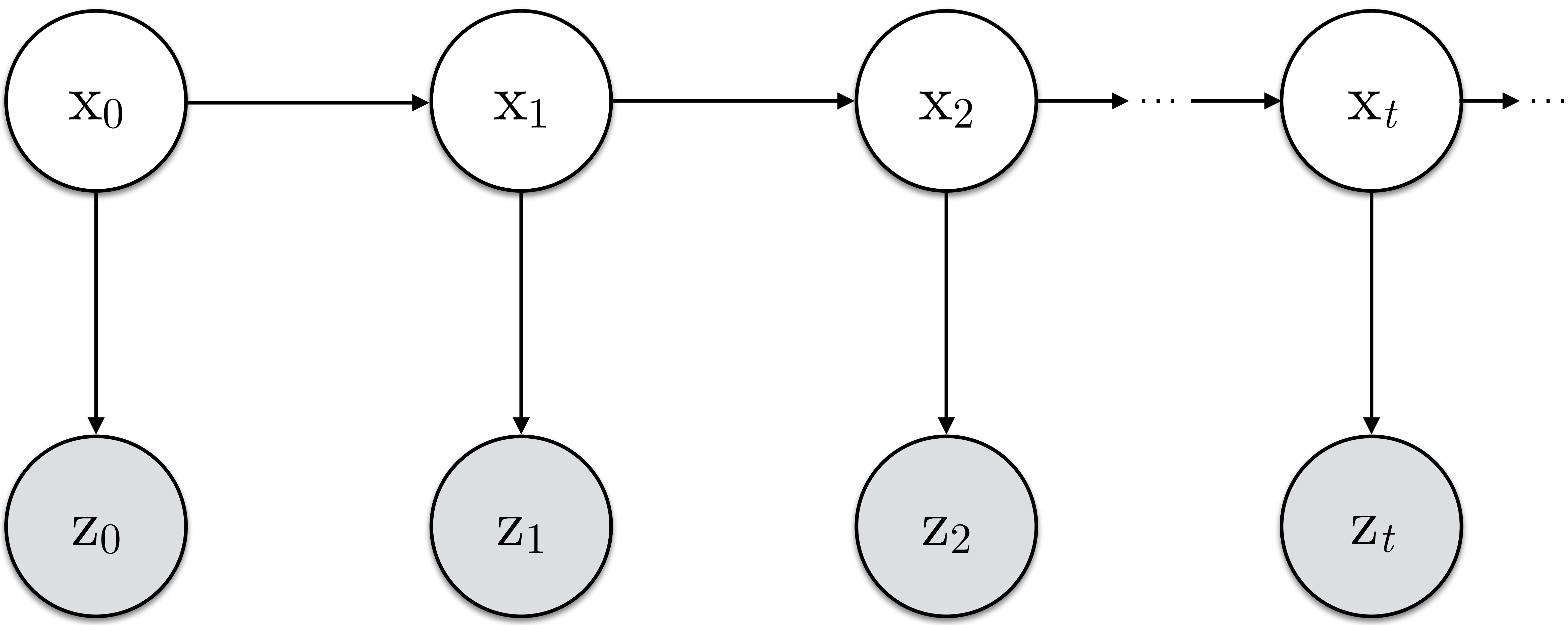
Planning, Learning and Intelligent Decision Making

Lecture 3

PADInt 2024

Hidden Markov model



Hidden Markov model

Markov state

The state at instant t is enough to predict the state at instant $t+1$:

$$\mathbb{P}\left[\mathrm{x}_{t+1} = y \mid \mathbf{x}_{0:t} = \boldsymbol{x}_{0:t}, \mathbf{z}_{0:t} = \boldsymbol{z}_{0:t}\right] = \mathbb{P}\left[\mathrm{x}_{t+1} = y \mid \mathrm{x}_t = x_t\right]$$

State-dependent observations

The state at instant t is enough to predict the observation at instant t:

$$\mathbb{P}\left[\mathrm{z}_t = z \mid \mathbf{x}_{0:t} = \boldsymbol{x}_{0:t}, \mathbf{z}_{0:t-1} = \boldsymbol{z}_{0:t-1}\right] = \mathbb{P}\left[\mathrm{z}_t = z \mid \mathrm{x}_t = x_t\right]$$

Estimation

- **Filtering:**

  Forward algorithm

  - Given a sequence of observations, estimate the final state

- **Smoothing:**

  Viterbi algorithm

  - Given a sequence of observations, estimate the sequence of states

- **Prediction:**

  - Given a sequence of observations, predict future states

Estimation

- **Filtering:**
  - Given a sequence of observations, estimate the final state
- **Smoothing:**
  - Given a sequence of observations, estimate the sequence of states
- **Prediction:**
  - Given a sequence of observations, predict future states

TÉCNICO LISBOA

Smoothing

Estimation

- **Filtering:**
  - Given a sequence of observations, estimate the final state
- **(Easier) Marginal smoothing:**
  - Given a sequence of observations, estimate **some state** in the middle
- **Prediction:**
  - Given a sequence of observations, predict future states

Smoothing

- We are given a sequence of observations $\boldsymbol{z}_{0:T}$

- We want to estimate, for $t < T$

What is some "middle" state?

$$\mathbb{P}_{\mu_0}\left[\mathrm{x}_t = x \mid \mathbf{z}_{0:T} = \boldsymbol{z}_{0:T}\right]$$

where μ_0 is the initial distribution, i.e.,

$$\mu_0(x) = \mathbb{P}\left[\mathrm{x}_0 = x\right]$$

Smoothing

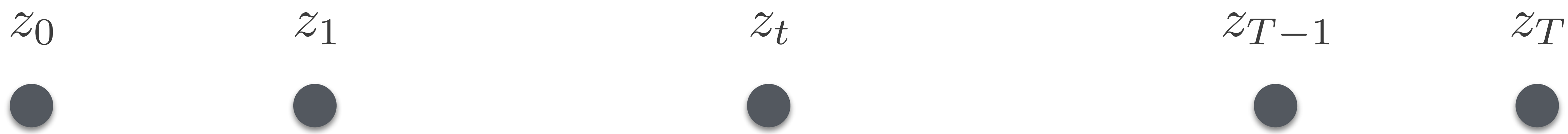
- We use the same notation:

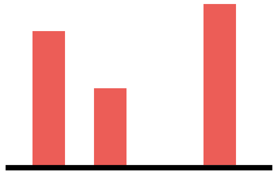
$$\mu_{t|0:T}(x) = \mathbb{P}_{\mu_0}\left[\mathrm{x}_t = x \mid \mathbf{z}_{0:T} = \boldsymbol{z}_{0:T}\right]$$

Distribution at time t

Given observations $0{:}T$

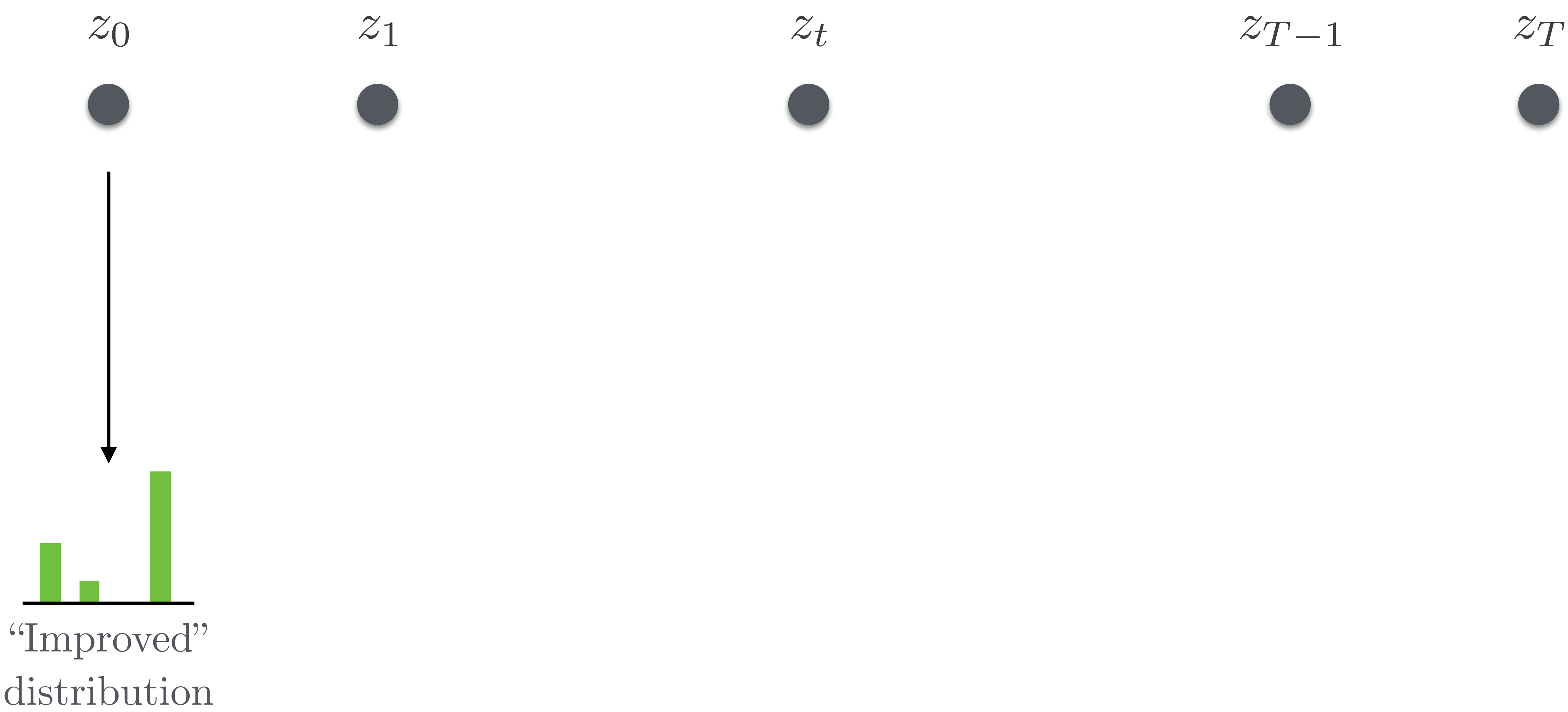
Similar idea



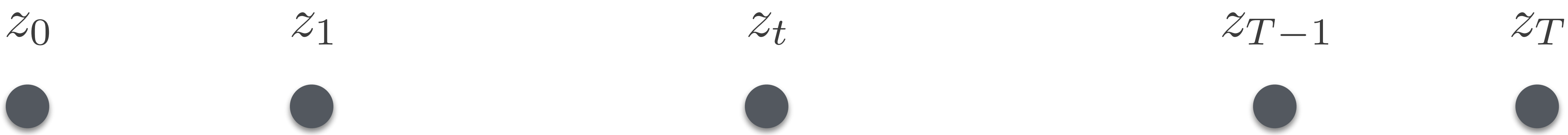


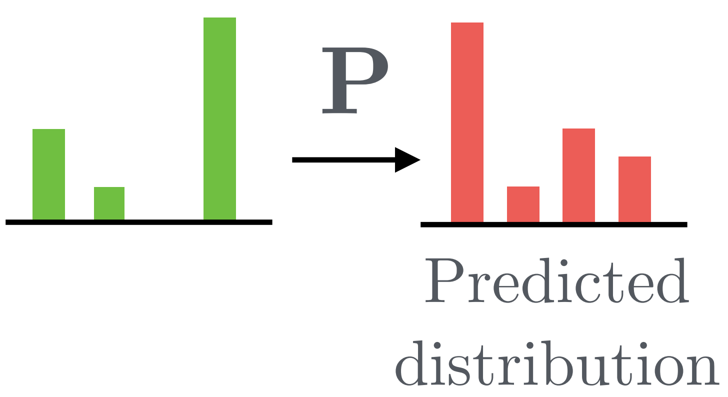
Initial
distribution

Similar idea

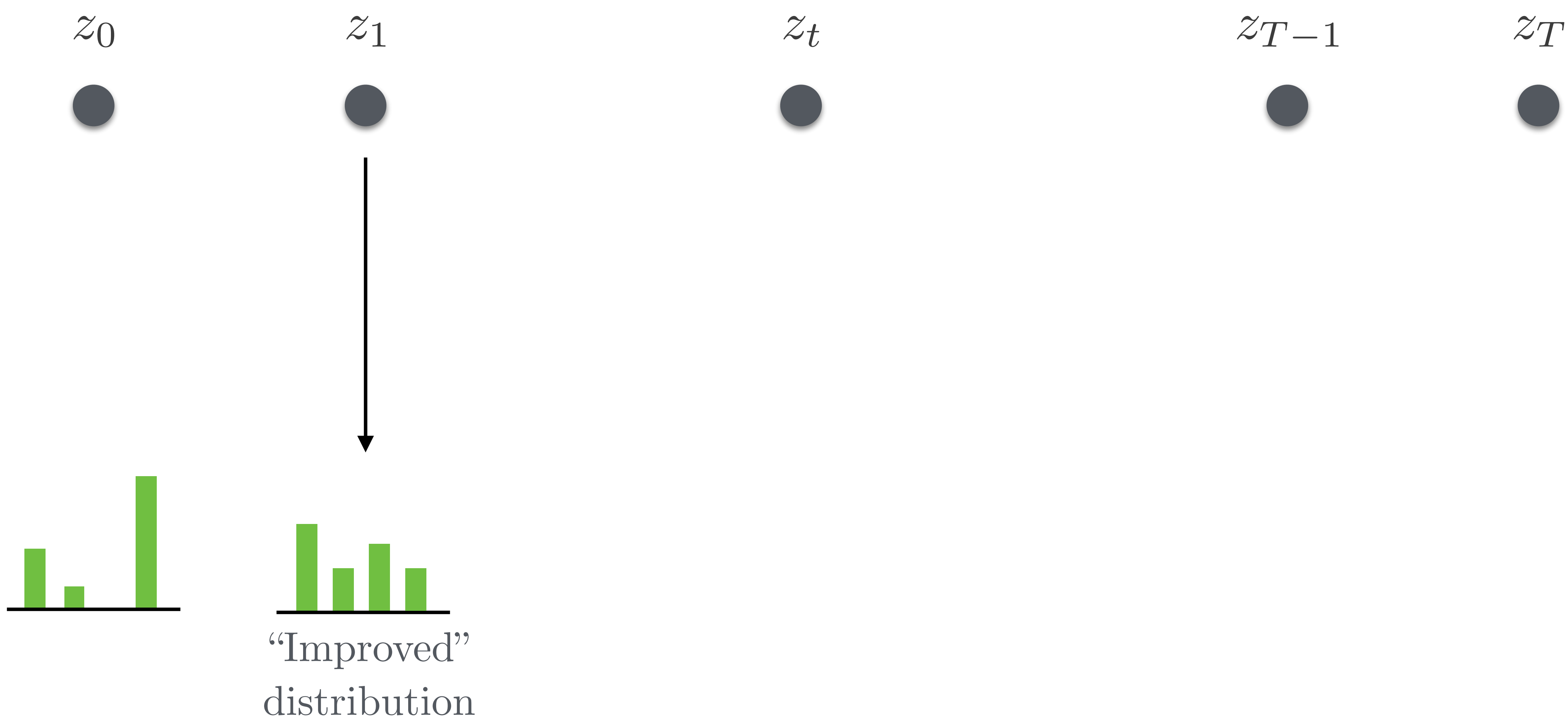


Similar idea

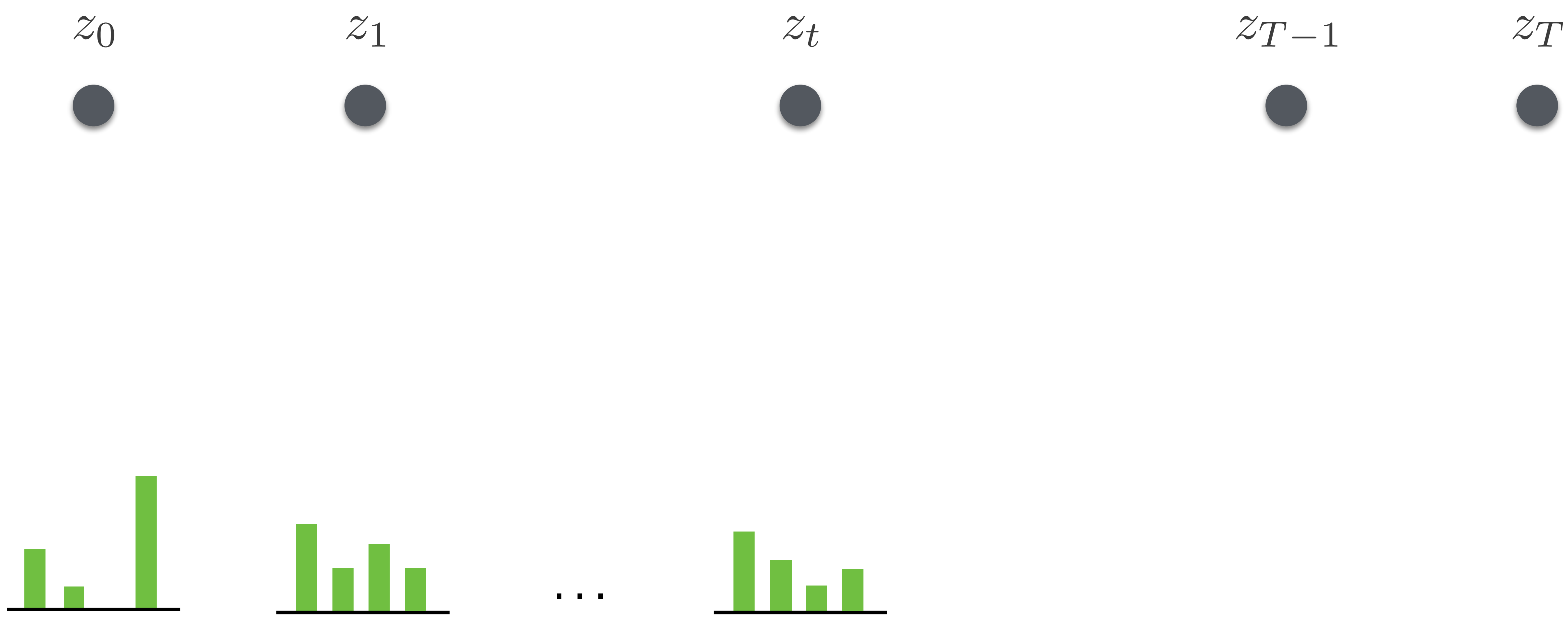




Similar idea

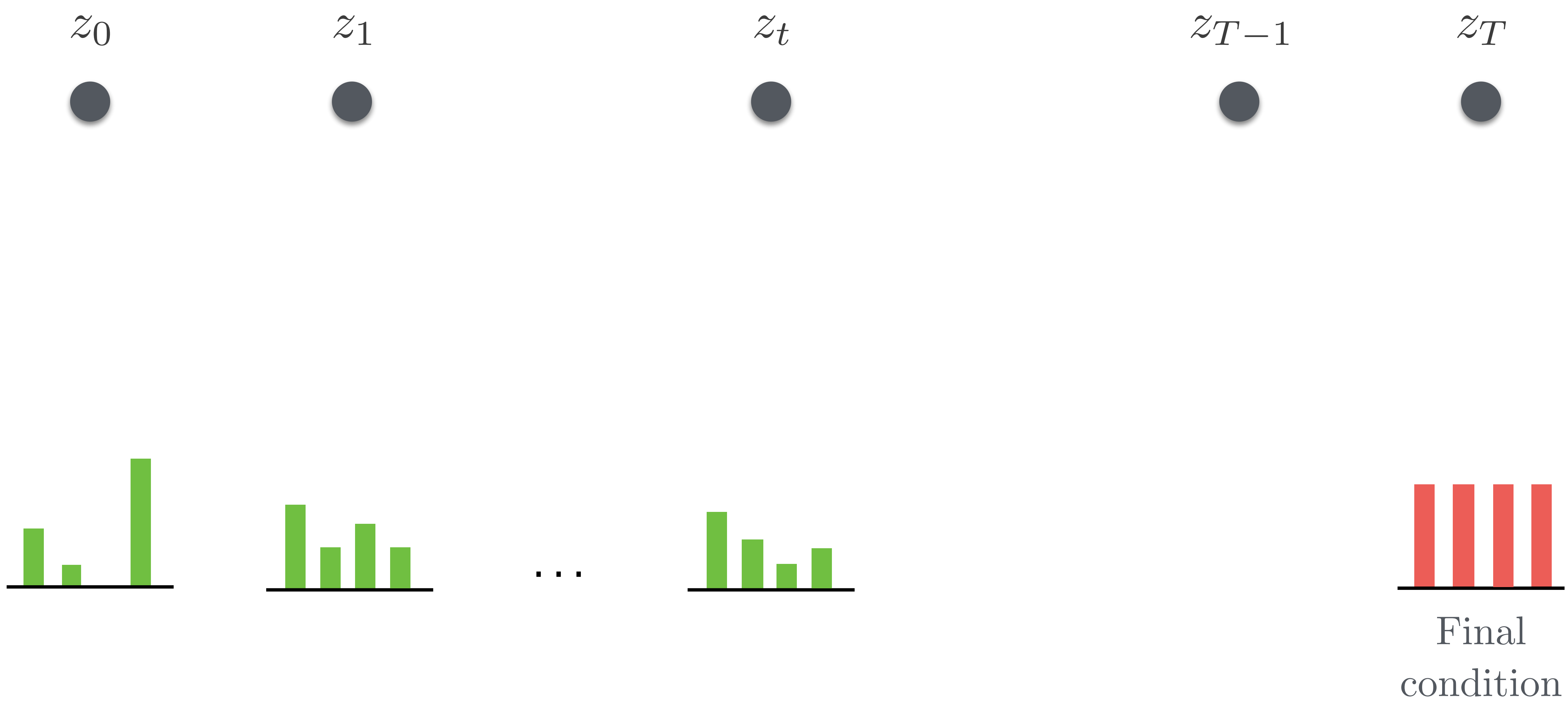


Similar idea

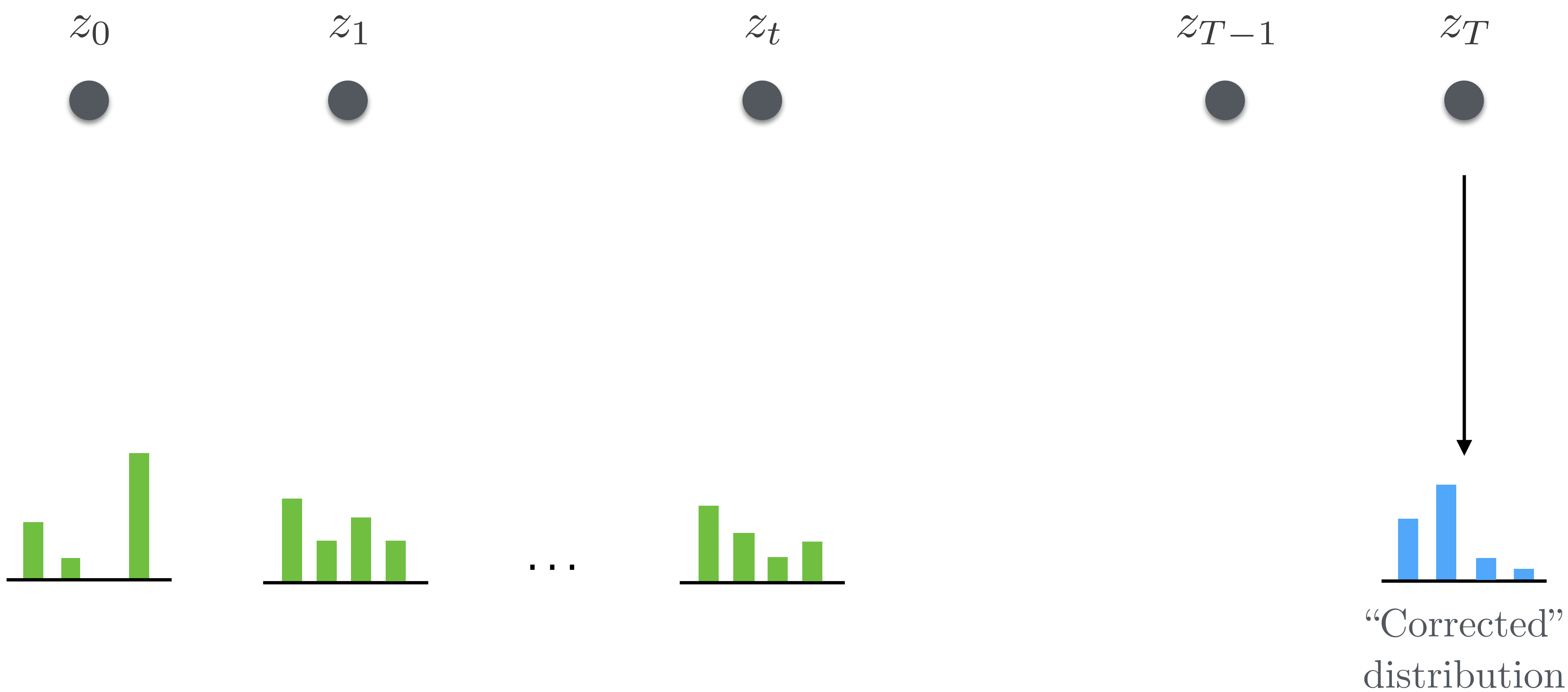


TÉCNICO LISBOA

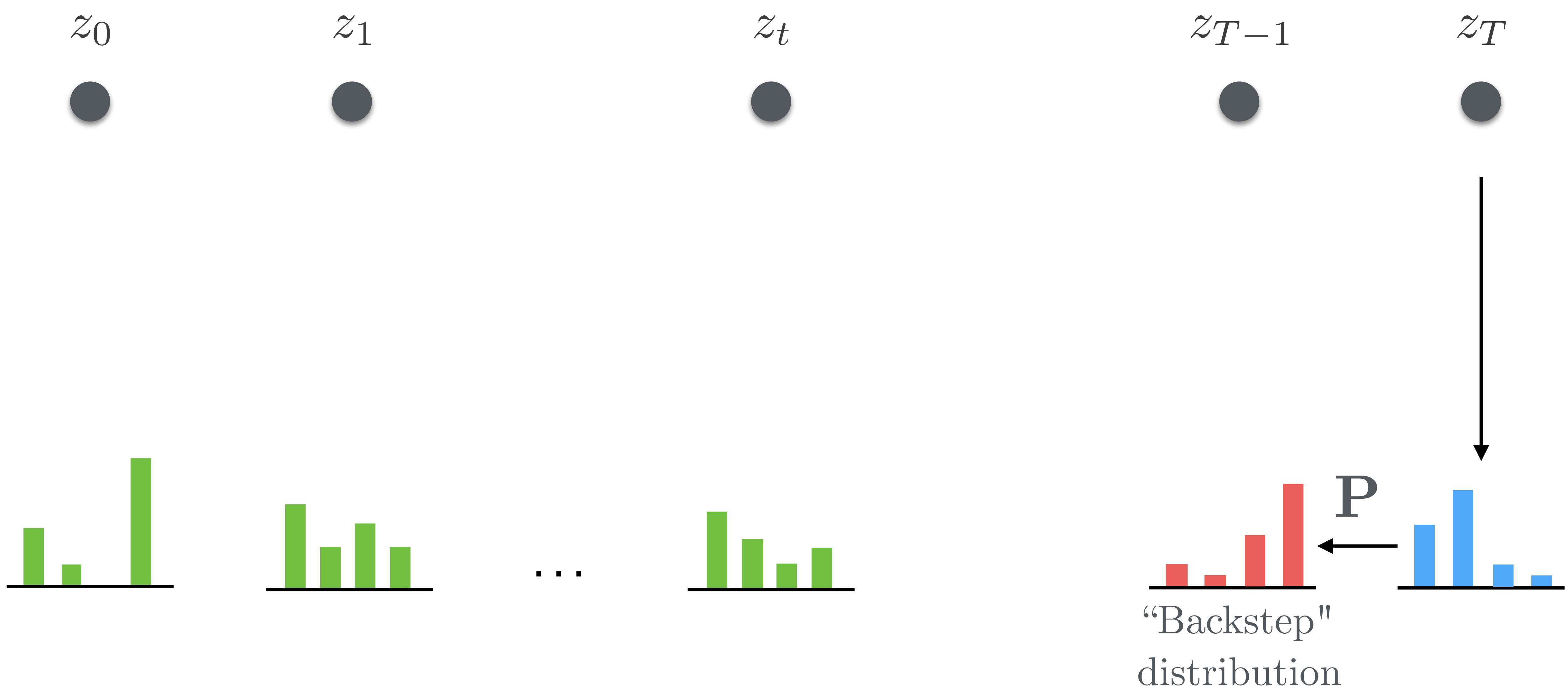
... with a twist



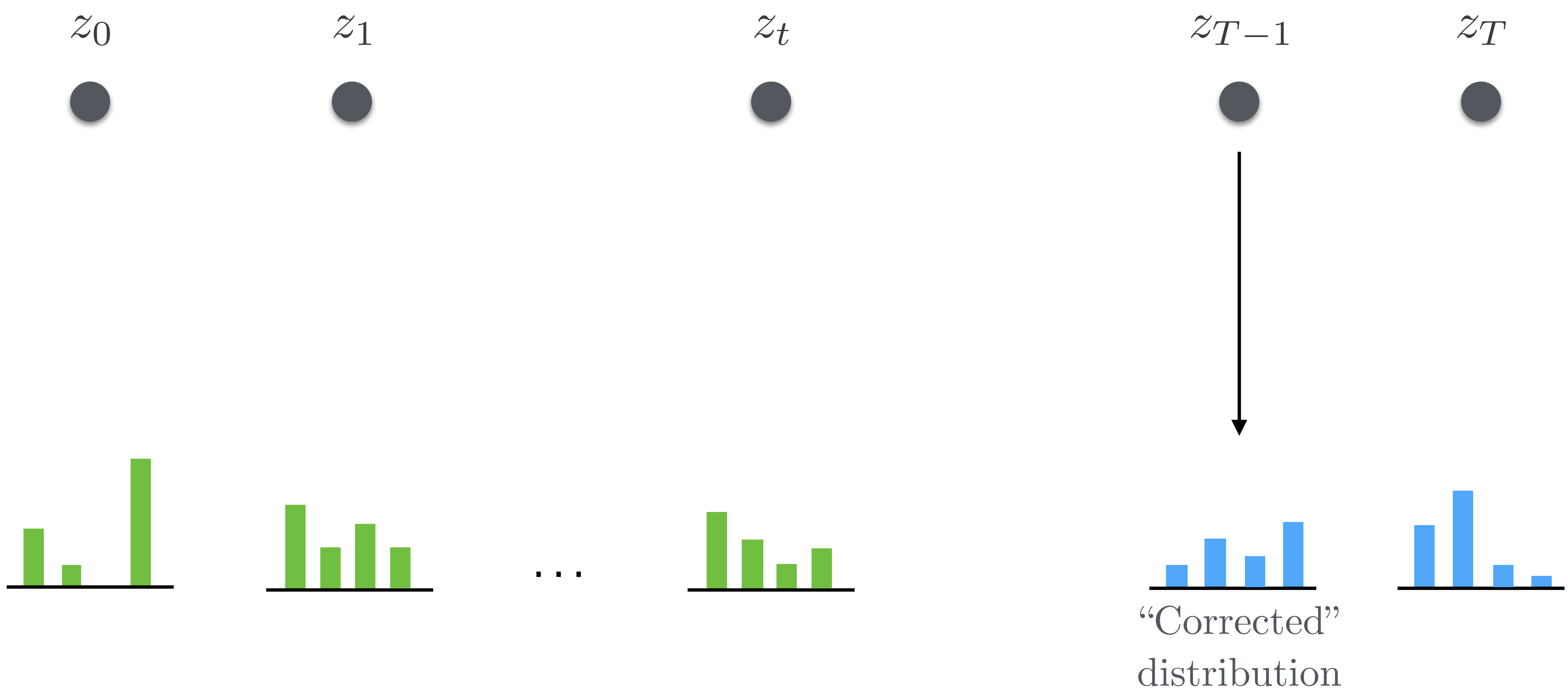
... with a twist



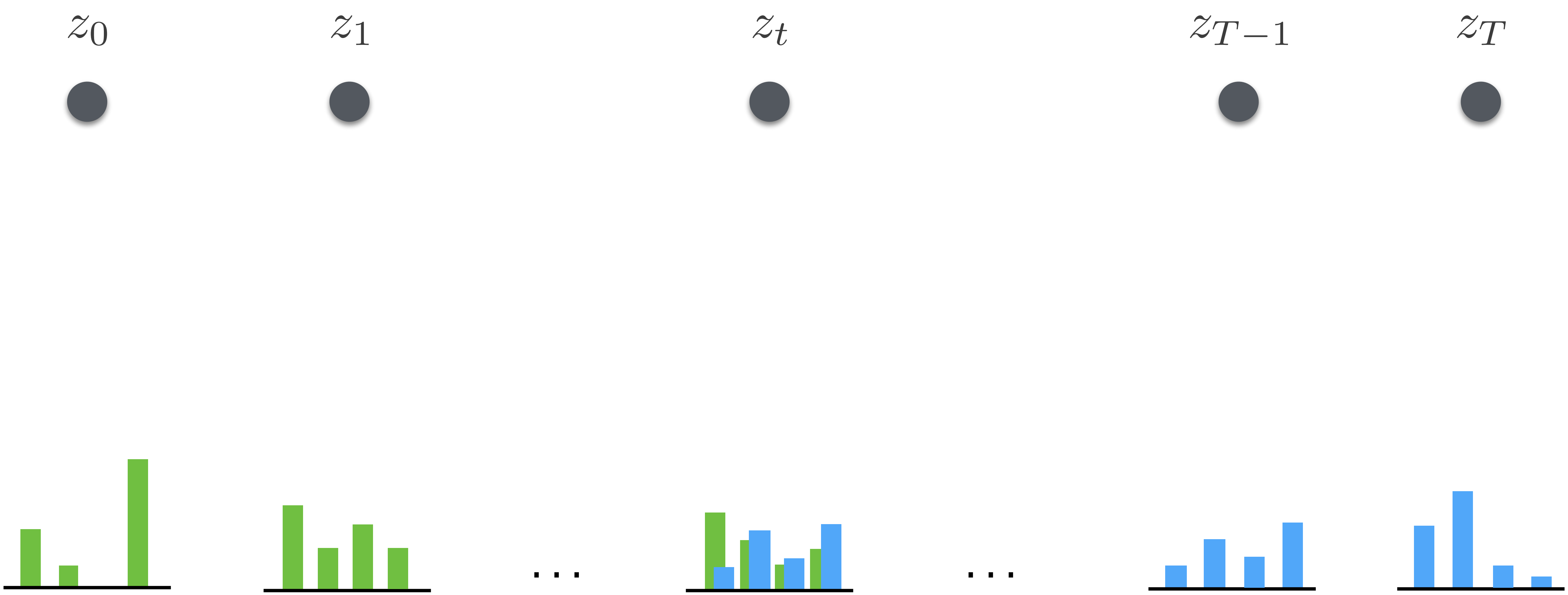
... with a twist



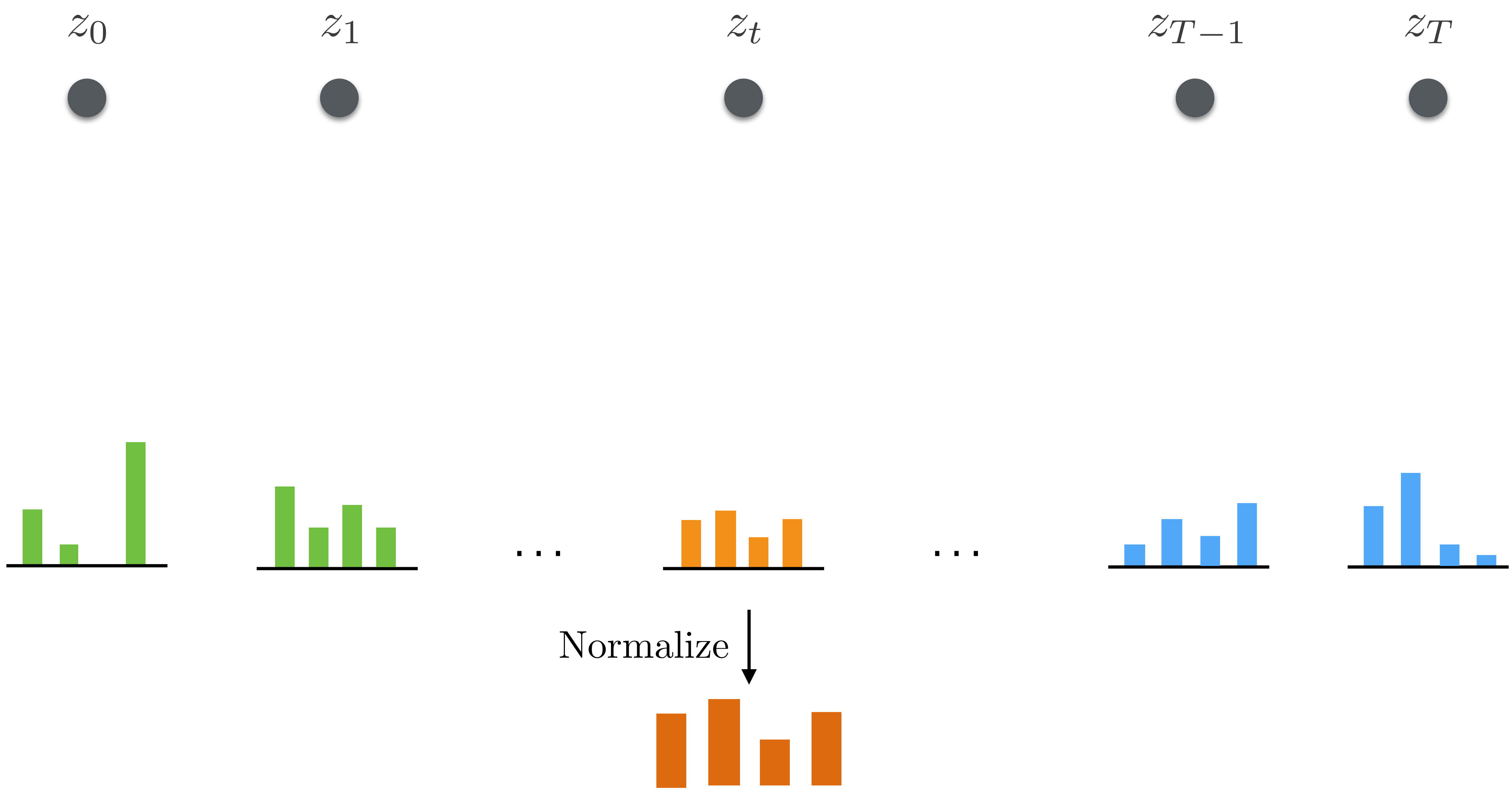
... with a twist



... with a twist



… with a twist



Backward mapping

Backward mapping

Given a sequence of observations $\boldsymbol{z}_{0:t}$, the backward mapping $\beta_t : \mathcal{X} \to \mathbb{R}$ is defined for each t as

$$\beta_t(x) = \mathbb{P}_{\mu_0}\left[\mathbf{z}_{t+1:T} = \boldsymbol{z}_{t+1:T} \;\middle|\; \mathrm{x}_t = x\right]$$

Conditional probability

How likely is it that I observe $\boldsymbol{z}_{t+1:T}$ knowing that I'm in x?

So what?

- Backward mapping has several useful properties

1. We can compute $\mu_{t|0:T}$ from α_t and β_t:

$$\mu_{t|0:T}(x) = \frac{\beta_t(x)\alpha_t(x)}{\sum_{y\in\mathcal{X}} \beta_t(y)\alpha_t(y)}$$

TÉCNICO LISBOA

So what?

- Backward mapping has several useful properties

1. We can compute $\mu_{t|0:T}$ from α_t and β_t

2. The backward mapping can be computed recursively:

$$\beta_t(x) = \sum_{y \in \mathcal{X}} \mathbf{O}(z_{t+1} \mid y)\beta_{t+1}(y)\mathbf{P}(y \mid x)$$

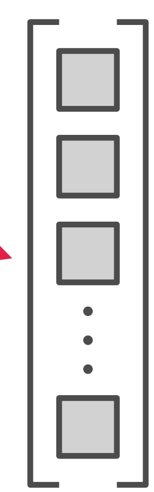
Vectorizing the backward update

- Let's look at the update

$$\beta_t(x) = \sum_{y \in \mathcal{X}} \mathbf{O}(z_{t+1} \mid y)\beta_{t+1}(y)\mathbf{P}(y \mid x)$$

using vectors:

We start with our vector at time $t+1$



β_{t+1}

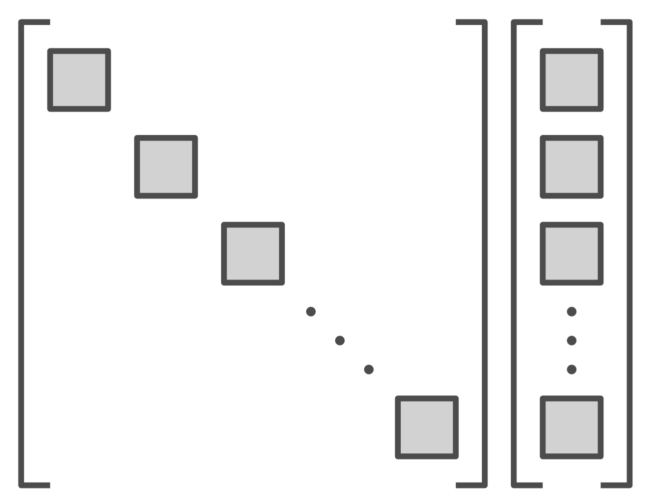
Vectorizing the backward update

- Let's look at the update

$$\beta_t(x) = \sum_{y \in \mathcal{X}} \mathbf{O}(z_{t+1} \mid y)\beta_{t+1}(y)\mathbf{P}(y \mid x)$$

using vectors:

We correct with the observation at time $t + 1$



$\mathrm{diag}(\mathbf{O}(z_{t+1} \mid \cdot))$ β_{t+1}

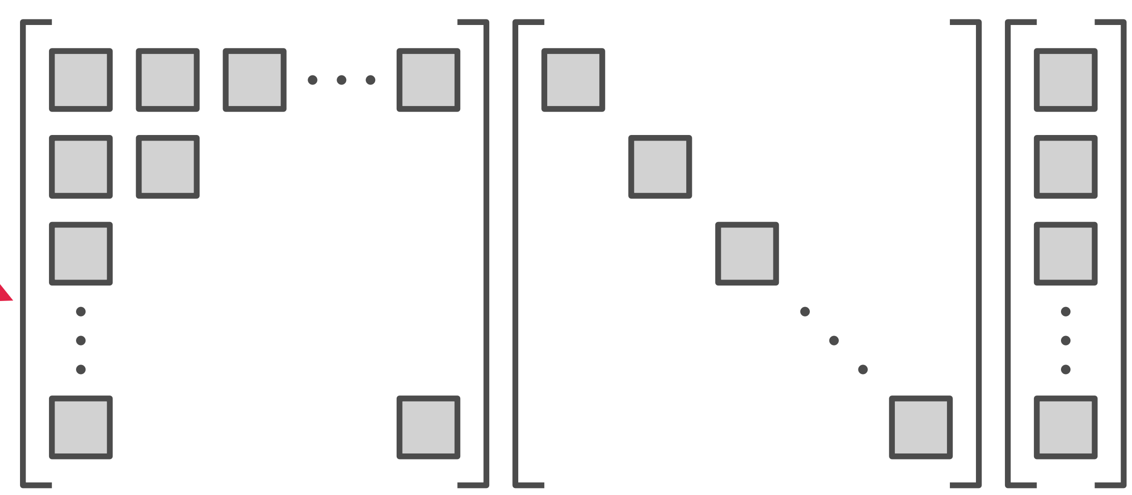
Vectorizing the backward update

- Let's look at the update

$$\beta_t(x) = \sum_{y \in \mathcal{X}} \mathbf{O}(z_{t+1} \mid y)\beta_{t+1}(y)\mathbf{P}(y \mid x)$$

using vectors:

We take a step back



$\mathbf{P}$ $\quad$ $\mathrm{diag}(\mathbf{O}(z_{t+1} \mid \cdot))$ $\quad$ β_{t+1}

TÉCNICO LISBOA

Forward-backward algorithm

Require: Observation sequence $\boldsymbol{z}_{0:T}$

1. Initialize $\boldsymbol{\alpha}_0 = \text{diag}(\mathbf{O}(z_0 \mid \cdot)\boldsymbol{\mu}_0^\top, \boldsymbol{\beta}_T = 1$

2. For $\tau = 1, \ldots, t$ do

$$\boldsymbol{\alpha}_\tau = \text{diag}(\mathbf{O}(z_\tau \mid \cdot))\mathbf{P}^\top \boldsymbol{\alpha}_{\tau-1}$$

Forward update

3. end for

4. For $\tau = T-1, \ldots, t$ do

$$\boldsymbol{\beta}_\tau = \mathbf{P}\text{diag}(\mathbf{O}(z_{\tau+1} \mid \cdot))\boldsymbol{\beta}_{\tau+1}$$

Backward update

5. end for

return $\boldsymbol{\alpha}_t \odot \boldsymbol{\beta}_t / \text{sum}(\alpha_\text{t} \odot \boldsymbol{\beta}_\text{t})$

Combine & normalize

Example: The urn problem

- Suppose that

$$\boldsymbol{\mu}_0 = \begin{bmatrix} 0.125 & 0.375 & 0.375 & 0.125 \end{bmatrix}$$

- We observe the sequence of observations

$$\boldsymbol{z}_{0:2} = \{w, w, b\}$$

- What is the state at time $t = 1$?

TÉCNICO LISBOA

Step 1: Initialize $\boldsymbol{\alpha}_0$

- $\boldsymbol{\alpha}_0 = \mathrm{diag}(\mathbf{O}(z_0 \mid \cdot)\boldsymbol{\mu}_0^\top$

$$\begin{array}{c} \begin{bmatrix} \boxed{\begin{matrix} 1 \\ 0 \\ 1 \\ 0 \end{matrix}} & \begin{matrix} 0 \\ 1 \\ 0 \\ 1 \end{matrix} \end{bmatrix} \\ \quad z_0 \qquad\quad \end{array} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.125 \\ 0.375 \\ 0.375 \\ 0.125 \end{bmatrix} \rightarrow \begin{bmatrix} 0.125 \\ 0 \\ 0.375 \\ 0 \end{bmatrix}$$

Step 2: Compute $\boldsymbol{\alpha}_1$

- $\boldsymbol{\alpha}_1 = \mathrm{diag}(\mathbf{O}(z_1 \mid \cdot))\mathbf{P}^\top \boldsymbol{\alpha}_0$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.2 & 0.2 & 0.05 & 0.05 \\ 0.6 & 0.6 & 0.15 & 0.15 \\ 0.15 & 0.15 & 0.6 & 0.6 \\ 0.05 & 0.05 & 0.2 & 0.2 \end{bmatrix} \begin{bmatrix} 0.125 \\ 0 \\ 0.375 \\ 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0.04 \\ 0 \\ 0.24 \\ 0 \end{bmatrix}$$

Step 3: Compute $\boldsymbol{\beta}_1$

- $\boldsymbol{\beta}_1 = \mathbf{P}\mathrm{diag}(\mathbf{O}(z_2 \mid \cdot))\boldsymbol{\beta}_2$

$$\begin{bmatrix} 0.2 & 0.6 & 0.15 & 0.05 \\ 0.2 & 0.6 & 0.15 & 0.05 \\ 0.05 & 0.15 & 0.6 & 0.2 \\ 0.05 & 0.15 & 0.6 & 0.2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} 0.65 \\ 0.65 \\ 0.35 \\ 0.35 \end{bmatrix}$$

Final step: Compute $\boldsymbol{\mu}_{1 \mid 0:2}$

- We finally get:

$$\begin{aligned}\mu_{1|0:2} &= \boldsymbol{\alpha}_1 \odot \boldsymbol{\beta}_1 / \mathrm{sum}(\boldsymbol{\alpha}_1 \odot \boldsymbol{\beta}_1) \\ &= \begin{bmatrix} 0.25 & 0 & 0.75 & 0 \end{bmatrix}\end{aligned}$$

TÉCNICO LISBOA

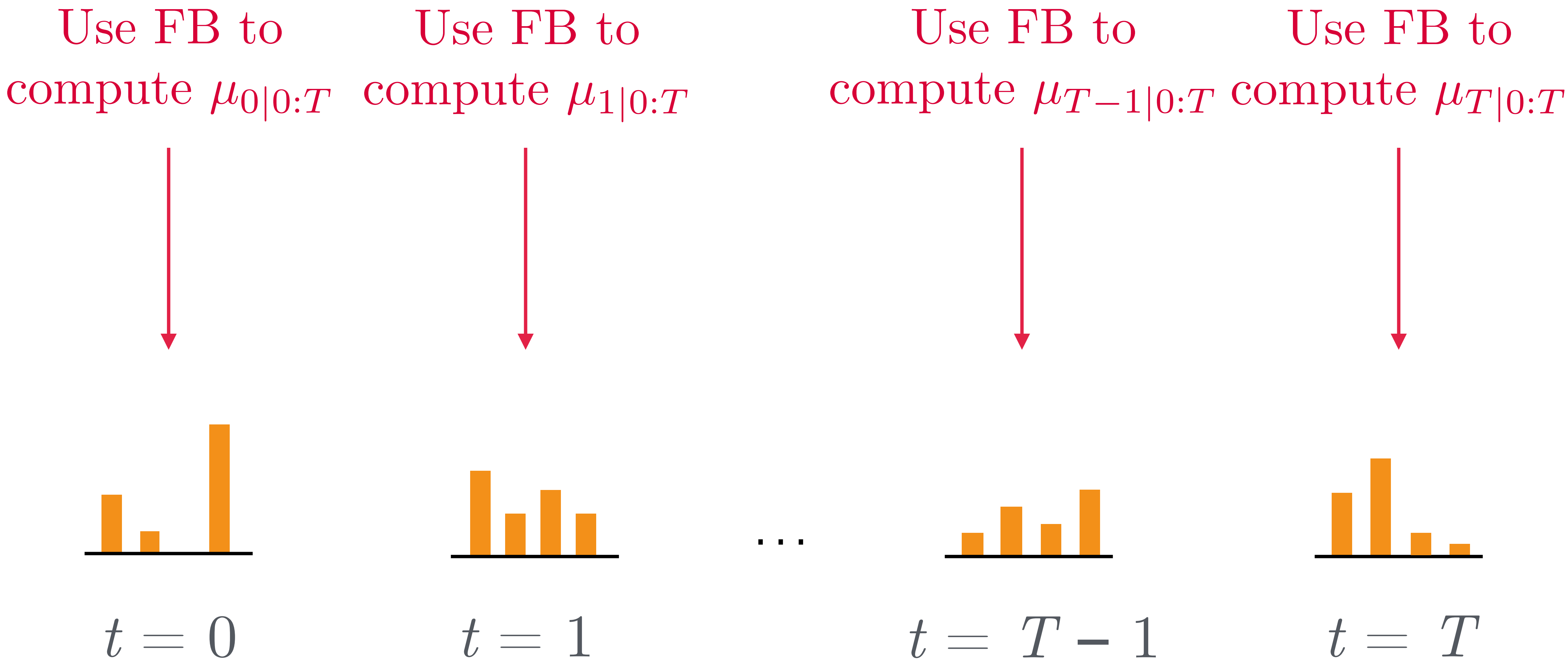
Estimation

- **Filtering:**
  - Given a sequence of observations, estimate the final state
- **(Joint) Smoothing:**
  - Given a sequence of observations, estimate the **whole** sequence of states (most likely sequence)
- **Prediction:**
  - Given a sequence of observations, predict future states

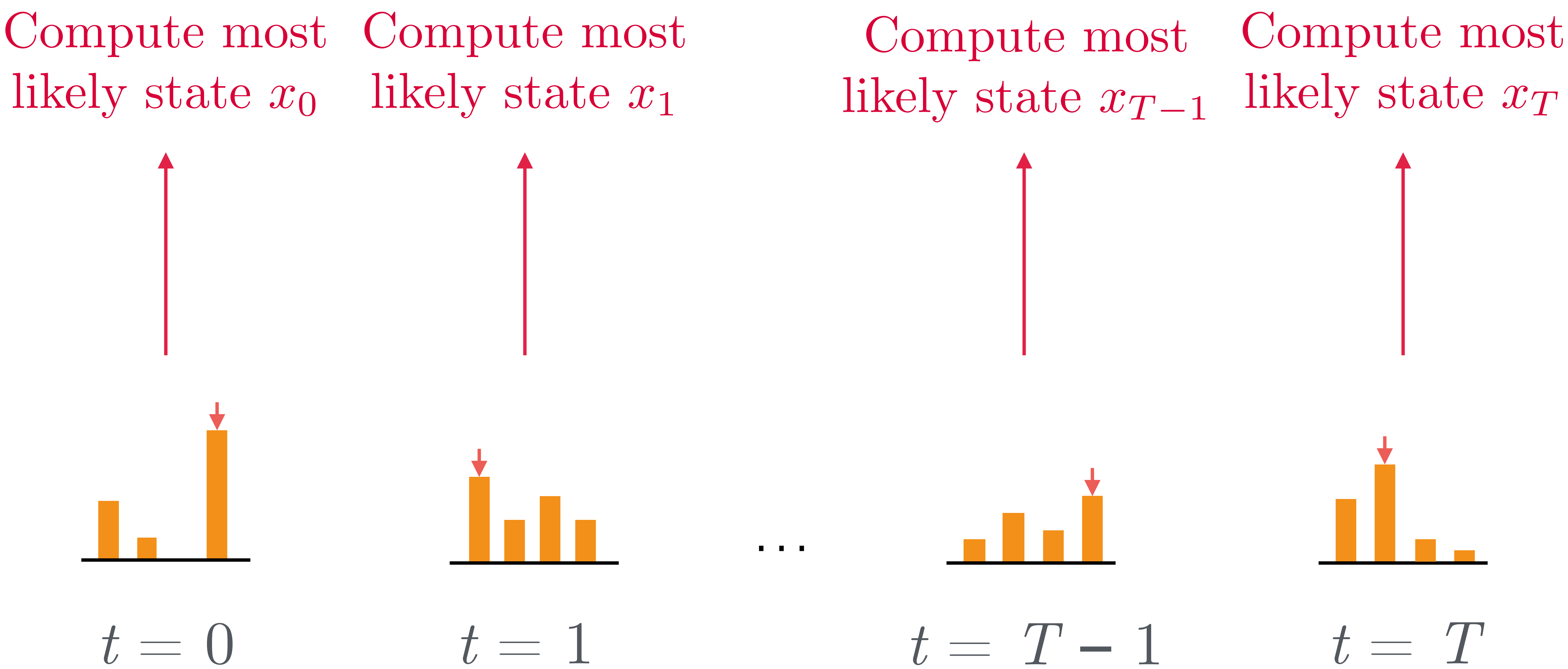
Any ideas?

TÉCNICO LISBOA

Naive approach



Naive approach

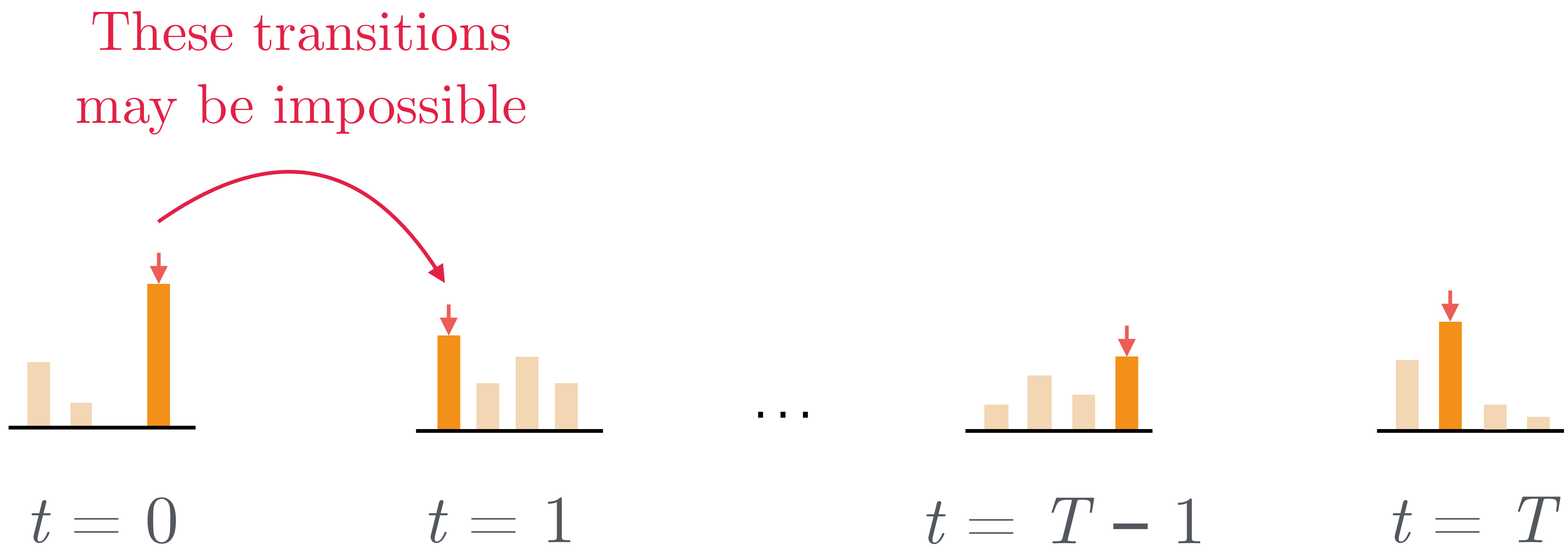


Naive approach

$t = 0$ $\qquad$ $t = 1$ $\qquad \cdots \qquad$ $t = T - 1$ $\qquad$ $t = T$

Problem?

Inconsistency...



Smoothing

- We are given a sequence of observations $\boldsymbol{z}_{0:T}$

- We want to estimate the **most likely sequence**, i.e.,

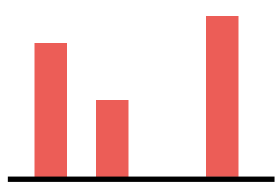
$$\boldsymbol{x}^*_{0:T} = \underset{\boldsymbol{x}_{0:T}}{\operatorname{argmax}} \mathbb{P}_{\mu_0} \left[\mathbf{x}_{0:T} = \boldsymbol{x}_{0:T} \mid \mathbf{z}_{0:T} = \boldsymbol{z}_{0:T} \right]$$

where $\boldsymbol{\mu}_0$ is the initial distribution, i.e.,

$$\mu_0(x) = \mathbb{P}\left[\mathrm{x}_0 = x \right]$$

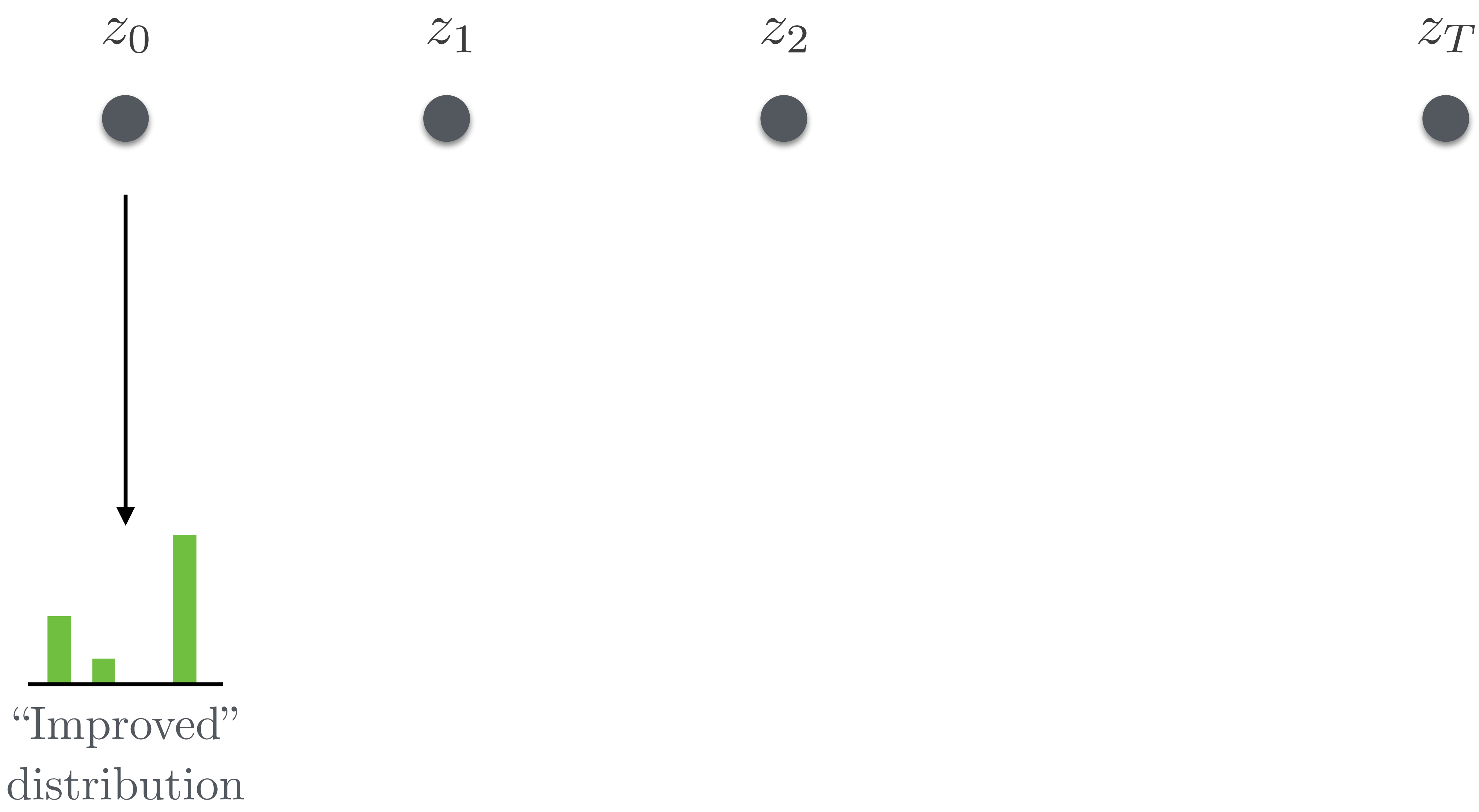
Idea

z_0 z_1 z_2 z_T



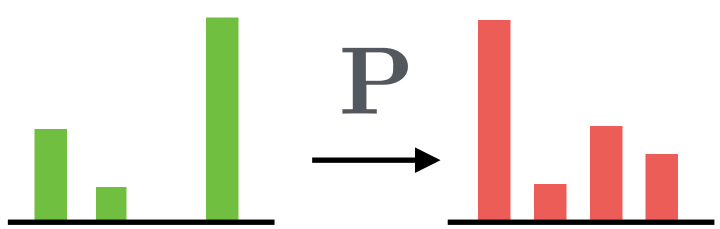
Initial distribution

Idea



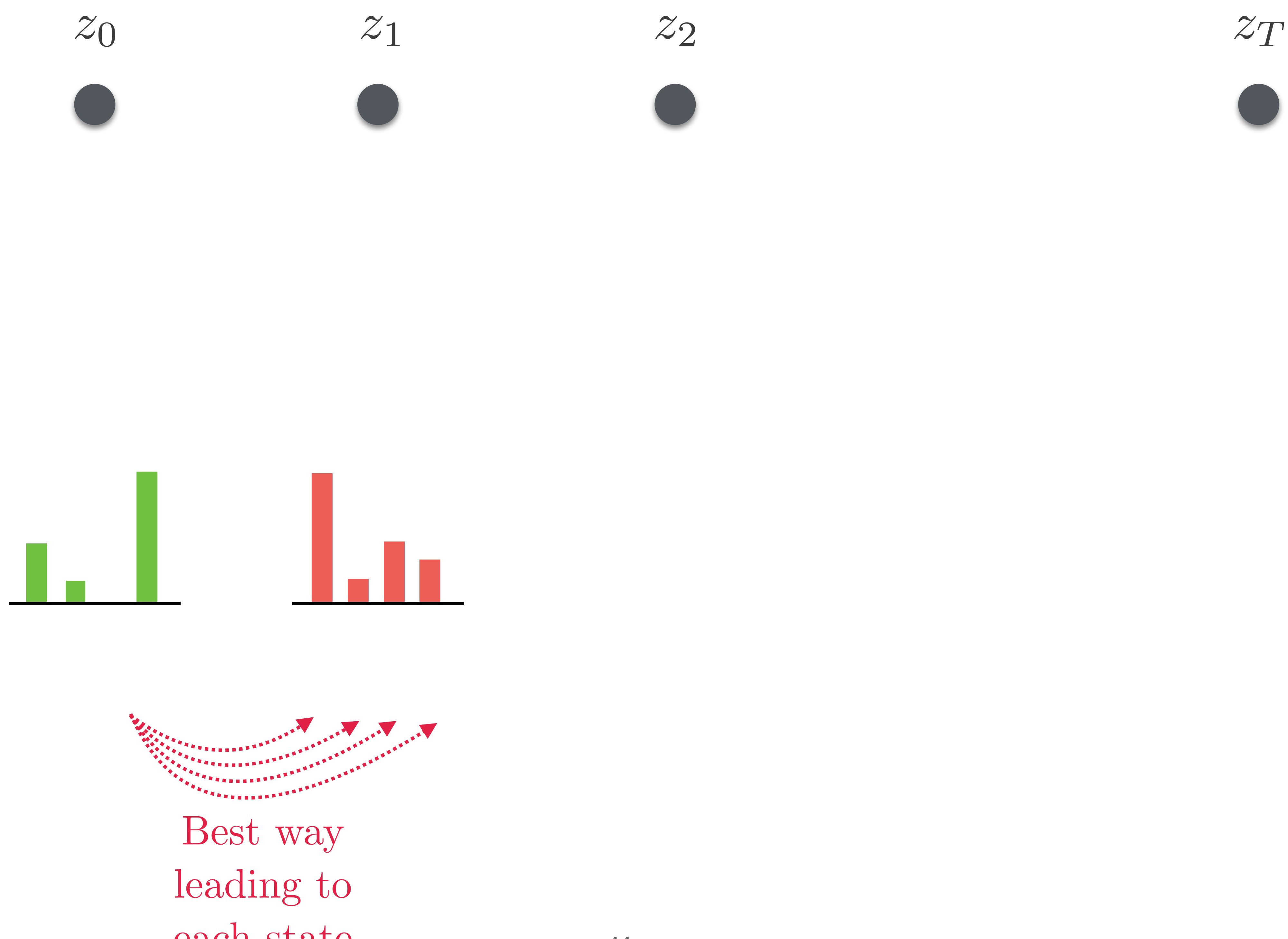
Idea

z_0 z_1 z_2 z_T

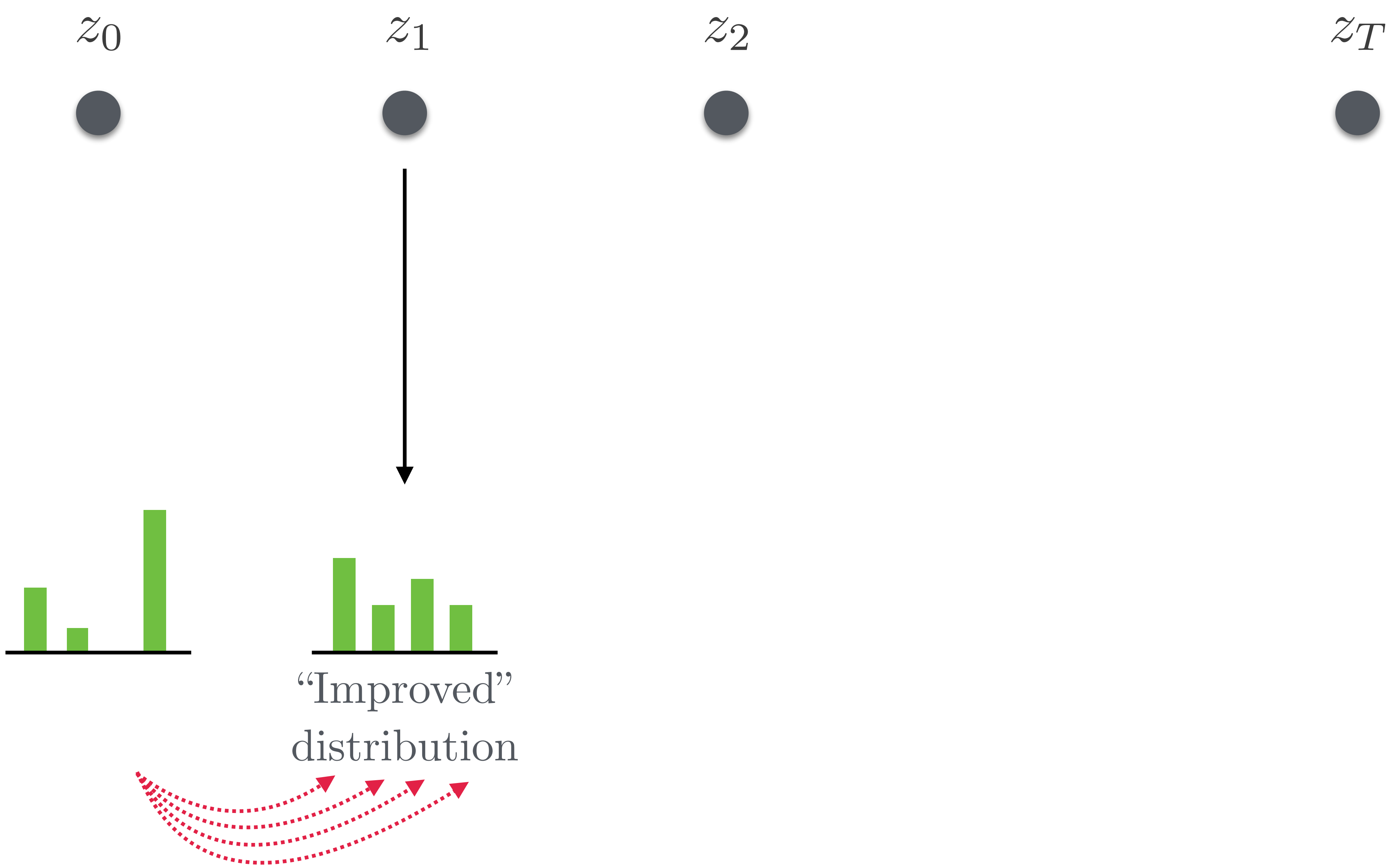


Predicted distribution

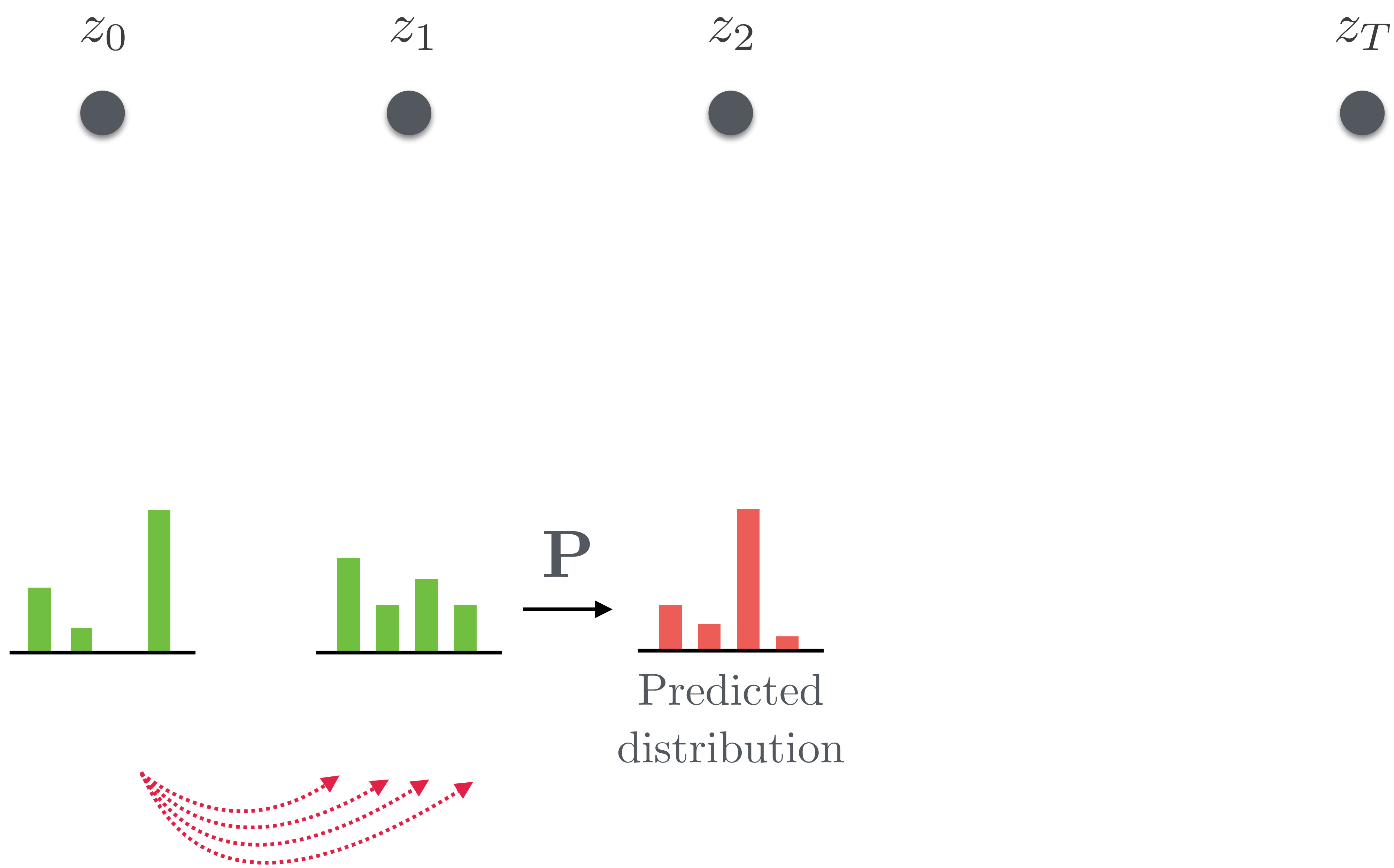
Idea



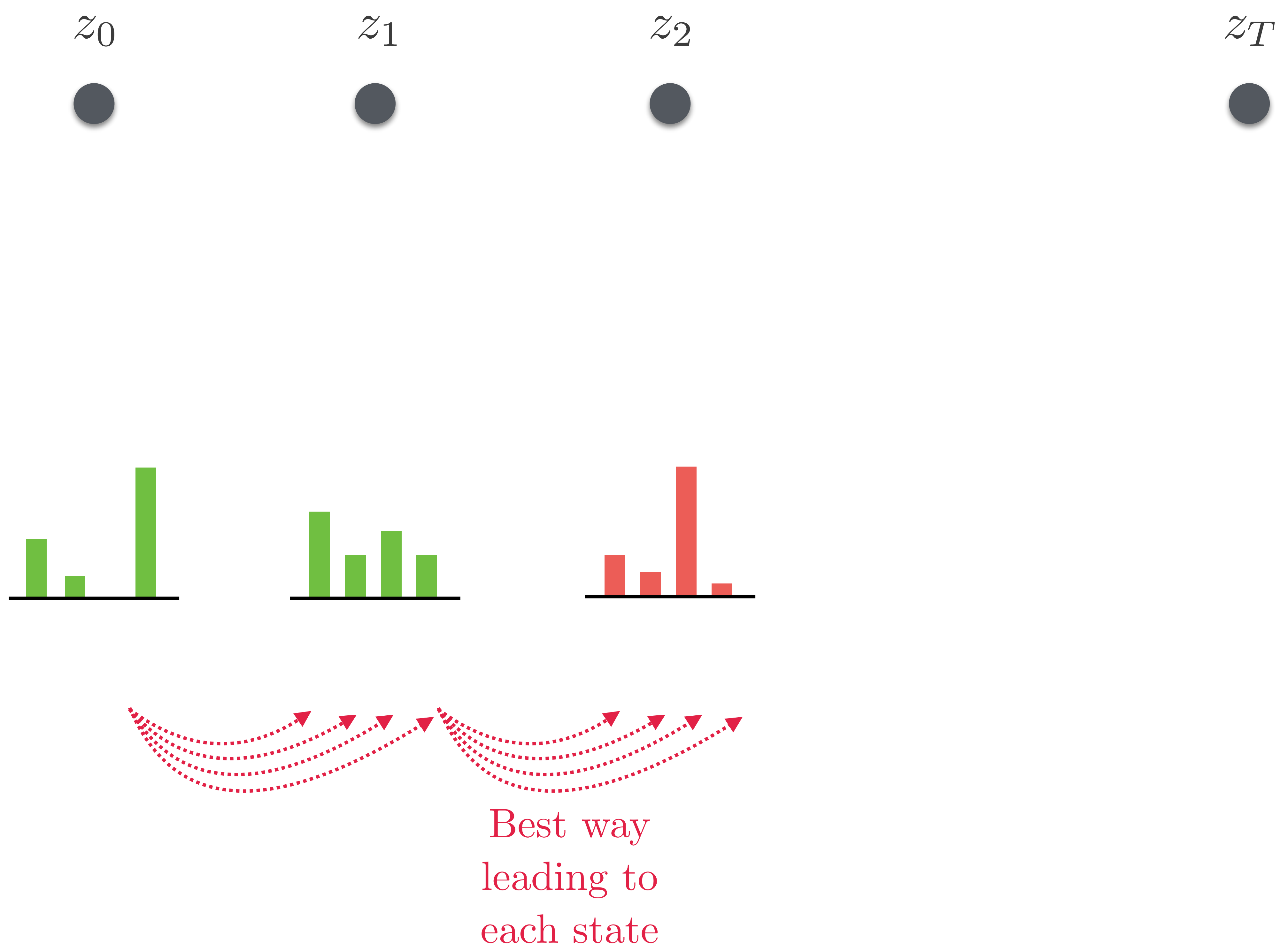
Idea



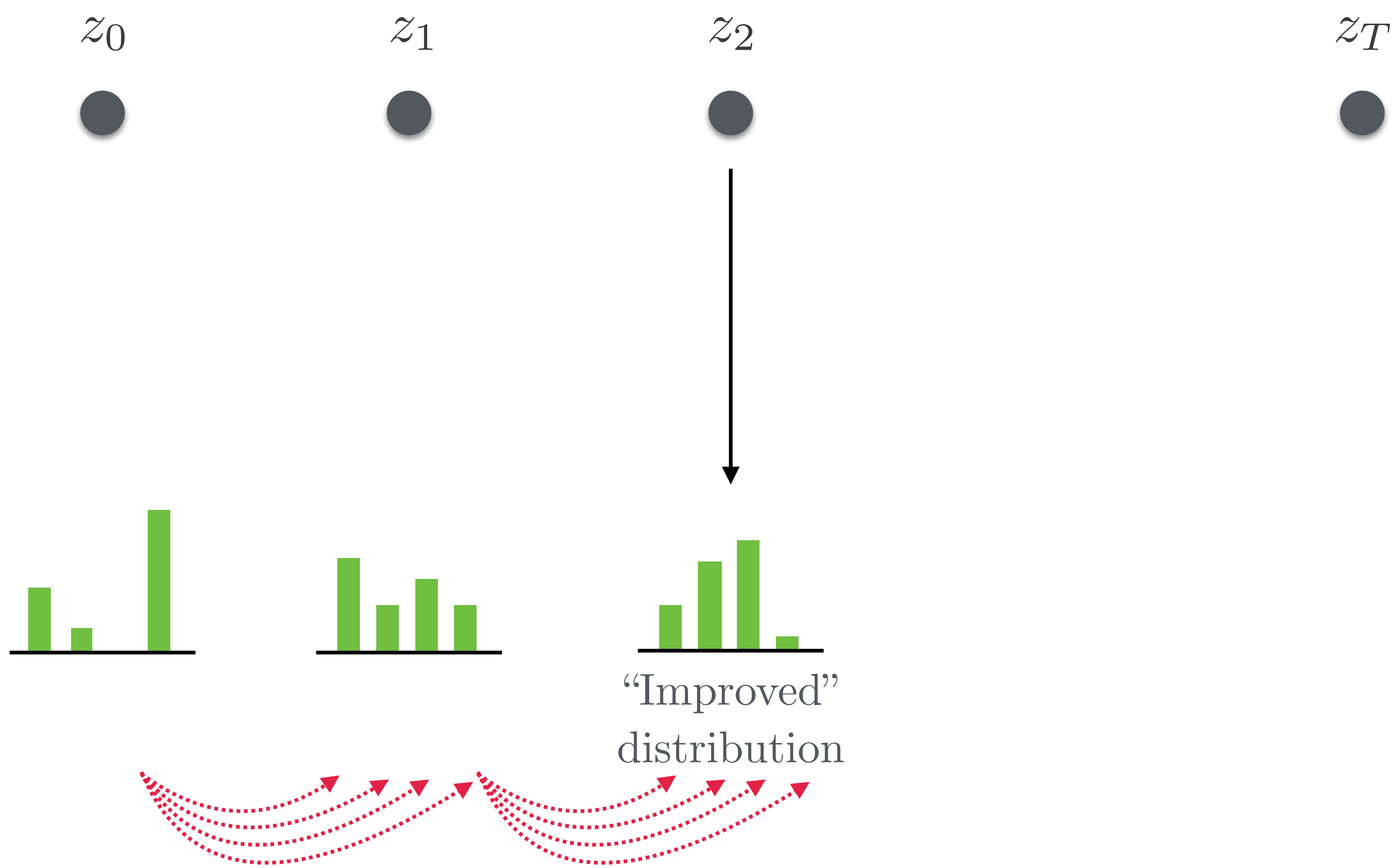
Idea



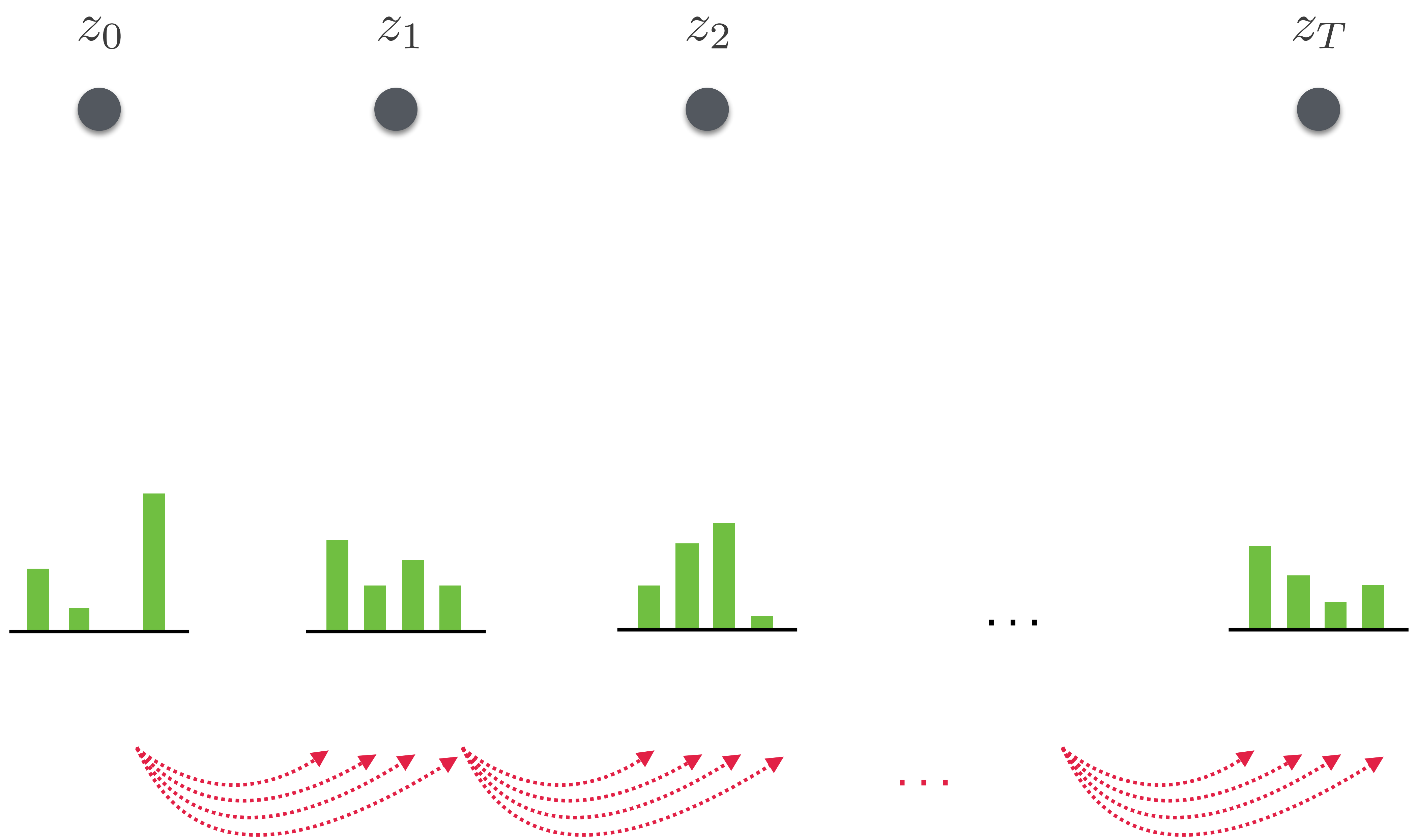
Idea



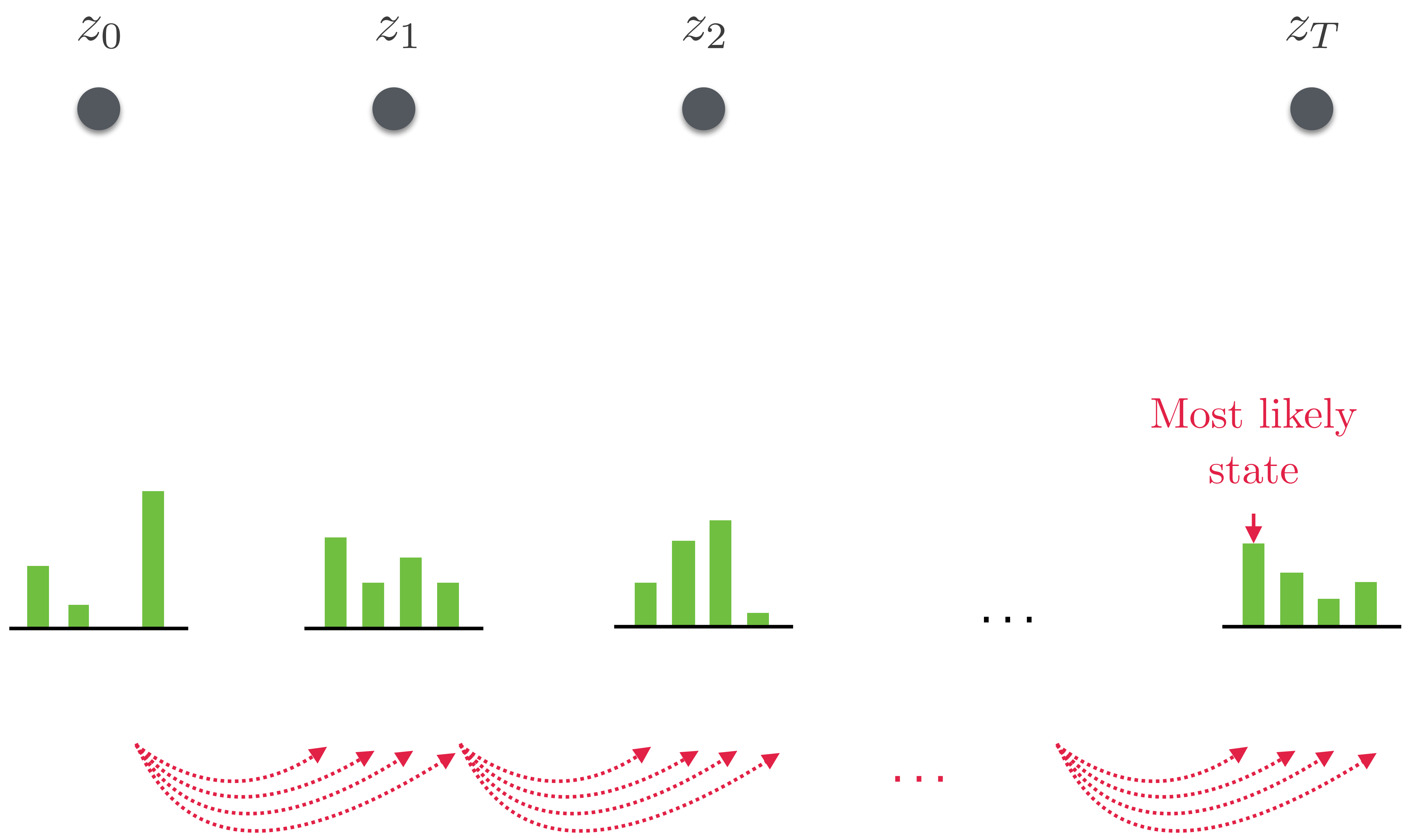
Idea



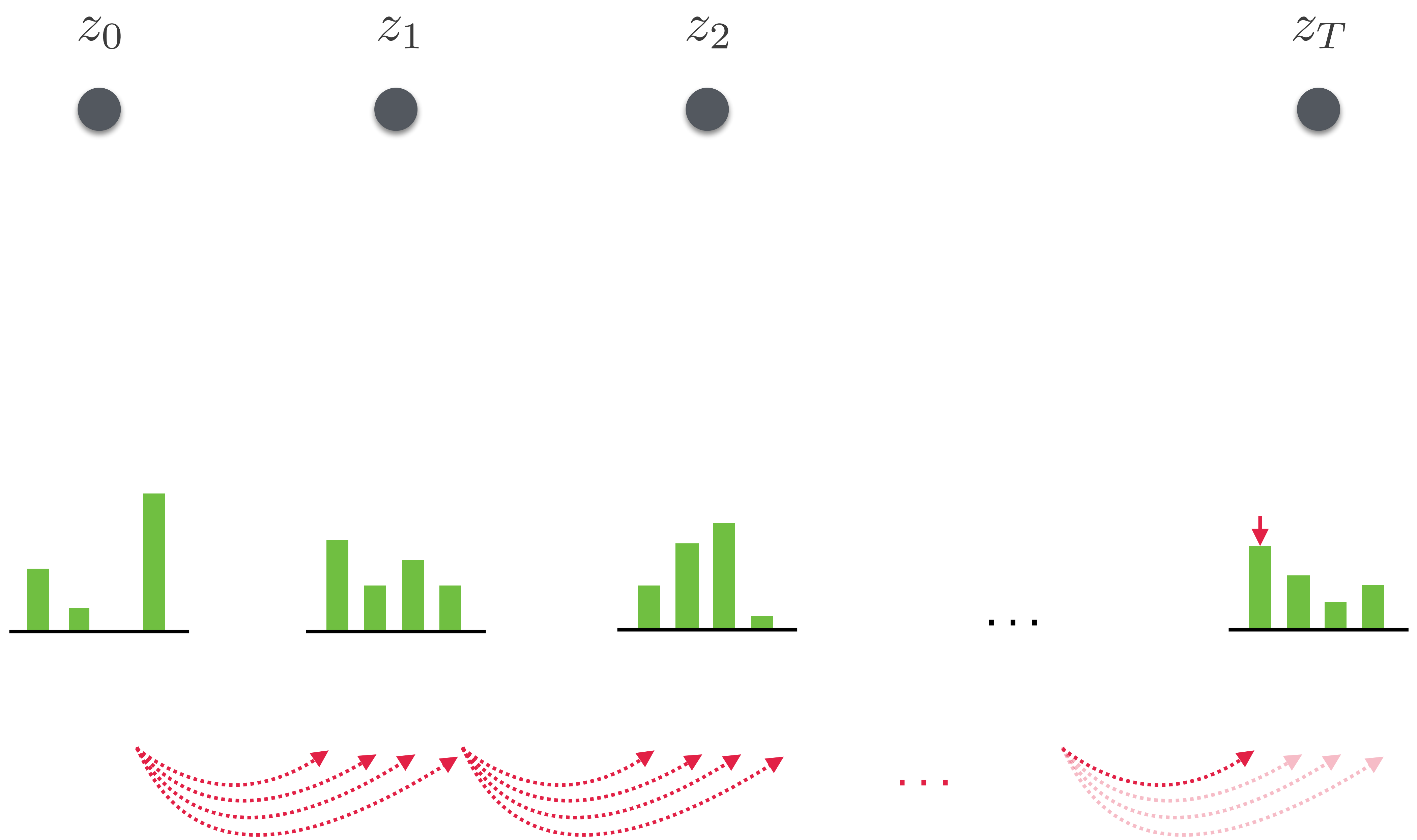
Idea



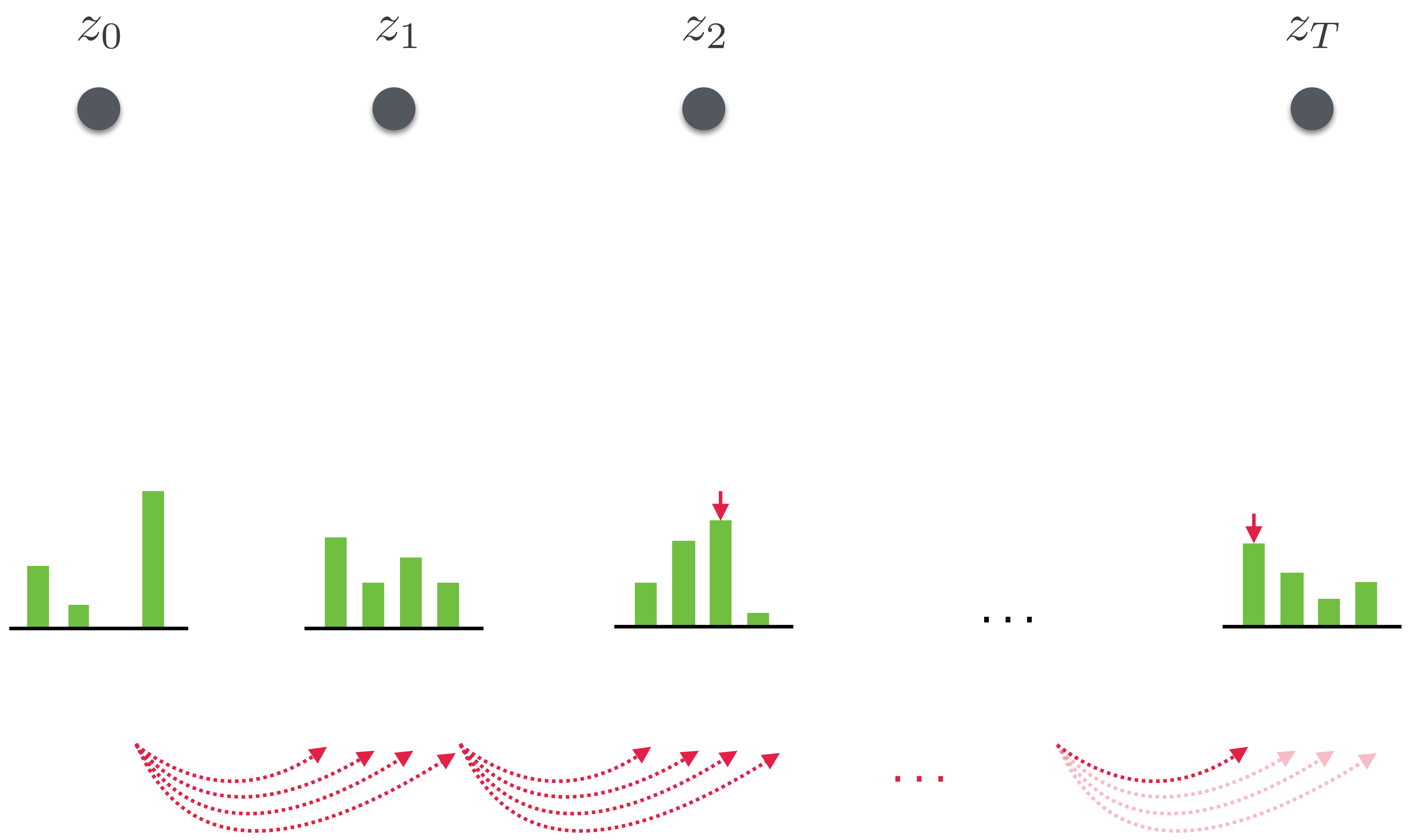
Idea



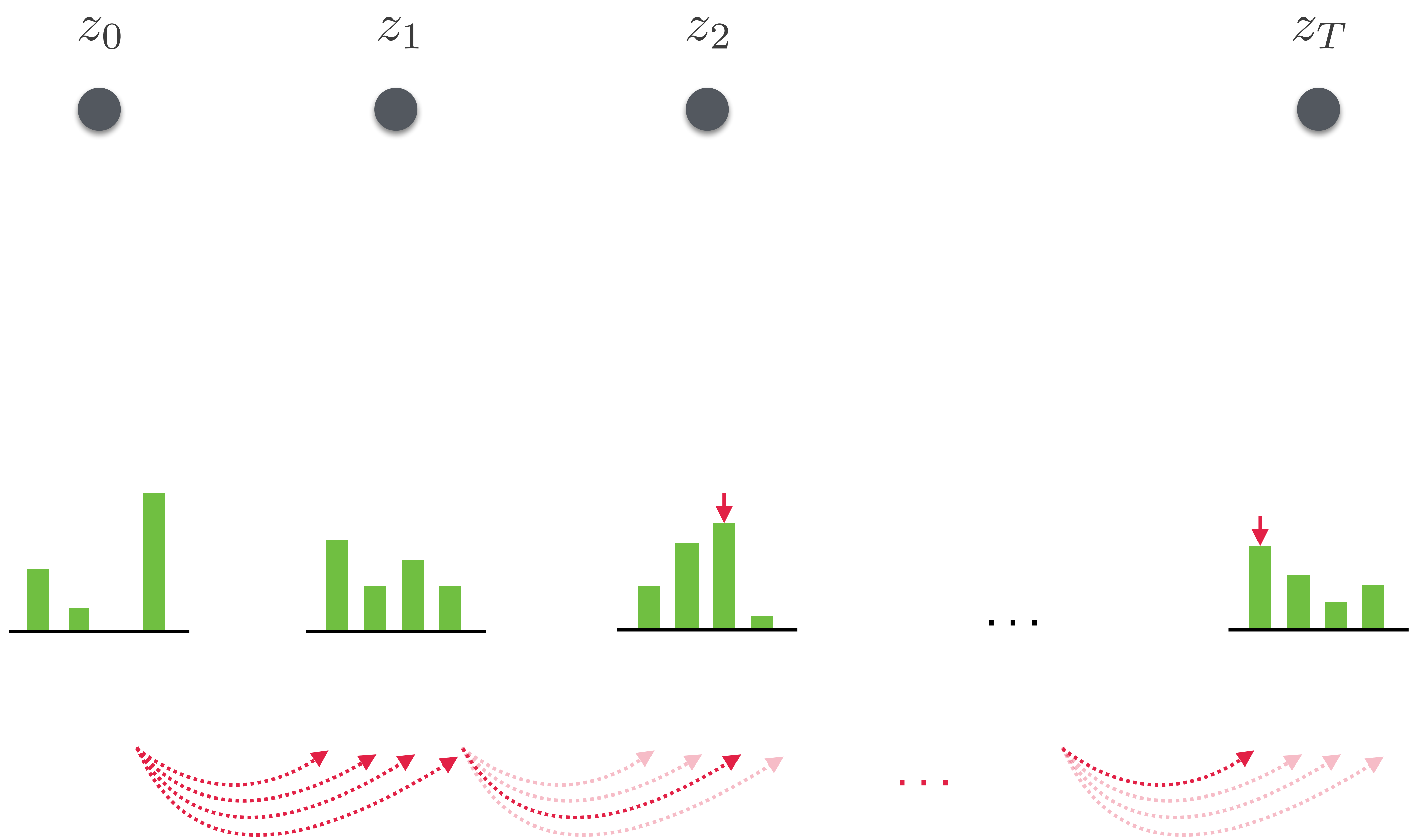
Idea



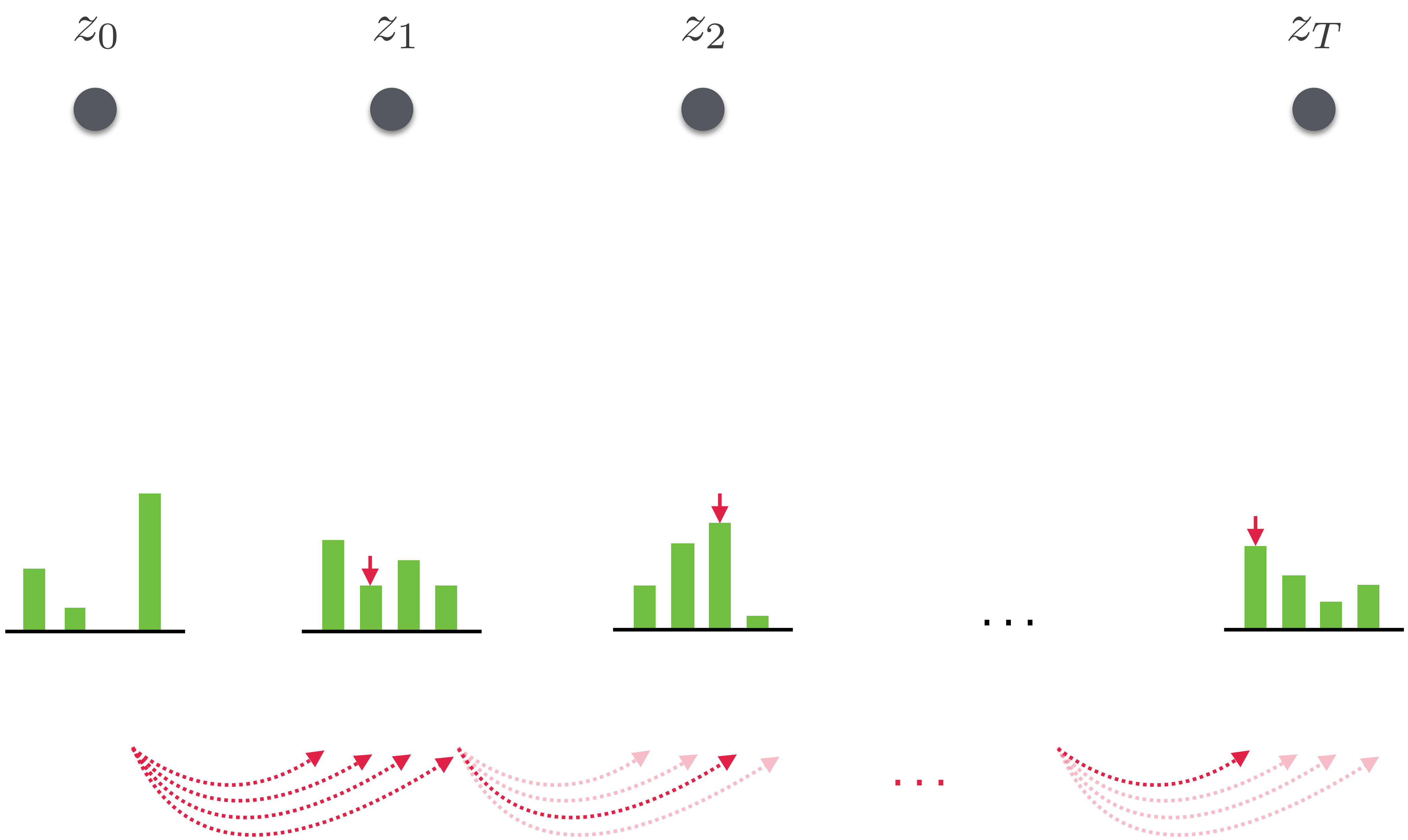
Idea



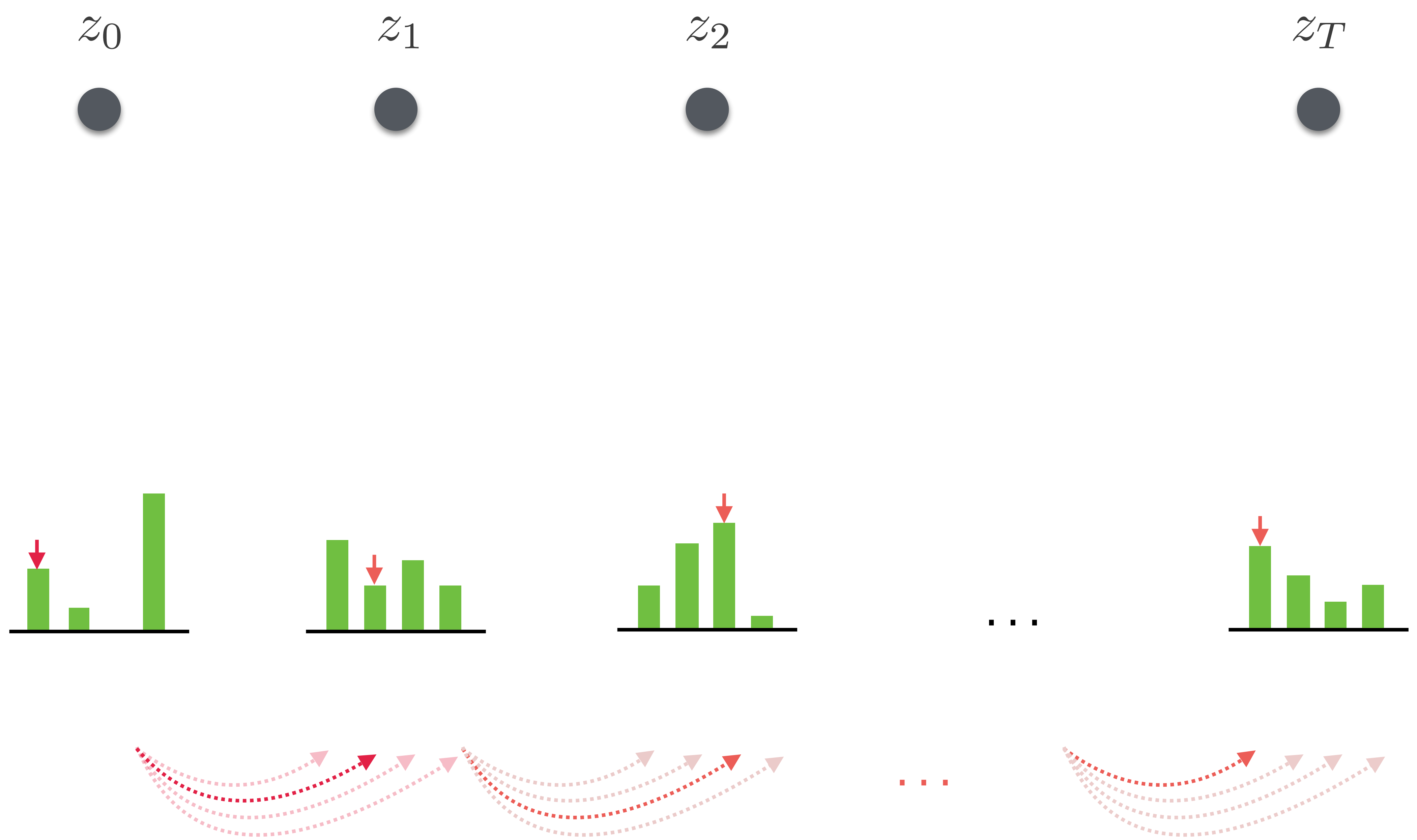
Idea



Idea



Idea



Maximizing forward mapping

Maximizing forward mapping

Given a sequence of observations $\boldsymbol{z}_{0:t}$, the maximizing forward mapping $\boldsymbol{m}_t : \mathcal{X} \to \mathbb{R}$ is defined for each t as

$$m_t(x) = \max_{\boldsymbol{x}_{0:t-1}} \mathbb{P}_{\mu_0}\left[\mathrm{x}_t = x, \mathbf{x}_{0:t-1} = \boldsymbol{x}_{0:t-1}, \mathbf{z}_{0:t} = \boldsymbol{z}_{0:t}\right]$$

Maximizing sequence
ending in x

Viterbi algorithm

Require: Observation sequence $\boldsymbol{z}_{0:T}$

1. Initialize $\boldsymbol{m}_0 = \text{diag}(\mathbf{O}(z_0 \mid \cdot)\boldsymbol{\mu}_0^\top$

2. For $\tau = 1, \ldots, T$ do

$$\boldsymbol{m}_t = \text{diag}(\mathbf{O}(z_t \mid \cdot)) \max\{\mathbf{P}^\top \text{diag}(\boldsymbol{m}_{t-1})\}$$

$$\boldsymbol{i}_t = \arg\max\{\mathbf{P}^\top \text{diag}(\boldsymbol{m}_{t-1})\}$$

3. end for

4. Let $x_T^* = \arg\max_{x \in \mathcal{X}} m_T(x)$

5. For $\tau = T-1, \ldots, 0$ do

$$x_t^* = i_{t+1}(x_{t+1}^*)$$

6. end for

return $\boldsymbol{x}_{0:T}^*$

Forward update

Index tracking

Backtrack

TÉCNICO LISBOA

Example: The urn problem

- Suppose that

$$\boldsymbol{\mu}_0 = \begin{bmatrix} 0.125 & 0.375 & 0.375 & 0.125 \end{bmatrix}$$

- We observe the sequence of observations

$$\boldsymbol{z}_{0:2} = \{w, w, b\}$$

- What is the most likely sequence up to time $t = 2$?

TÉCNICO LISBOA

Step 1: Initialize $\boldsymbol{m}_0$

- $\boldsymbol{m}_0 = \mathrm{diag}(\mathbf{O}(z_0 \mid \cdot))\boldsymbol{\mu}_0^\top$

$$\begin{array}{c}\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \\ \begin{array}{cc} z_0 & \end{array}\end{array} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.125 \\ 0.375 \\ 0.375 \\ 0.125 \end{bmatrix} \rightarrow \begin{bmatrix} 0.125 \\ 0 \\ 0.375 \\ 0 \end{bmatrix}$$

Step 2: Compute $\boldsymbol{m}_1$

- $\boldsymbol{m}_1 = \operatorname{diag}(\mathbf{O}(z_1 \mid \cdot)) \max\{\boxed{\mathbf{P}^\top \operatorname{diag}(\boldsymbol{m}_0)}\}$

$$
\begin{bmatrix}
0.2 & 0.2 & 0.05 & 0.05 \\
0.6 & 0.6 & 0.15 & 0.15 \\
0.15 & 0.15 & 0.6 & 0.6 \\
0.05 & 0.05 & 0.2 & 0.2
\end{bmatrix}
\begin{bmatrix}
0.125 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0.375 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
\rightarrow
\begin{bmatrix}
0.025 & 0 & 0.019 & 0 \\
0.075 & 0 & 0.056 & 0 \\
0.019 & 0 & 0.225 & 0 \\
0.006 & 0 & 0.075 & 0
\end{bmatrix}
$$

Step 2: Compute m_1

- $\boldsymbol{m}_1 = \mathrm{diag}(\mathbf{O}(z_1 \mid \cdot))\boxed{\max\{\mathbf{P}^\top \mathrm{diag}(\boldsymbol{m}_0)\}}$

$$\begin{bmatrix} \boxed{0.025} & 0 & 0.019 & 0 \\ \boxed{0.075} & 0 & 0.056 & 0 \\ 0.019 & 0 & \boxed{0.225} & 0 \\ 0.006 & 0 & \boxed{0.075} & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0.025 \\ 0.075 \\ 0.225 \\ 0.075 \end{bmatrix}$$

Step 2: Compute $\boldsymbol{m}_1$

- $\boldsymbol{m}_1 = \boxed{\mathrm{diag}(\mathbf{O}(z_1 \mid \cdot)}\max\{\mathbf{P}^\top \mathrm{diag}(\boldsymbol{m}_0)\}$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.025 \\ 0.075 \\ 0.225 \\ 0.075 \end{bmatrix} \longrightarrow \begin{bmatrix} 0.025 \\ 0 \\ 0.225 \\ 0 \end{bmatrix}$$

TÉCNICO LISBOA

Step 3: Compute i_1

- $i_1 = \arg\max\{\mathbf{P}^\top \mathrm{diag}(\boldsymbol{m}_0)\}$

$$\begin{bmatrix} \boxed{0.025} & 0 & 0.019 & 0 \\ \boxed{0.075} & 0 & 0.056 & 0 \\ 0.019 & 0 & \boxed{0.225} & 0 \\ 0.006 & 0 & \boxed{0.075} & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 \\ 1 \\ 3 \\ 3 \end{bmatrix}$$

Step 4: Compute m_2

- $\boldsymbol{m}_2 = \mathrm{diag}(\mathbf{O}(z_2 \mid \cdot) \max\{\mathbf{P}^\top \mathrm{diag}(\boldsymbol{m}_1)\}$

$$\begin{bmatrix} 0.2 & 0.2 & 0.05 & 0.05 \\ 0.6 & 0.6 & 0.15 & 0.15 \\ 0.15 & 0.15 & 0.6 & 0.6 \\ 0.05 & 0.05 & 0.2 & 0.2 \end{bmatrix} \begin{bmatrix} 0.025 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0.225 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0.005 & 0 & 0.011 & 0 \\ 0.015 & 0 & 0.034 & 0 \\ 0.004 & 0 & 0.135 & 0 \\ 0.001 & 0 & 0.045 & 0 \end{bmatrix}$$

Step 4: Compute m_2

- $\boldsymbol{m}_2 = \mathrm{diag}(\mathbf{O}(z_2 \mid \cdot) \max\{\mathbf{P}^\top \mathrm{diag}(\boldsymbol{m}_1)\}$

$$\begin{bmatrix} 0.005 & 0 & \boxed{0.011} & 0 \\ 0.015 & 0 & \boxed{0.034} & 0 \\ 0.004 & 0 & \boxed{0.135} & 0 \\ 0.001 & 0 & \boxed{0.045} & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0.011 \\ 0.034 \\ 0.135 \\ 0.045 \end{bmatrix}$$

TÉCNICO LISBOA

Step 4: Compute $\boldsymbol{m}_2$

- $\boldsymbol{m}_2 = \text{diag}(\mathbf{O}(z_2 \mid \cdot) \max\{\mathbf{P}^\top \text{diag}(\boldsymbol{m}_1)\}$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.011 \\ 0.034 \\ 0.135 \\ 0.045 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 \\ 0.034 \\ 0 \\ 0.045 \end{bmatrix}$$

Step 5: Compute i_2

- $\boldsymbol{i}_2 = \arg\max\{\mathbf{P}^\top \mathrm{diag}(\boldsymbol{m}_1)\}$

$$\begin{bmatrix} 0.005 & 0 & 0.011 & 0 \\ 0.015 & 0 & 0.034 & 0 \\ 0.004 & 0 & 0.135 & 0 \\ 0.001 & 0 & 0.045 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 3 \\ 3 \\ 3 \\ 3 \end{bmatrix}$$

Step 6: Maximize m_2

- $x_2^* = \arg\max_{x \in \mathcal{X}} m_2(x)$

$$\boldsymbol{m}_2$$

$$\begin{bmatrix} 0 \\ 0.034 \\ 0 \\ \boxed{0.045} \end{bmatrix}$$

$$x_2^* = 4$$

Step 7: Backtrack

- $x_t^* = i_{t+1}(x_{t+1}^*)$

$$
\begin{array}{ccc}
\boldsymbol{i}_1 & \boldsymbol{i}_2 & \boldsymbol{m}_2 \\
\begin{bmatrix} 1 \\ 1 \\ \boxed{3} \\ 3 \end{bmatrix} &
\begin{bmatrix} 3 \\ 3 \\ 3 \\ \boxed{3} \end{bmatrix} &
\begin{bmatrix} 0 \\ 0.034 \\ 0 \\ \boxed{0.045} \end{bmatrix} \\
x_1^* = 3 & x_1^* = 3 & x_2^* = 4
\end{array}
$$

Finally...

- Most likely sequence:

$$x_1^* = 3 \qquad\qquad x_1^* = 3 \qquad\qquad x_2^* = 4$$

$$(2,\ w) \qquad\qquad (2,\ w) \qquad\qquad (2,\ b)$$

TÉCNICO LISBOA

Estimation

- **Filtering:**
  - Given a sequence of observations, estimate the final state
- **Smoothing:**
  - Given a sequence of observations, estimate the sequence of states
- **Prediction:**
  - Given a sequence of observations, predict future states

TÉCNICO LISBOA

Prediction

- We are given a sequence of observations $\boldsymbol{z}_{0:T}$

- We want to estimate, for $t > T$

$$\mathbb{P}_{\mu_0}\left[\mathrm{x}_t = x \mid \mathbf{z}_{0:T} = \boldsymbol{z}_{0:T}\right]$$

where μ_0 is the initial distribution, i.e.,

$$\mu_0(x) = \mathbb{P}\left[\mathrm{x}_0 = x\right]$$

Prediction

- Easy:

  - We compute $\boldsymbol{\mu}_{T|0:T}$ using the forward algorithm

  - We use the Markov property:

$$\boldsymbol{\mu}_{T+1|0:T} = \boldsymbol{\mu}_{T|0:T}\mathbf{P}$$

TÉCNICO LISBOA

Making decisions

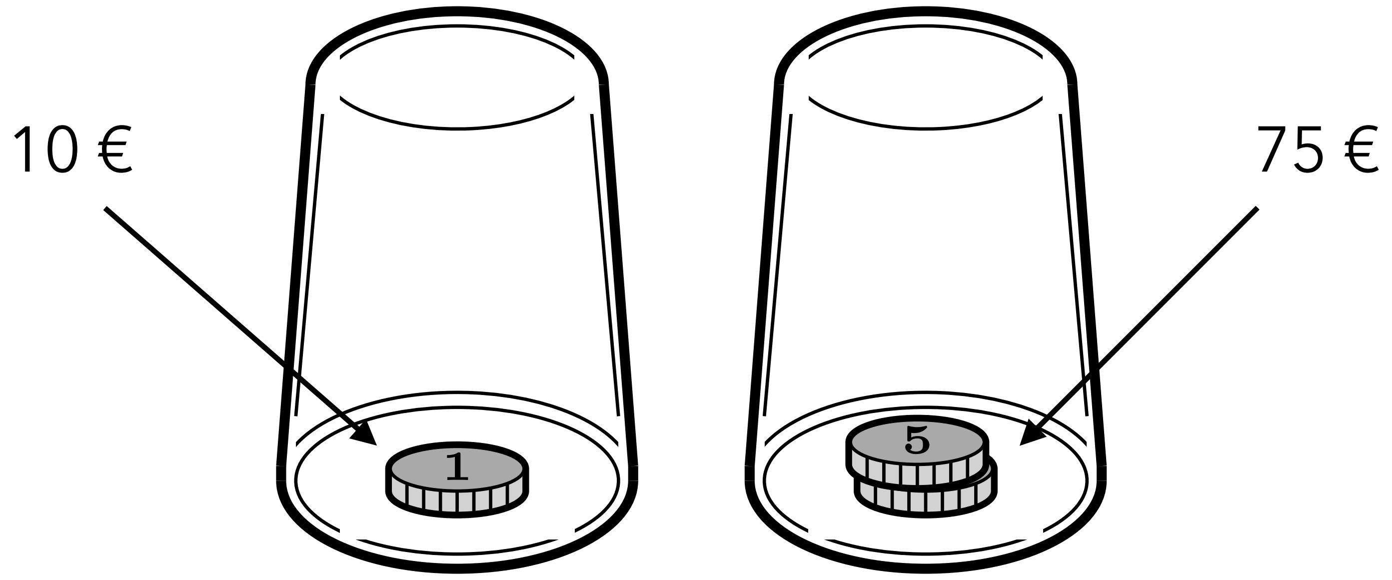
Decision I

- You are a participant in a contest
  - You are currently winning **50 EUR**
  - You are at the last stage of the game

TÉCNICO LISBOA

Decision I

- You are offered a choice:
  - Keep your prize
  - Choose one of the cups



Which one will you choose?

TÉCNICO LISBOA

Decision II

- You are a participant in a contest
  - You are currently winning a **50EUR-valued meal** in a steak house
  - You are at the last stage of the game

Decision I

- You are offered a choice:
  - Keep your price
  - Choose one of the cups

Which one will you choose?

Decisions and preferences

- The prizes in both situations have the same monetary value
- In the first situation, the more valuable prize will be selected
- In the second situation, different choices are possible...

Decisions and preferences

- If you are allergic to cheese, you may prefer to keep your current prize
- If you love pizza, you may prefer to go for the pizza house meal

Is this “irrational” in some way?

Decisions and preferences

- Decisions often involve factors other than value:
  - Many decisions are based on personal **preferences**

Can we program “preferences” into a computer?

Can we treat decision making algorithmically?

Preferences

What is a strict preference?

- A strict preference is a **binary relation** between **outcomes**
  - When outcome x is **preferred** to y, we write $x > y$
  - If an individual prefers outcome x to y, that means that it would be willing to pay some "fair amount" to have x instead of y

TÉCNICO LISBOA

What is a strict preference?

- Is any binary relation a valid preference?
  - **No!**

Example 1.

Suppose that, for some decision maker, $x \succ y$ and $y \succ x$.

We could charge the agent some amount a to swap x for y. And then swap some amount b to swap y for x. And then charge a to swap x for y again, and so on...

Not reasonable

What is a strict preference?

- Is any binary relation a valid preference?
  - **No!**

Preferences must not be **symmetric**

Example 1.

Suppose that, for some decision maker, $x > y$ and $y > x$.

We could charge the agent some amount a to swap x for y. And then swap some amount b to swap y for x. And then charge a to swap x for y again, and so on...

What is a strict preference?

- Is any binary relation a valid preference?
  - **No!**

Example 2.

Suppose that, for some decision maker, $x \succ y$, $y \succ z$ and $z \succ x$.

We could charge the agent some amount a to swap z for x. And then pay some amount b to swap y for z. And then charge c to swap x for y, and so on...

Not reasonable

TÉCNICO LISBOA

What is a strict preference?

- Is any binary relation a valid preference?
  - **No!**

Preferences must be **negative transitive**

Example 2.

Suppose that, for some decision maker, $x > y$, $y > z$ and $z > x$.

We could charge the agent some amount a to swap z for x. And then pay some amount b to swap y for z. And then charge c to swap x for y, and so on...

TÉCNICO LISBOA

What is a strict preference?

- A **strict preference** is a **binary relation** between **outcomes** such that
  - It is anti-symmetric:

$$x \succ y \Rightarrow y \not\succ x$$

  - It is negative transitive:

$$x \not\succ y \text{ and } y \not\succ z \text{ then } x \not\succ z$$

TÉCNICO LISBOA

What is a strict preference?

- A **strict preference** is a **binary relation** between **outcomes** such that
  - It is anti-symmetric:

$$x \succ y \Rightarrow y \not\succ x$$

  - It is negative transitive:

$$x \succ y \text{ then either } z \succ y \text{ or } x \succ z$$

TÉCNICO LISBOA

Related relations

- If "$\succ$" is a strict preference on some set of outcomes $\mathcal{X}$...
  - ... we write "$x \succ y$" if outcome $\boldsymbol{x}$ **is preferred to** $\boldsymbol{y}$ ("x is better than y")
  - ... we write "$x \prec y$" if outcome $\boldsymbol{y}$ **is preferred to** $\boldsymbol{x}$ (or "x is worse than y")
  - ... if $x \nsucc y$ and $x \nprec y$ (x is neither better nor worse than y), we say that **the two outcomes are indifferent**, and write $x \sim y$

Related relations

- If "$\succ$" is a strict preference on some set of outcomes $\mathcal{X}$...
  - ... if "$x \succ y$" or "$x \sim y$" we can write "$x \succcurlyeq y$" (**$\boldsymbol{x}$ is not worse than $\boldsymbol{y}$**).
  - ... if "$x \prec y$" or "$x \sim y$", we can write "$x \preccurlyeq y$" (**$\boldsymbol{x}$ is not better than $\boldsymbol{y}$**).

Properties of preferences

- Given any two outcomes, exactly one of the following holds:
  - $x \succ y$
  - $x \sim y$
  - $x \prec y$

- Equivalently, given any two outcomes, either $x \succcurlyeq y$ or $y \succcurlyeq x$

The relation $\succcurlyeq$ is complete

Properties of preferences

- Given any outcomes x, y and z, if $x \succ y$ and $y \succ z$, then $x \succ z$

The relation $\succ$ is transitive

Properties of preferences

- Given any outcomes x, y and z, if $x \succcurlyeq y$ and $y \succcurlyeq z$, then $x \succcurlyeq z$

The relation $\succcurlyeq$ is transitive

TÉCNICO LISBOA

Properties of preferences

- Indifference is
  - Reflexive (i.e., $x \sim x$)
  - Symmetric (i.e., if $x \sim y$, then $y \sim x$)
  - Transitive (i.e., if $x \sim y$ and $y \sim z$, then $x \sim z$)

The relation $\sim$ is an equivalence relation

Properties of preferences

Summarizing:

- $\succ$ is **anti-symmetric**, **transitive** and **negative transitive**
- $\succcurlyeq$ is **complete** and **transitive**
- $\sim$ is **reflexive**, **symmetric** and **transitive** (i.e., is an equivalence relation)

TÉCNICO LISBOA

Rational preference

- A relation is called a **rational preference** if it is **complete** and **transitive**
  - **The relation $\succcurlyeq$ is complete and transitive and hence rational**

TÉCNICO LISBOA

Utility

Computational considerations

- Computationally, preferences are cumbersome to maintain
- Preferences express an **ordering** between outcomes

Order preserving function

Existence

- Does an order preserving function exist?
  - **Yes!**
  - ... if the preference is rational

TÉCNICO LISBOA

Existence

Theorem

Let $\mathcal{X}$ be a set of possible outcomes, and $\succeq$ a rational preference on $\mathcal{X}$. Then, there is a function $u : \mathcal{X} \to \mathbb{R}$ such that $u(x) \geq u(y)$ if and only if $x \succeq y$ for all $x, y \in \mathcal{X}$.

u is called a *utility* function

Making decisions

- We can use utility functions in computational decision-making
  - Given a set of alternatives/actions $\mathcal{A}$
  - $X(a)$ is the outcome associated with action $a \in \mathcal{A}$
  - The **value** of action a is

$$Q(a) \stackrel{\text{def}}{=} u(X(a))$$

Utility of associated outcome

TÉCNICO LISBOA

Making decisions

- We can use utility functions in computational decision-making
  - Given a set of alternatives/actions $\mathcal{A}$
  - $X(a)$ is the outcome associated with action $a \in \mathcal{A}$
  - The **value** of action a is

$$Q(a) = \sum_{x \in \mathcal{X}} u(x)\mathbb{I}[x = X(a)]$$

1 if condition is true
0 otherwise

Making decisions

- Select actions with maximum value:

$$\operatorname*{argmax}_{a \in \mathcal{A}} Q(a)$$

TÉCNICO LISBOA

Uncertainty

Handling uncertainty

- What if action outcomes are uncertain?
    - We write $P(x \mid a)$ to denote the probability of outcome x when action a was selected:

$$P(x \mid a) = \mathbb{P}[\mathrm{x} = x \mid \mathrm{a} = a]$$

    - The **value** of an action was:

$$Q(a) = \sum_{x \in \mathcal{X}} u(x)\mathbb{I}[x = X(a)]$$

Handling uncertainty

- What if action outcomes are uncertain?
  - We write $P(x \mid a)$ to denote the probability of outcome x when action a was selected:

$$P(x \mid a) = \mathbb{P}[\mathrm{x} = x \mid \mathrm{a} = a]$$

  - The **expected value** of an action is now:

$$Q(a) = \sum_{x \in \mathcal{X}} u(x) P(x \mid a)$$

$$= \mathbb{E}\left[u(\mathrm{x}) \mid a\right]$$

TÉCNICO LISBOA

Formulating our problem

- One such decision problem is described as a tuple $(\mathcal{X}, \mathcal{A}, P, u)$
  - $\mathcal{X}$ is the set of possible outcomes
  - $\mathcal{A}$ is the set of available actions
  - For each outcome x and action a, $P(x \mid a)$ is the probability of outcome x when action a is selected
  - For each outcome x, $u(x)$ is the utility of x

Examples

The weather example

The weather example

- You must decide whether to take an umbrella before leaving home
  - Carrying an umbrella is inconvenient
  - You don't want to get soaked because of rain

The weather example

- The weather forecast is:
  - Rain with probability 0.3
  - Sun with probability 0.7

TÉCNICO LISBOA

The weather example

- At any moment,
  - You will be outside with a probability 0.5
  - You will be indoors with a probability 0.5

TÉCNICO LISBOA

Outcomes

- What are the possible outcomes?

  (A) Don't carry umbrella; get home dry

  (B) Carry umbrella; get home dry

  (C) Don't carry umbrella; get home soaked

  (D) Carry umbrella; get home soaked

Outcomes

- What are the possible outcomes?

  (A) Don't carry umbrella; get home dry

  (B) Carry umbrella; get home dry

  (C) Don't carry umbrella; get home soaked

  (D) ~~Carry umbrella; get home soaked~~

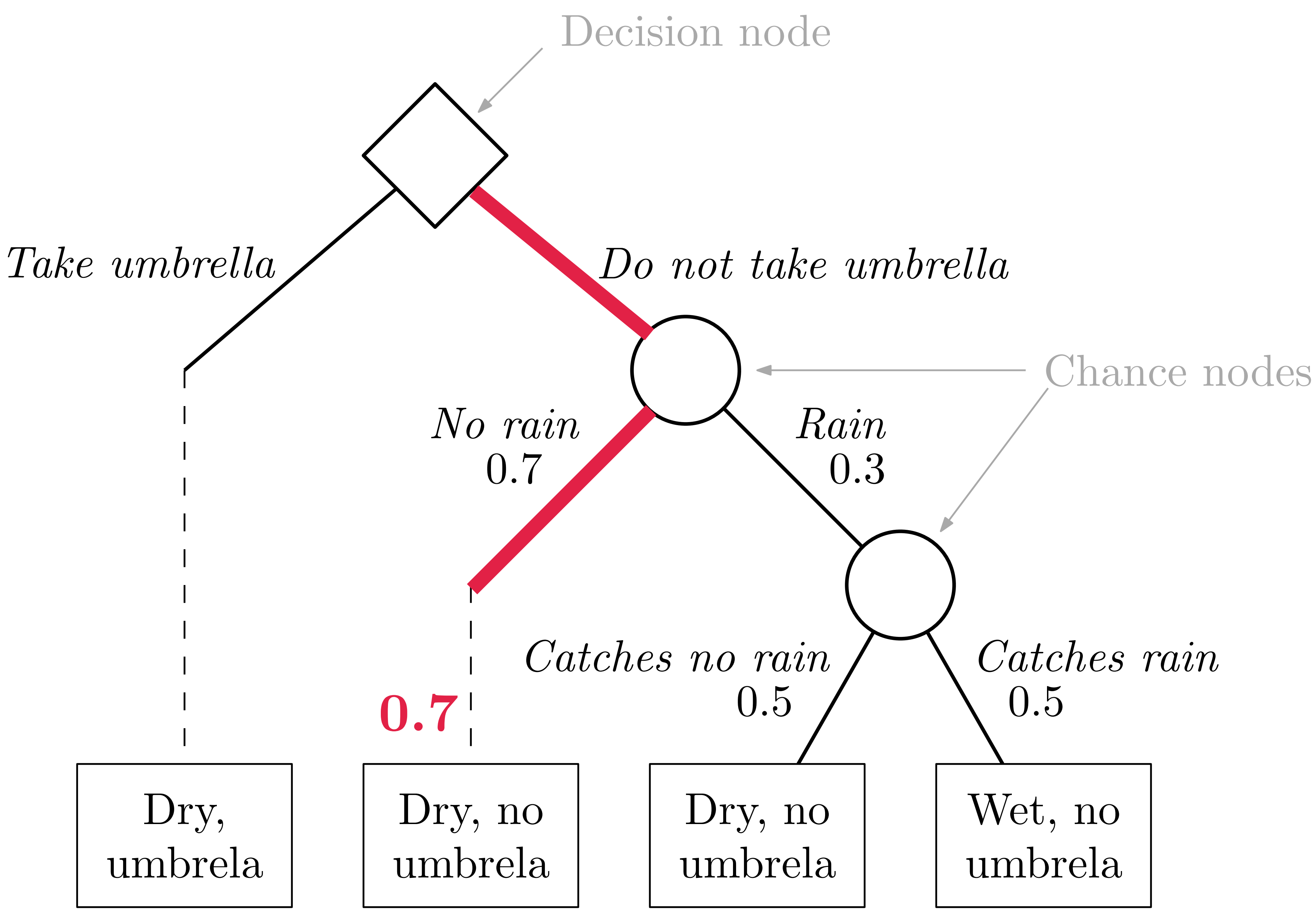
Actions

- What are the possible actions?

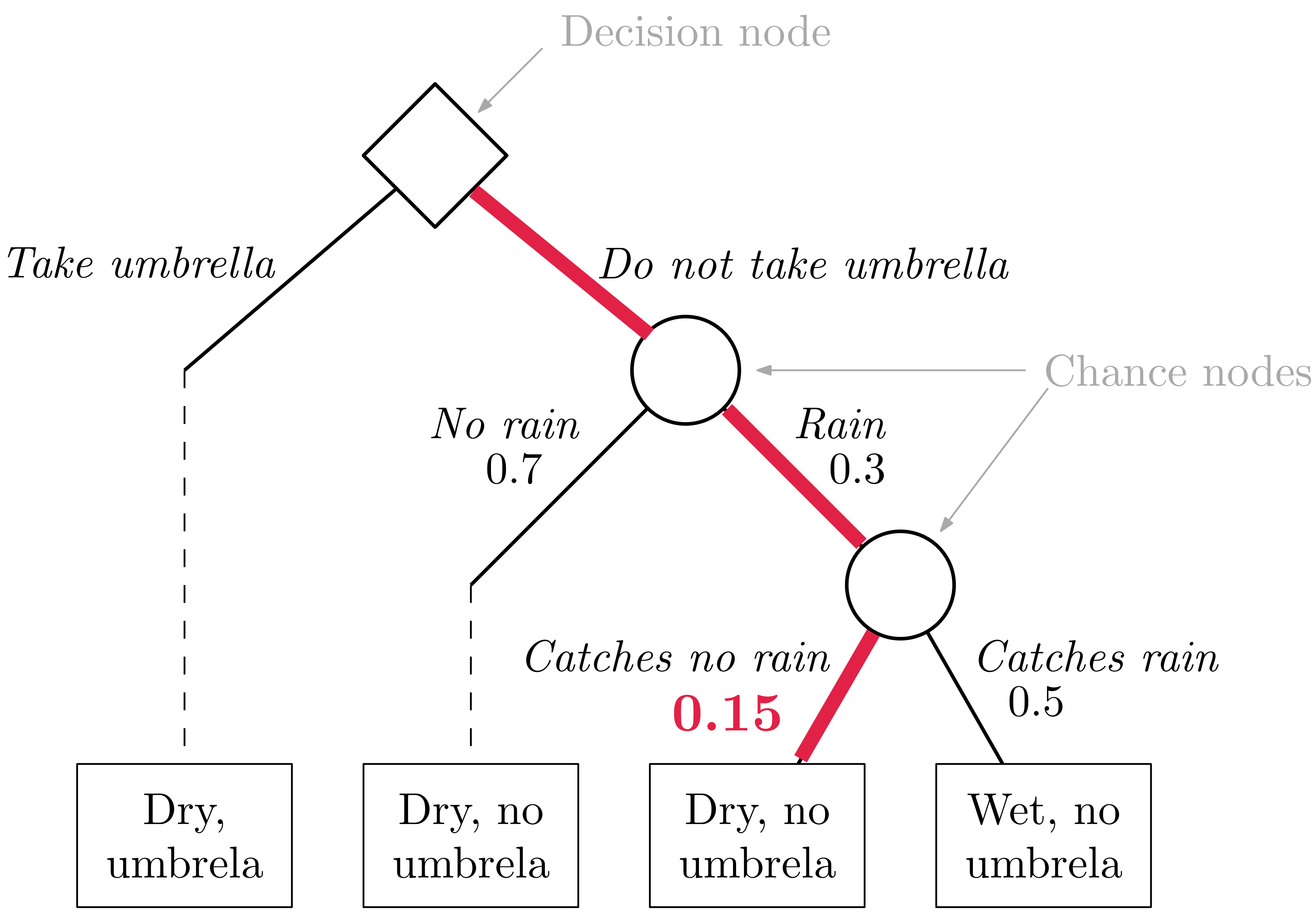
  (A) Take the umbrella

  (B) Don't take the umbrella

Decision tree

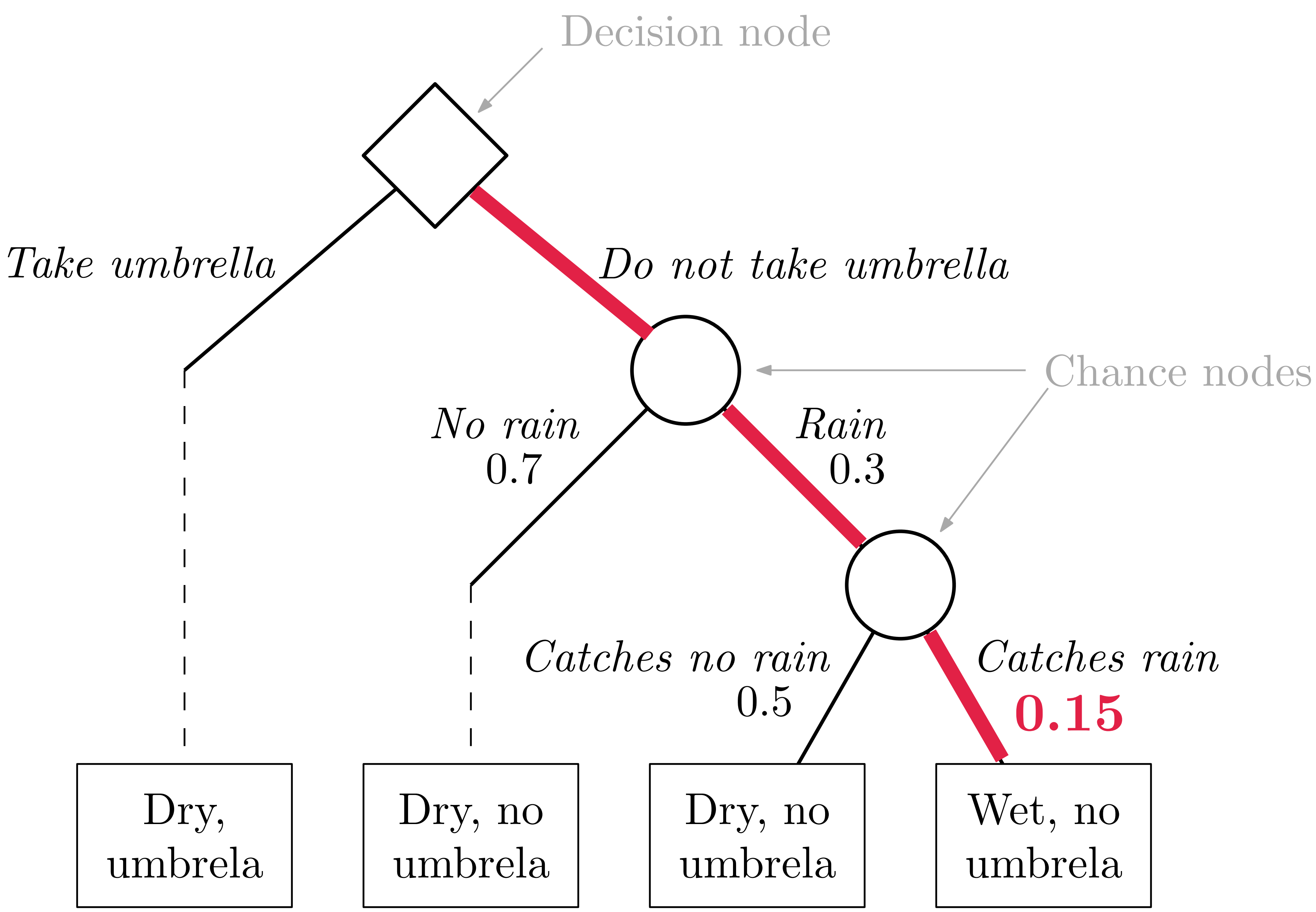


Decision tree



TÉCNICO LISBOA

Decision tree

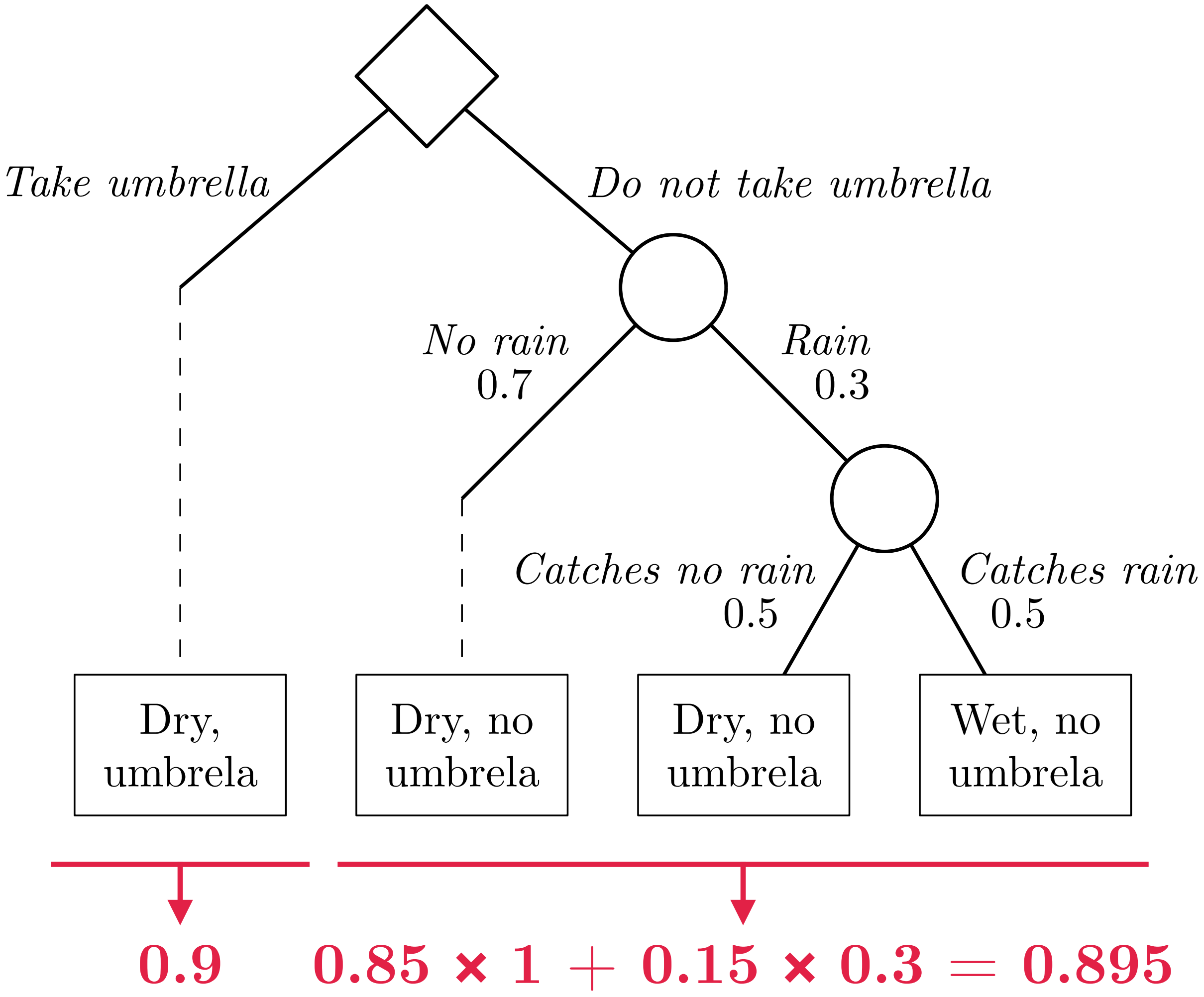


Outcome probabilities

- If $a =$ "Take the umbrella",
  - $P(\text{Carry umbrella, get home dry} \mid a) = 1$
  - $P(x \mid a) = 0$, otherwise
- If $a =$ "Don't take the umbrella",
  - $P(\text{Don't carry umbrella, get home dry} \mid a) = 0.85$
  - $P(\text{Don't carry umbrella, get home soaked} \mid a) = 0.15$
  - $P(x \mid a) = 0$, otherwise

Expected value

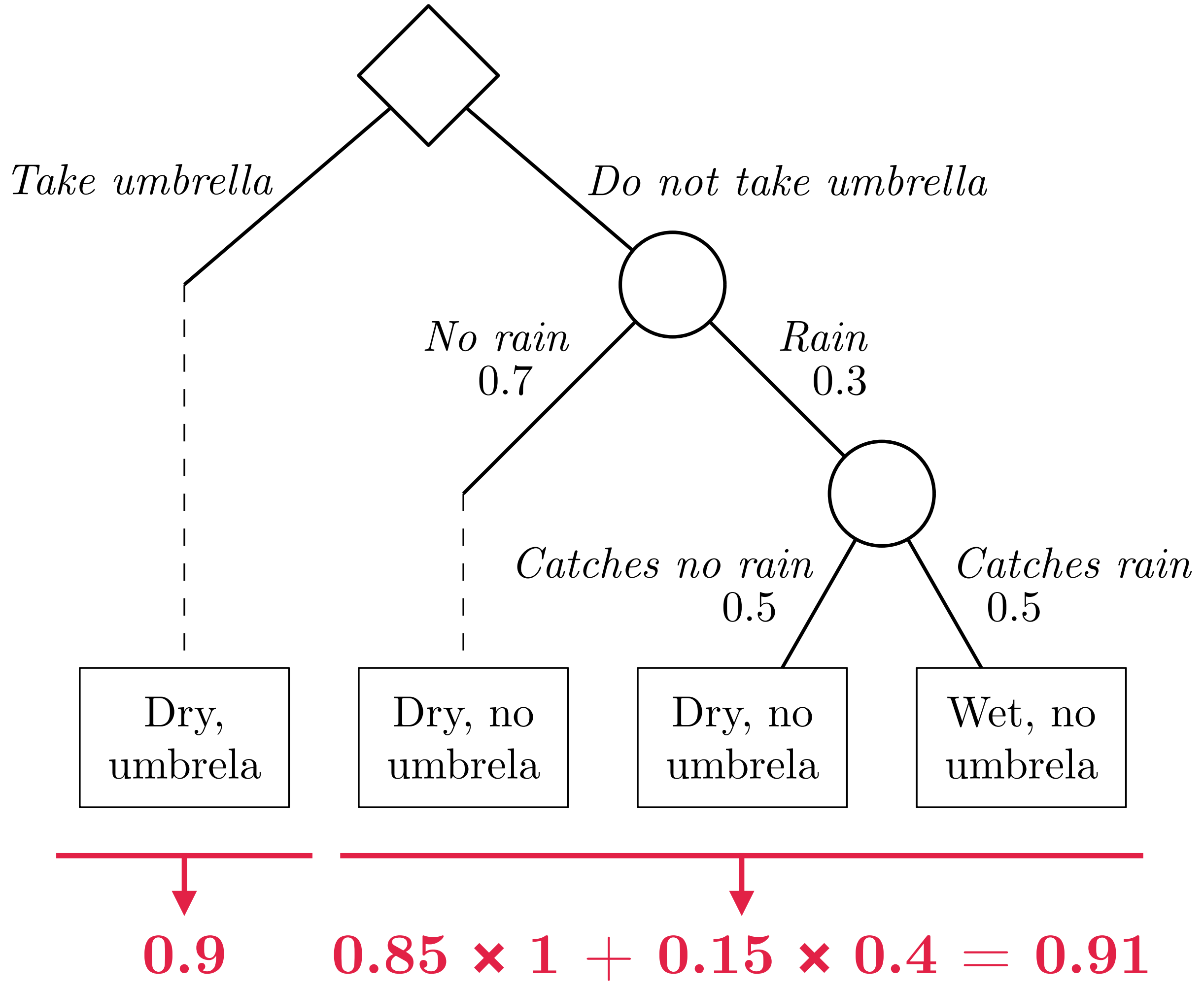
- Multiply leave probabilities by corresponding value
  - $u(\text{No umbrella, dry}) = 1$
  - $u(\text{Umbrella, dry}) = 0.9$
  - $u(\text{No umbrella, wet}) = 0.3$
  - $u(\text{Umbrella, wet}) = 0$



- Best action: Take umbrella

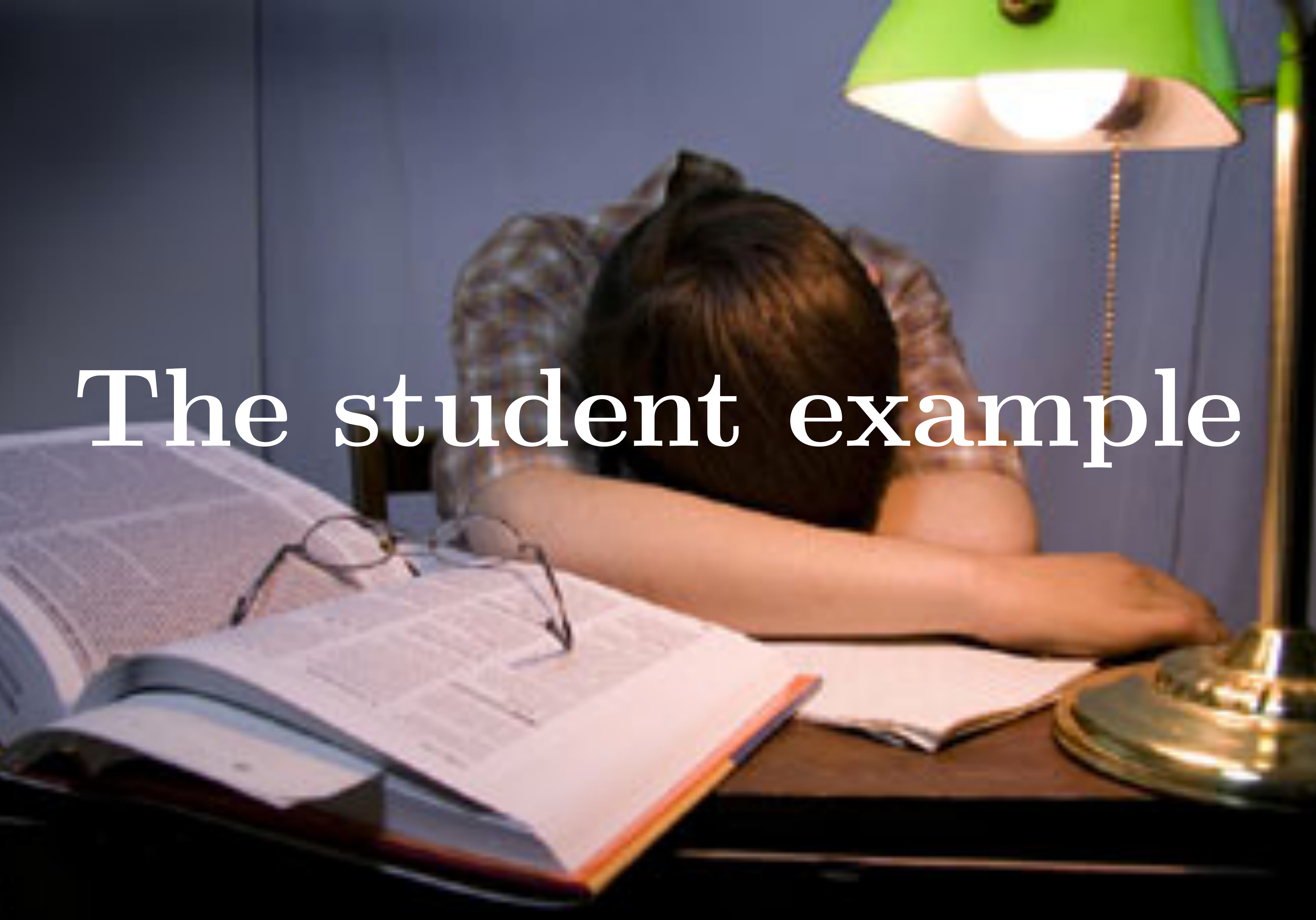
Expected value

- Multiply leave probabilities by corresponding value
  - $u(\text{No umbrella, dry}) = 1$
  - $u(\text{Umbrella, dry}) = 0.9$
  - $u(\text{No umbrella, wet}) = 0.4$
  - $u(\text{Umbrella, wet}) = 0$



- Best action: Don't take umbrella

The student example



The student example

- A freshman finished the final project for a course
- On her way to submit the project report, she realizes that half the pages are missing!

Alternatives

- She has two alternatives:

  (A) Return home and print the remaining pages

  (B) Print the remaining pages at the University

The student example

- If she returns home...
  - There is a 0.6 probability that she'll arrive late at the University, due to traffic

The student example

- If she prints in the University...
  - There is a 0.3 probability that she can't find a printer in time (she'll submit an incomplete report)
  - There is a 0.5 chance that the printer is busy (she'll submit the report late)

TÉCNICO LISBOA

The student example

- If she submits the report late, she'll lose 2 points
- If she submits an incomplete report, she'll lose 3 points

Outcomes

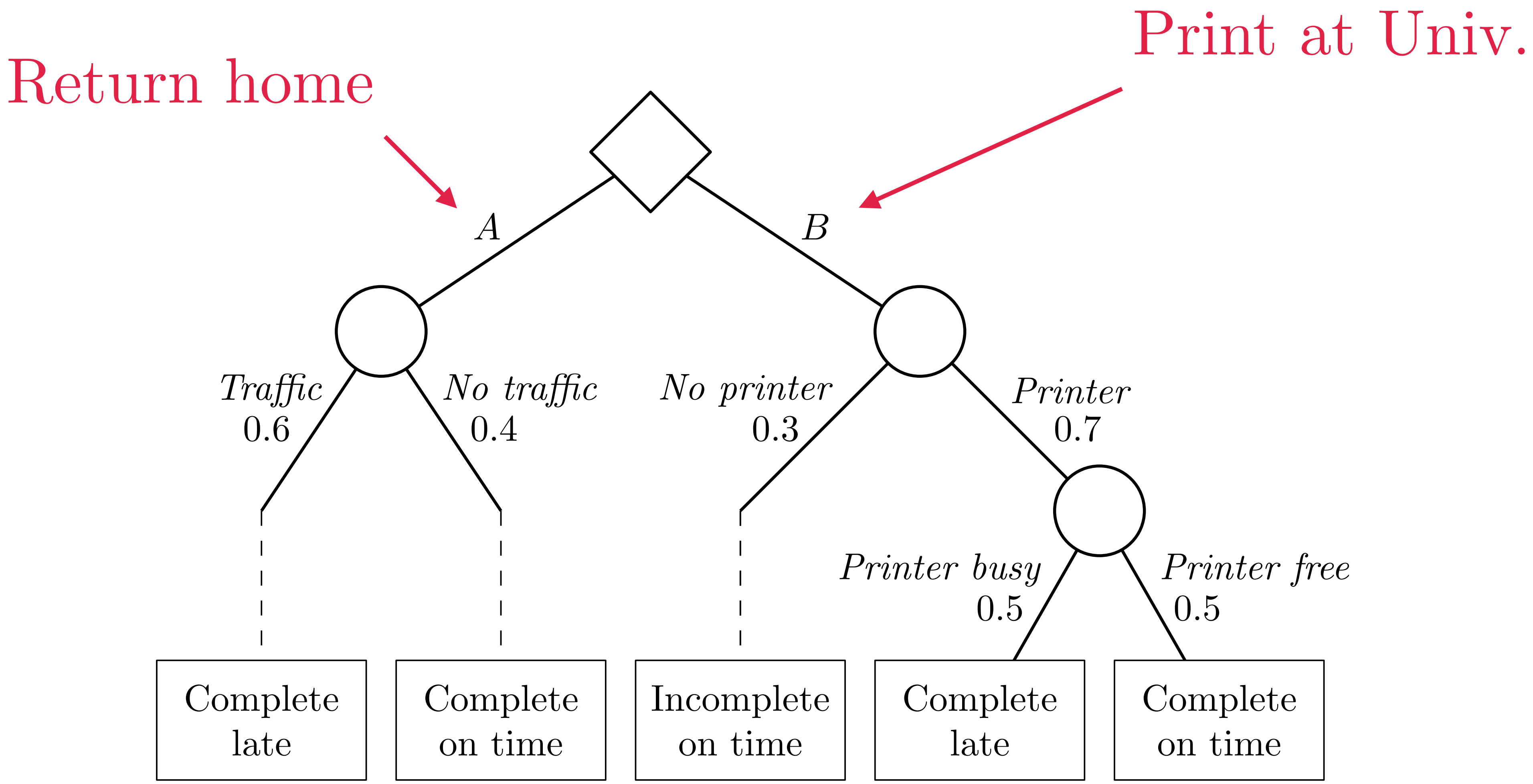
- What are the possible outcomes?

  (A) Report is complete and on time (CT, utility of 0)

  (B) Report is complete but late (CL, utility of -2)

  (C) Report is incomplete (IT, utility of -3)

Decision tree



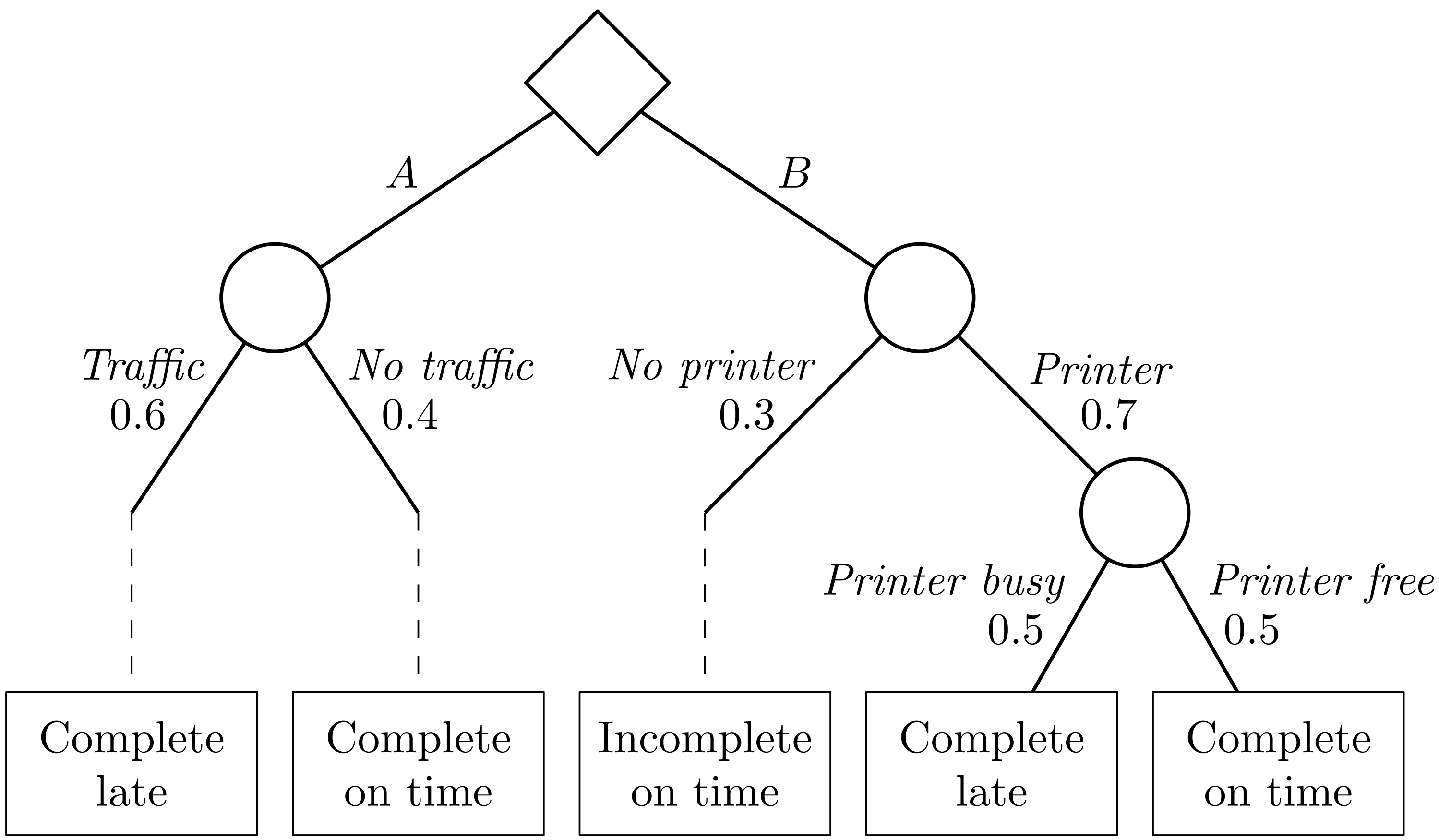
$$P(\mathrm{CT} \mid \mathrm{H}) = 0.4 \qquad P(\mathrm{CT} \mid \mathrm{U}) = 0.7 \times 0.5 = 0.35$$

$$P(\mathrm{CL} \mid \mathrm{H}) = 0.6 \qquad P(\mathrm{CL} \mid \mathrm{U}) = 0.7 \times 0.5 = 0.35$$

$$P(\mathrm{IT} \mid \mathrm{U}) = 0.3$$

TÉCNICO LISBOA

Expected value



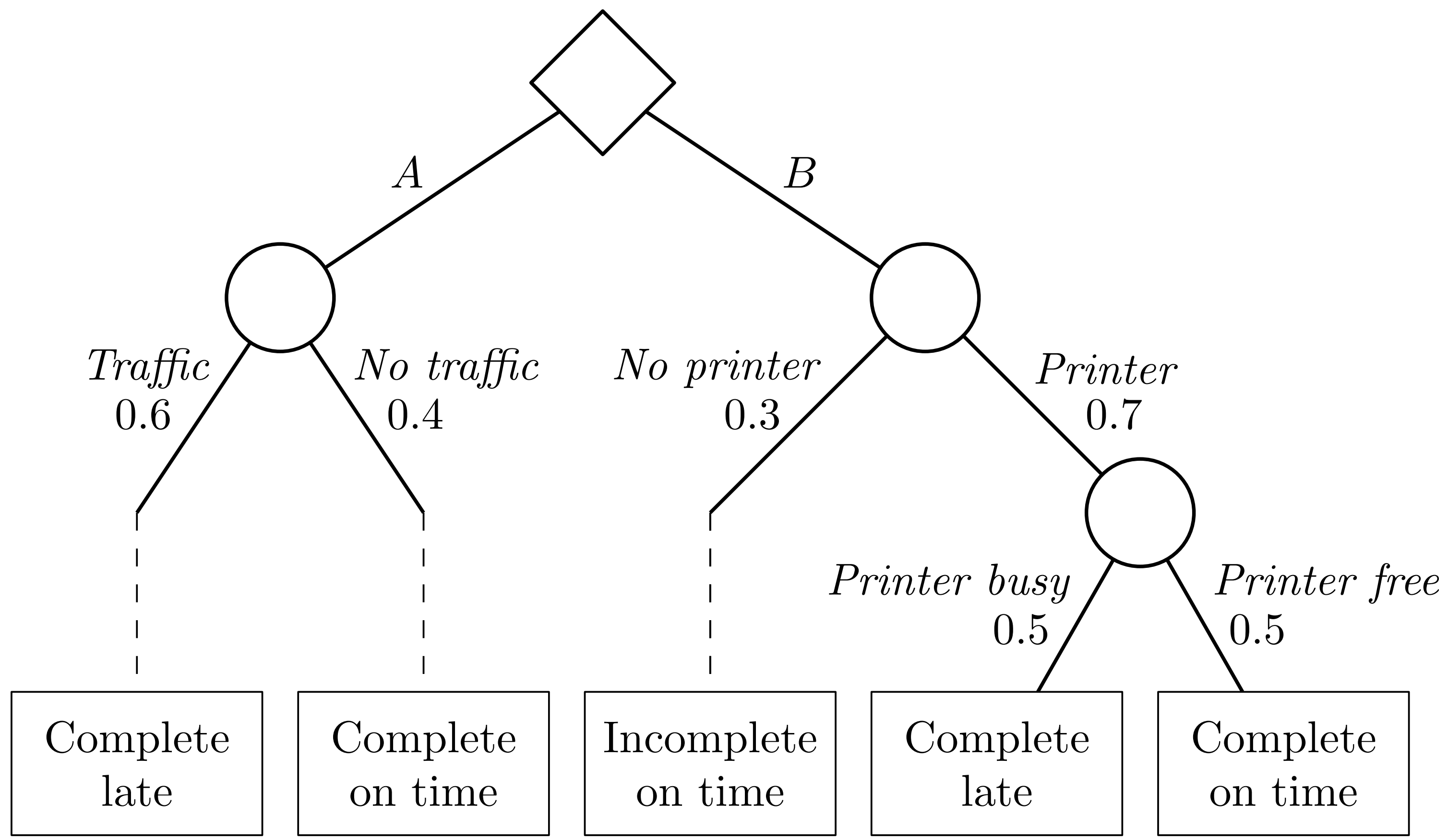
$$\mathrm{Q(H)} = 0.4 \times 0 + 0.6 \times (-2) = -1.2$$

$$P(\mathrm{CT} \mid \mathrm{U}) = 0.7 \times 0.5 = 0.35$$

$$P(\mathrm{CL} \mid \mathrm{U}) = 0.7 \times 0.5 = 0.35$$

$$P(\mathrm{IT} \mid \mathrm{U}) = 0.3$$

Expected value



$$Q(H) = 0.4 \times 0 + 0.6 \times (-2) = -1.2 \quad Q(U) = 0.35 \times 0 + 0.35 \times (-2) + 0.3 \times (-3) = -1.6$$

She should return home!

St. Petersburg paradox

St. Petersburg paradox

- A casino in St. Petersburg offers the following game
  - The initial prize is 2 rubles
  - A fair coin is tossed
  - If it comes out "tails", the game ends, and you get the prize
  - If it comes "heads", the prize doubles and the game continues

How much should a player pay to enter for the game to be fair?

Expected value

- Expected value:

  $P(\mathrm{T}) = 1/2 \longrightarrow$ Game ends with prize 2

  $P(\mathrm{HT}) = 1/4 \longrightarrow$ Game ends with prize 4

  $P(\mathrm{HHT}) = 1/8 \longrightarrow$ Game ends with prize 8

  ...

  $P(\mathrm{HH...HT}) = 1/2^n \longrightarrow$ Game ends with prize 2^n

$$\text{Expected value} = \frac{1}{2} \times 2 + \frac{1}{4} \times 4 + \ldots + \frac{1}{2^n} \times 2^n + \ldots = \sum_{n=1}^{\infty} 1 = \infty$$

TÉCNICO LISBOA

Would you take this bet?