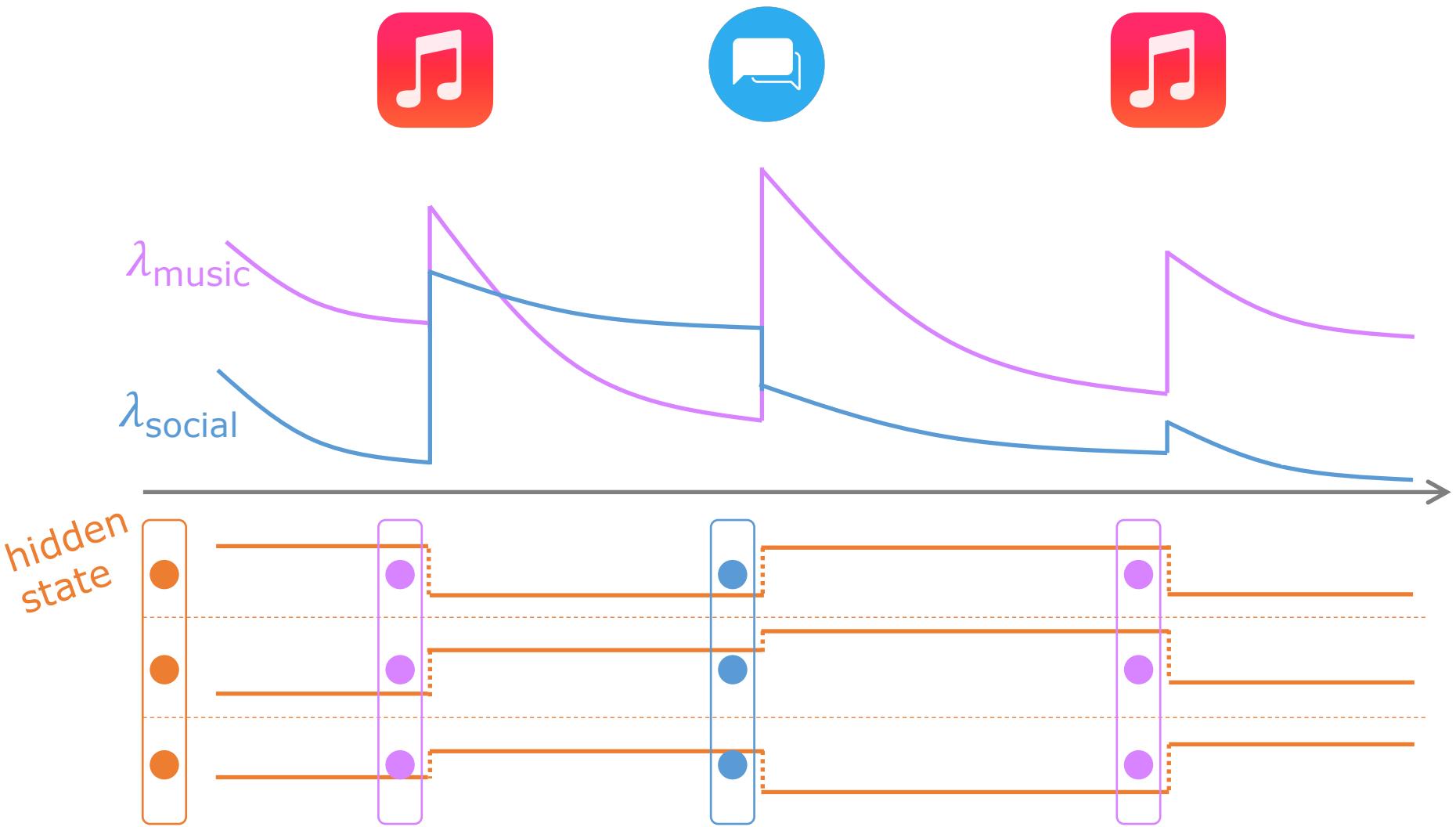
RNN/LSTM Transformer



RNN/LSTM Transformer

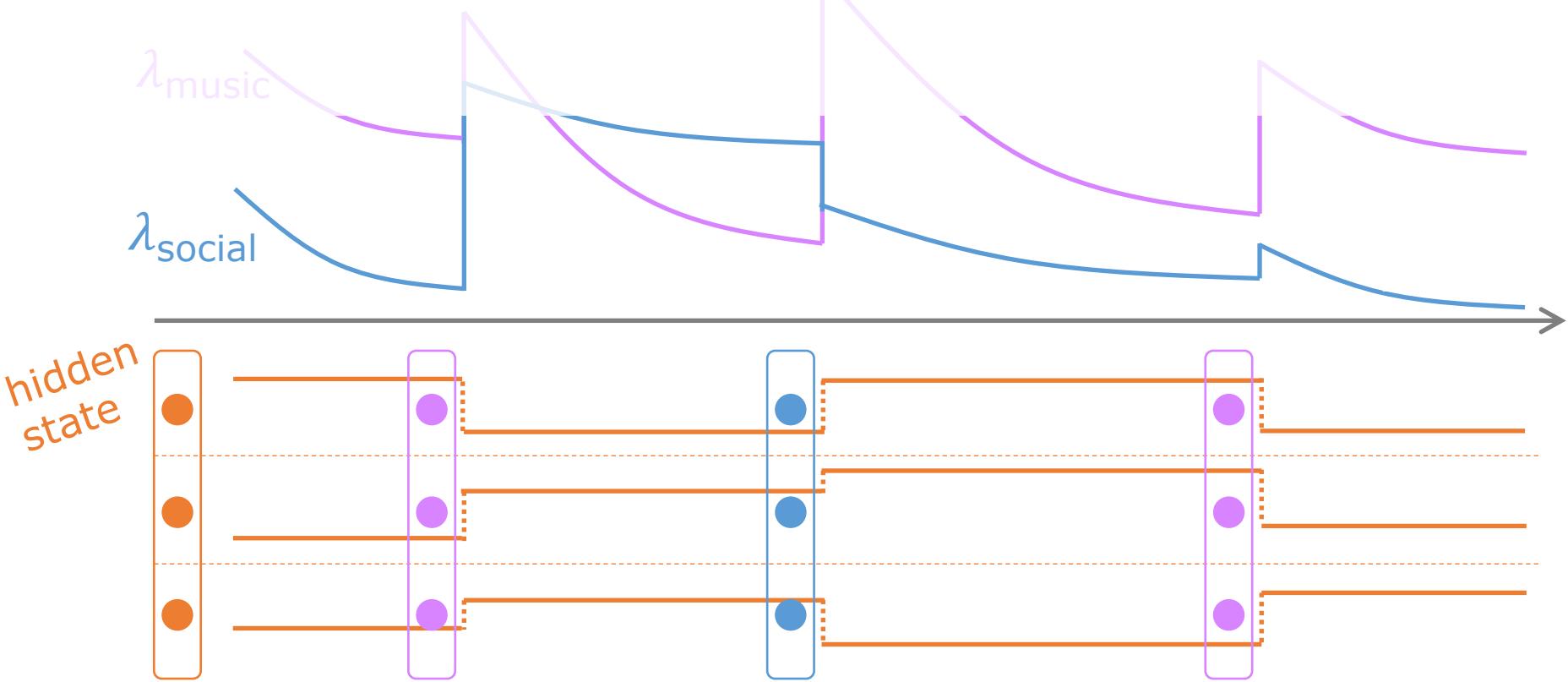


RNN/LSTM Transformer



RNN/LSTM Transformer

Transformer Hawkes Process (THP) Zuo et al 2020



Let's write some code

<http://bburl/tpp-lab-p2>

Any Questions?

<http://bburl/tpp-slides-p2>

<http://bburl/tpp-lab-p2>