Dynamic Bayesian Networks II

Particle filters

Szymon Jaroszewicz

Institute of Computer Science
Polish Academy of Sciences
Warsaw, Poland

Ph. D. Programme 2012/2013

PhD STUDIES

KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

Inference in general DBNs

- We saw that various types of inference can be performed efficiently on Hidden Markov Models
- What about general Dynamic Bayesian Networks?
- Do those algorithms still work?

- In principle: yes
- But the hidden state of a DBN is described by several variables
- Exact inference can only possible when those variables have some independence structure
- but...

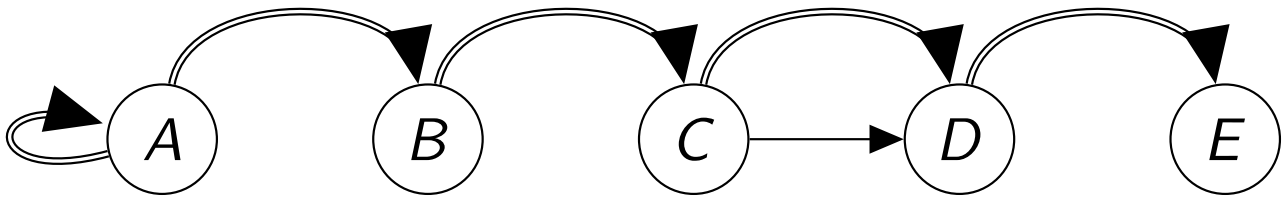
PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Inference in general DBNs – the entanglement problem

The entaglement problem

As we move with time more and more dependencies appear between variables. They become entangled

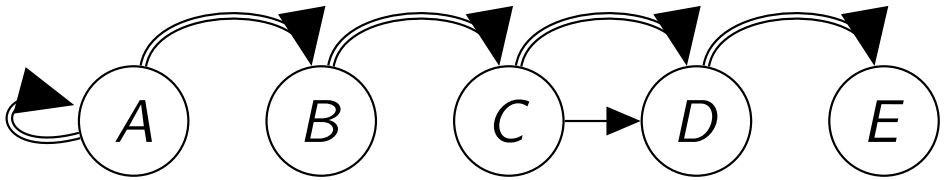
- Consider a rather sparse DBN with few dependencies within/between time slices

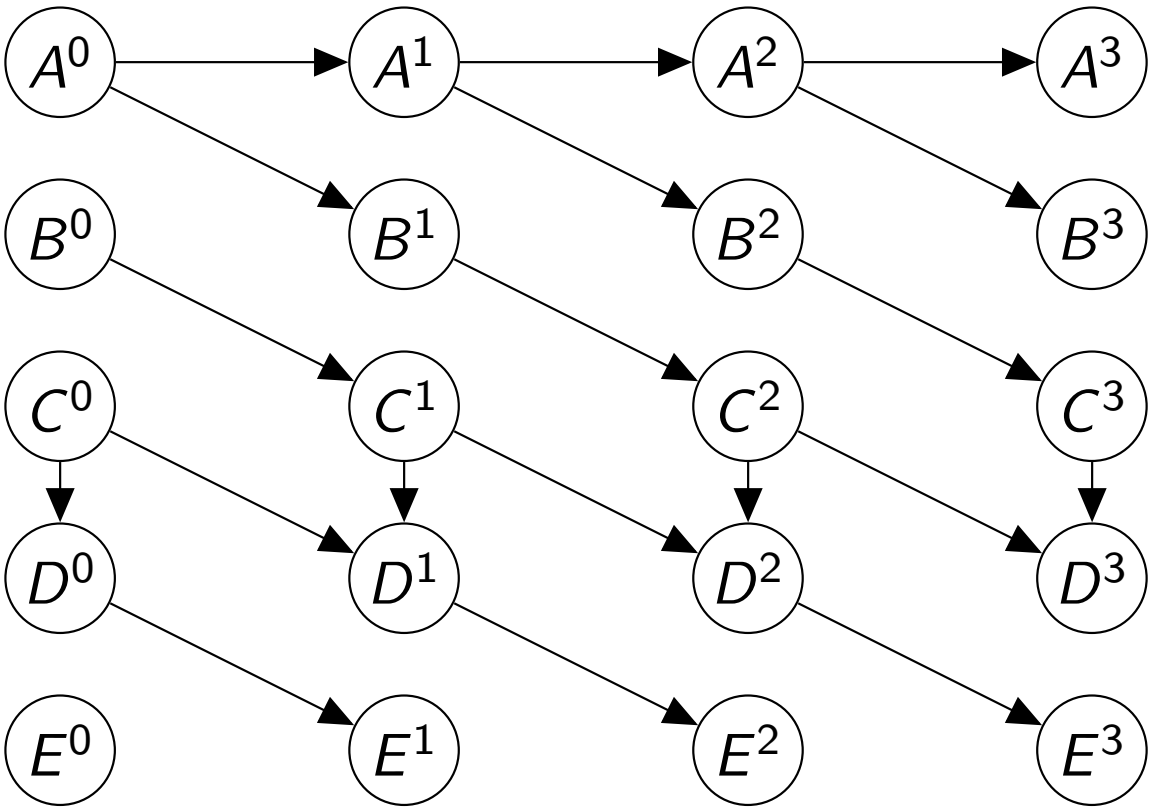


- Let's unroll it

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

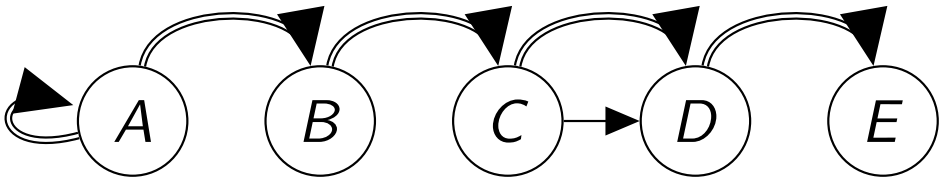
Inference in general DBNs – the entanglement problem

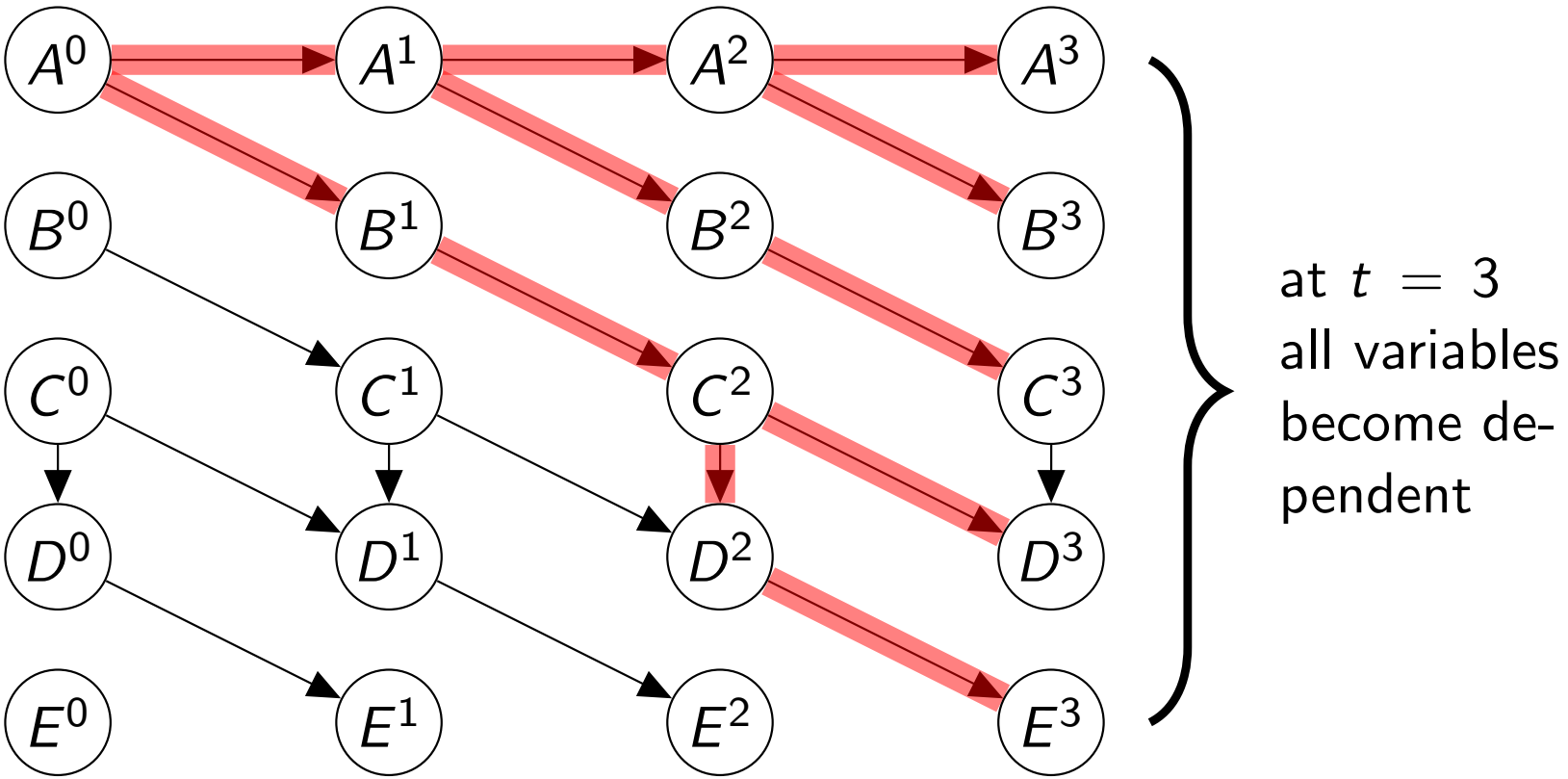




PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Inference in general DBNs – the entanglement problem

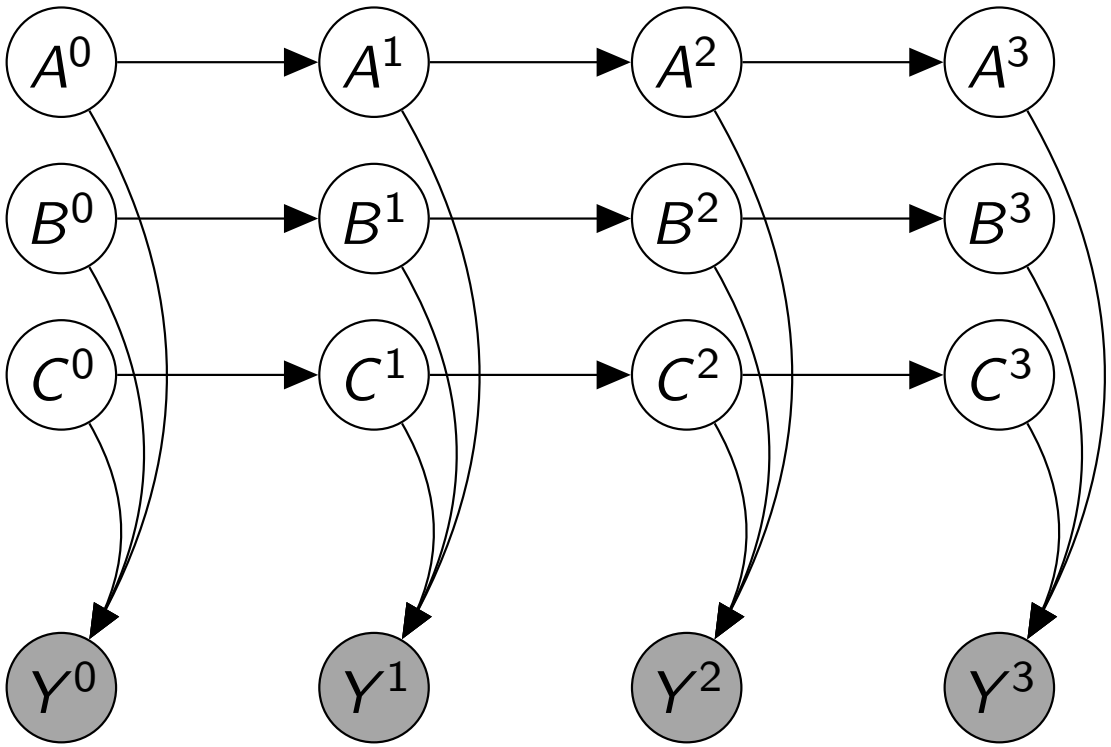




PhD STUDIES
KAPITAŁ LUDZKI NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA EUROPEJSKI FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Inference in general DBNs – the entanglement problem

- When observed outputs are present, the situation is even worse



- 3 otherwise independent internal variables become immediately dependent when the output is observed

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Inference in general DBNs – the entanglement problem

- The only real option is to somehow represent the full joint distribution of states at every time step!
- If there are more than a few internal variables, there is no hope of being able to do exact inference
- We have to approximate!

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Inference in general DBNs – approximating the joint distribution

Great many approximations have been proposed

- Mixtures of Gaussians
- Product approximations (the Boyen-Koller filter)
- Sampling (particle filters)
- ...

We will briefly mention the Boyen-Koller filter, and spend more time on particle filters

But first let's talk about the filtering task in general BDNs

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Inference in general DBNs – filtering

- We performed several different types of inference on HMMs
- It is possible to do them on the DBNs also
- But the most frequently discussed problem is filtering, that is computing

$$P(\mathbf{X}^t|\mathbf{y}^{1..t})$$

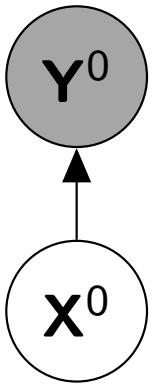
  the estimate of current state based on all previous outputs
- Most algorithms are called filters

- Given the solution at time t we want to update it to time $t+1$
- This is in fact the forward algorithm for DBNs
- Practical motivation: tracking the system's hidden state in real time

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

The forward algorithm for DBNs

- Main idea: start at $t = 0$ and proceed recursively to $t = 1, 2, \ldots$
- The time slice of the DBN at $t = 0$:



  where $\mathbf{X}$ and $\mathbf{Y}$ are sets of variables
- Joint distribution

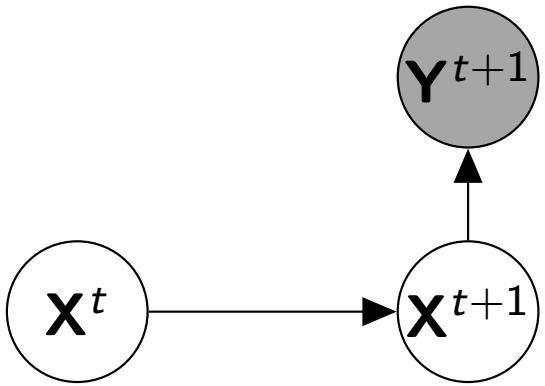
$$P(\mathbf{Y}^0 = \mathbf{y}^0, \mathbf{X}^0) = P(\mathbf{X}^0)P(\mathbf{y}^0|\mathbf{X}^0)$$

- How to get $P(\mathbf{X}^0|\mathbf{y}^0)$? Just normalize
- So the base case is easy

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

The forward algorithm for DBNs

- Suppose we have computed $P(\mathbf{X}^t|\mathbf{y}^{0..t})$
- How do we compute $P(\mathbf{X}^{t+1}|\mathbf{y}^{0..t+1})$?
- The relevant slice of the DBN:



- Joint distribution:

$$P(\mathbf{X}^t|\mathbf{y}^{0..t})P(\mathbf{X}^{t+1}|\mathbf{X}^t)P(\mathbf{y}^{t+1}|\mathbf{X}^{t+1})$$

- Eliminate $\mathbf{X}^t$ and normalize to get $P(\mathbf{X}^{t+1}|\mathbf{y}^{0..t+1})$
- Proceed to the next time step

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

The forward algorithm for DBNs

- The elimination/normalization is a two step process

$$P(\mathbf{X}^{t+1}|\mathbf{y}^{0..t}) = \sum_{\mathbf{X}^t} P(\mathbf{X}^t|\mathbf{y}^{0..t})P(\mathbf{X}^{t+1}|\mathbf{X}^t)$$

$$P(\mathbf{X}^{t+1}|\mathbf{y}^{0..t+1}) = \frac{P(\mathbf{y}^{t+1}|\mathbf{X}^{t+1})P(\mathbf{X}^{t+1}|\mathbf{y}^{0..t})}{\sum_{\mathbf{X}^{t+1}} P(\mathbf{y}^{t+1}|\mathbf{X}^{t+1})P(\mathbf{X}^{t+1}|\mathbf{y}^{0..t})}$$

- Both sums a hard to compute if $|\mathbf{X}|$ is large

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

The forward algorithm for DBNs

- DBNs are frequently used with continuous variables (tracking)
- For continuous variables the update steps become

$$P(\mathbf{X}^{t+1}|\mathbf{y}^{0..t}) = \int P(\mathbf{X}^{t}|\mathbf{y}^{0..t})P(\mathbf{X}^{t+1}|\mathbf{X}^{t})\, d\mathbf{X}^{t}$$

$$P(\mathbf{X}^{t+1}|\mathbf{y}^{0..t+1}) = \frac{P(\mathbf{y}^{t+1}|\mathbf{X}^{t+1})P(\mathbf{X}^{t+1}|\mathbf{y}^{0..t})}{\int P(\mathbf{y}^{t+1}|\mathbf{X}^{t+1})P(\mathbf{X}^{t+1}|\mathbf{y}^{0..t})\, d\mathbf{X}^{t+1}}$$

- P is now a probability density function
- Both integrals can be hard to compute even if $|\mathbf{X}|$ is small (depending on actual distributions)

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

The forward algorithm for DBNs

- Due to the entaglement problem those sums/integrals cannot be computed using variable elimination
- It is necessary to approximate the state distribution $P(\mathbf{X})$
- We'll discuss three approaches
  - the Kalman filter – sketch only
  - product decomposition (the BK filter) – sketch only
  - particle filters

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

The Kalman filter

- If for some matrices $\mathbf{A}$, $\mathbf{B}$
  - $\mathbf{X}^0 \sim N(\mu_0, \mathbf{\Sigma}_0)$
  - $\mathbf{X}^{t+1}|\mathbf{X}^t = \mathbf{A}\mathbf{X}^t + N(\mathbf{0}, \mathbf{\Sigma}_1)$
  - $\mathbf{Y}^t|\mathbf{X}^t = \mathbf{B}\mathbf{X}^t + N(\mathbf{0}, \mathbf{\Sigma}_2)$

  then the DBN is called a Kalman filter
- Exact inference is possible using linear algebra
- Kalman filters are very popular in control, video image analysis, etc.

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

The Boyen-Koller (BK) filter – sketch

- Introduced in Xavier Boyen and Daphne Koller in *Tracable Inference for Complex Stochastic Processes*, UAI'98. The paper is here
- The approach: factorize the joint distribution of states even though this violates the dependencies between variables
- We thus approximate the full joint distribution by a product of smaller distributions
- How does this affect the error of the forward algorithm?
- One worries that it will get worse and worse as the algorithm proceeds
  ... and the results will quickly become useless

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

The Boyen-Koller (BK) filter – sketch

A positive result

Under certain assumptions the error never grows above a certain threshold

- Main idea:
  - as time goes by the system forgets the initial state so the initial distribution and its approximation become closer and closer to each other
    - similar to Markov Chains converging to stationary distributions
  - if this convergence is fast enough, it will offset the growing approximation error

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

The Boyen-Koller (BK) filter – sketch

Definition

Let $P(\mathbf{X}^{t+1}|\mathbf{X}^t)$ be the transition probabilities for the internal state of the DBN. The minimum mixing rate of the DBN is

$$\gamma = \min_{\mathbf{x}_1,\mathbf{x}_2} \sum_{\mathbf{x}} \min\{P(\mathbf{x}|\mathbf{x}_1), P(\mathbf{x}|\mathbf{x}_2)\}$$

- It is the minimum probability that, starting in two different states at time t, we end up in the same state at $t+1$

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

The Boyen-Koller (BK) filter – sketch

Weakly interacting set of processes

Suppose the internal state $\mathbf{X}$ can be partitioned into blocks $\mathbf{X}_1 \cup \ldots \cup \mathbf{X}_k$ such that

- there are no edges within the same time slice between $\mathbf{X}_i$ and $\mathbf{X}_j$, $i \neq j$
- each $\mathbf{X}_i$ influences at most q other blocks at time $t+1$
- at time $t+1$ each $\mathbf{X}_i$ is influenced by at most r other blocks

- Of course $\mathbf{X}_i^{t+1}$ usually depends on $\mathbf{X}_i^t$
- Intuition: each $\mathbf{X}_i$ forms a subprocess within the DBN and the amount of interaction between subprocesses is limited
- Idea: approximate $P(\mathbf{X})$ using $\tilde{P}(\mathbf{X}) = \prod P(\mathbf{X}_i)$

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

The Boyen-Koller (BK) filter – sketch

Theorem

Suppose that:

- the internal state of a DBN can be decomposed into k weakly interacting subprocesses such that each subprocess has a minimum mixing rate at least γ
- for all t, using the approximation $\tilde{P}$ increases the error by at most ε:

$$KL(P : \tilde{P}) - KL(P : \hat{P}) \leq \varepsilon$$

where $\hat{P}$ is the unapproximated estimate of P

Then

$$\mathrm{E}KL(P : \tilde{P}) \leq \varepsilon/(\gamma/r)^q$$

$KL(P : Q)$ is the Kullback-Leibler divergence between distributions P and Q

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

The Boyen-Koller (BK) filter – sketch

- Conclusion: the approximation error does not grow above certain threshold no matter how long we perform the filtering
- In other words: approximation errors do not accumulate above $\varepsilon/(\gamma/r)^q$

- Very nice theoretically
- In practice works well if the internal state can really be decomposed into weakly interacting subprocesses

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Particle filters

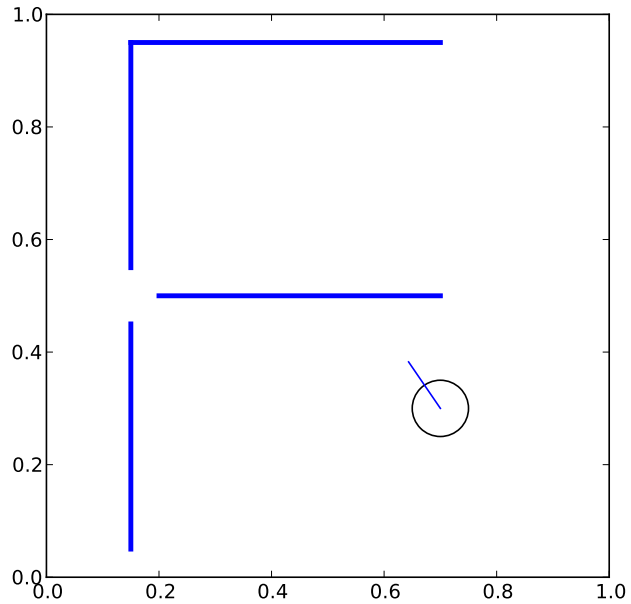
PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Example: robot localization problem

Let us begin with a motivating example

Given:

- the terrain map
- the history of moves a robot has made
- the history of robot's sensor readings



PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Example: robot localization problem

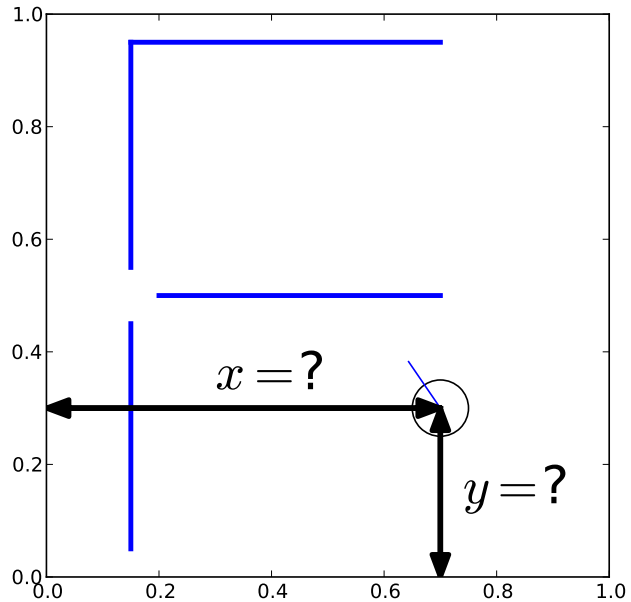
Let us begin with a motivating example

Given:

- the terrain map
- the history of moves a robot has made
- the history of robot's sensor readings

Find:

- Robot's position on the map



PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

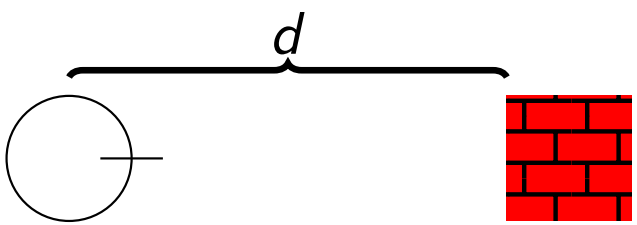
Example: robot localization problem

- Robot's position (internal state of the DBN):

$$\mathbf{X}^t = \begin{pmatrix} x^t \\ y^t \\ \alpha^t \end{pmatrix}$$

- There is only one sensor which measures the distance from obstacles

$$\mathbf{Y} = \frac{1}{d^2(\mathbf{X})} + N(0, \epsilon)$$



- $d(\mathbf{X})$ is the distance to the nearest obstacle in the direction the sensor is pointing
- it is a function of $\mathbf{X}$ and of the coordinates of walls on the map

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Example: robot localization problem

- We assume the uniform a-priori distribution

$$x^0, y^0 \sim U(0,1), \quad \alpha^0 \sim U(0, 2\pi)$$

- $P(\mathbf{X}^{t+1}|\mathbf{X}^t)$ depends on the move the robot made at time t
- Since the moves are not perfectly precise, the new position has random errors
- Possible moves:
  - rotate by angle β

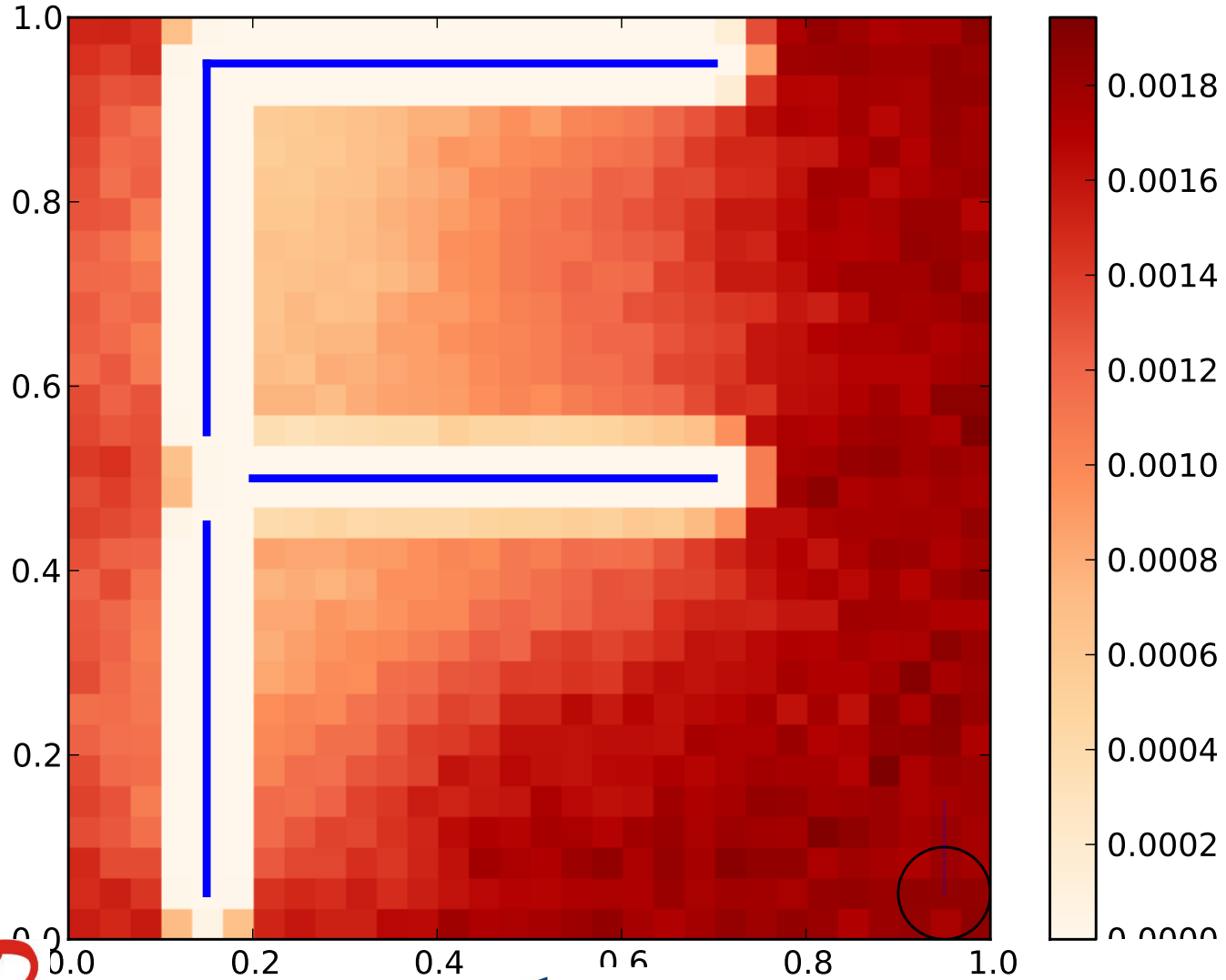
$$\mathbf{X}^t = \begin{pmatrix} x \\ y \\ \alpha \end{pmatrix}, \qquad \mathbf{X}^{t+1} = \begin{pmatrix} x \\ y \\ \alpha + \beta + \epsilon \end{pmatrix}$$

  - move forward by p

$$\mathbf{X}^t = \begin{pmatrix} x \\ y \\ \alpha \end{pmatrix}, \qquad \mathbf{X}^{t+1} = \begin{pmatrix} x + (p + \epsilon)\cos\alpha \\ y + (p + \epsilon)\sin\alpha \\ \alpha \end{pmatrix}$$

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Example: robot localization problem



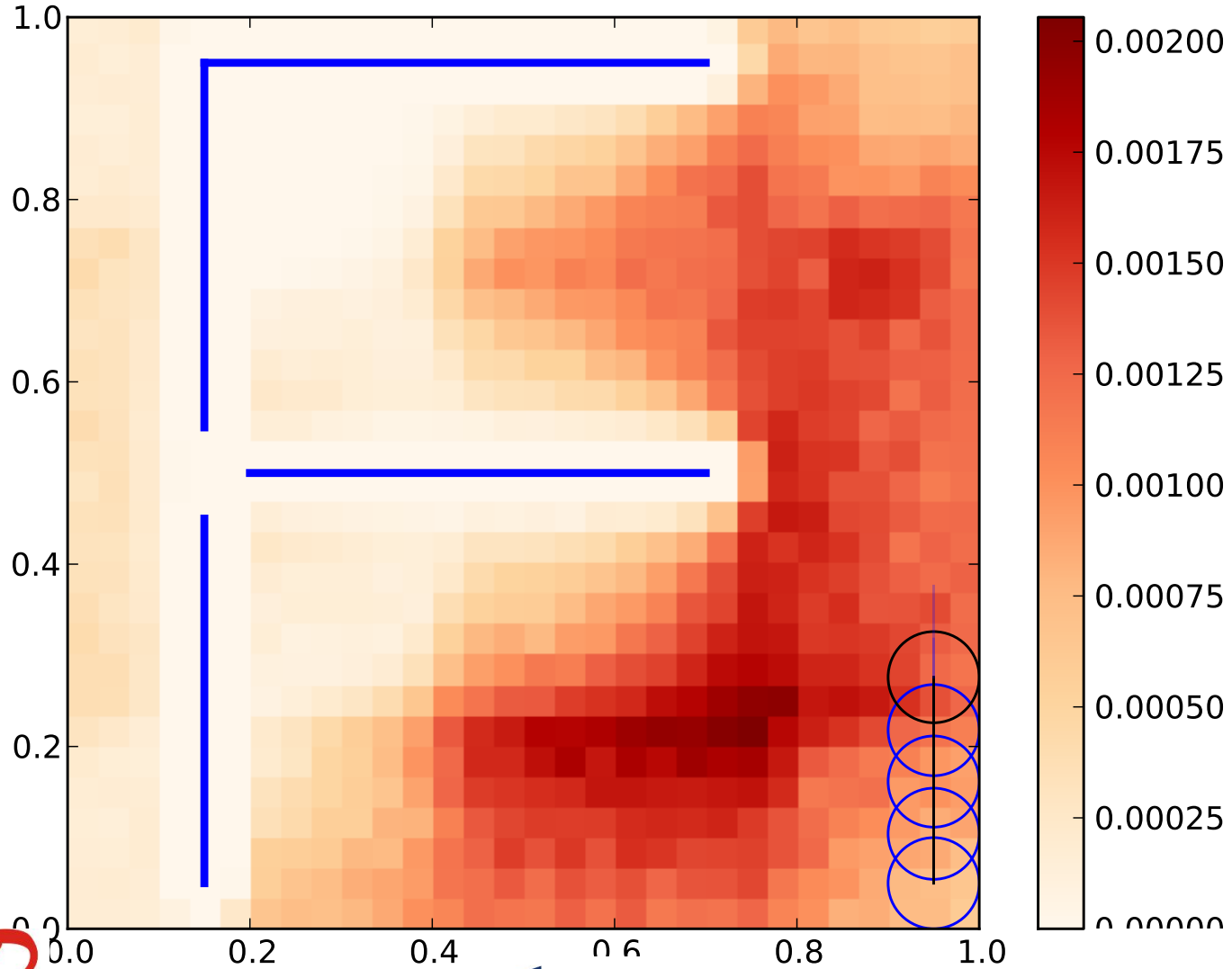
PhD STUDIES

KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

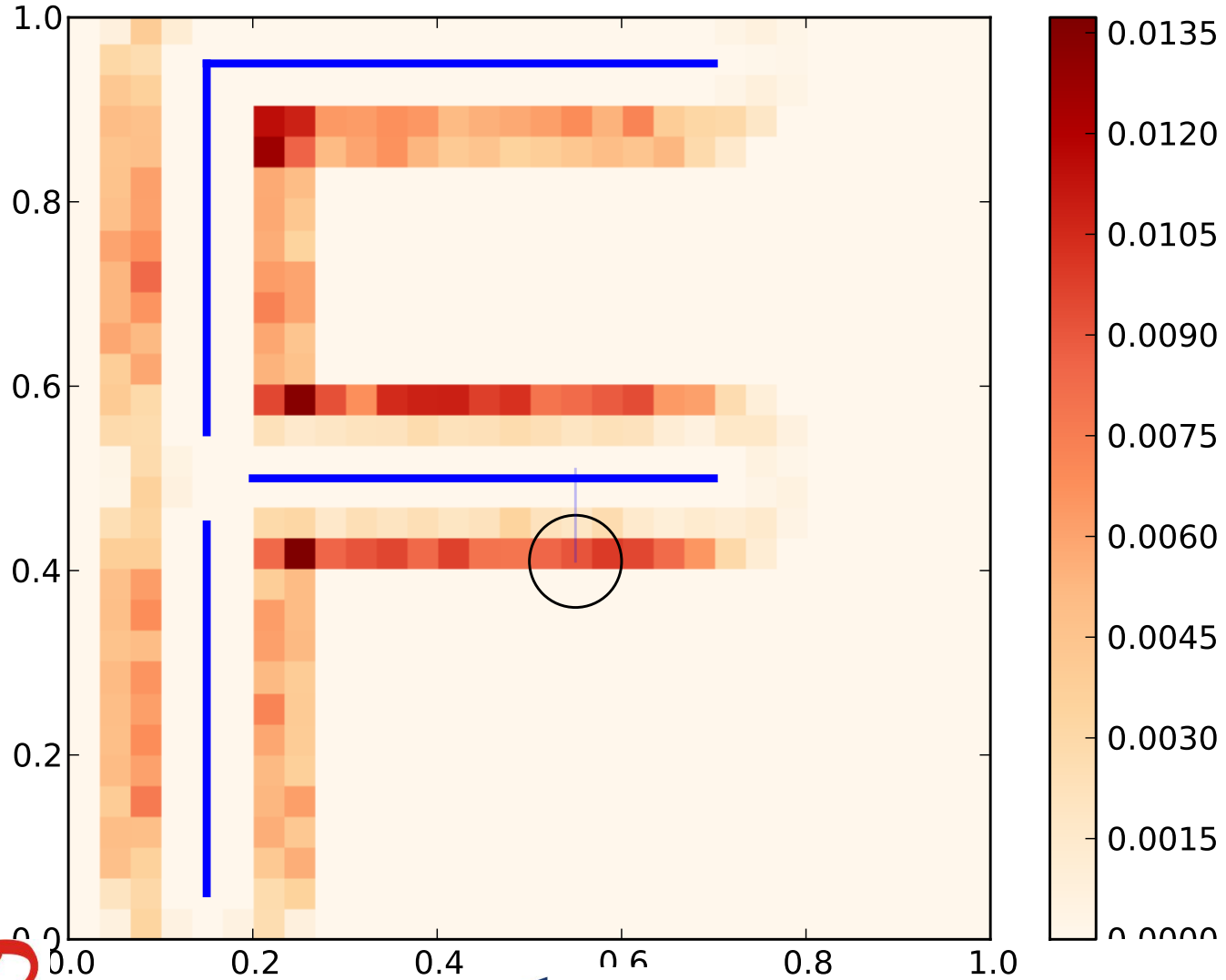
Example: robot localization problem



PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

Example: robot localization problem



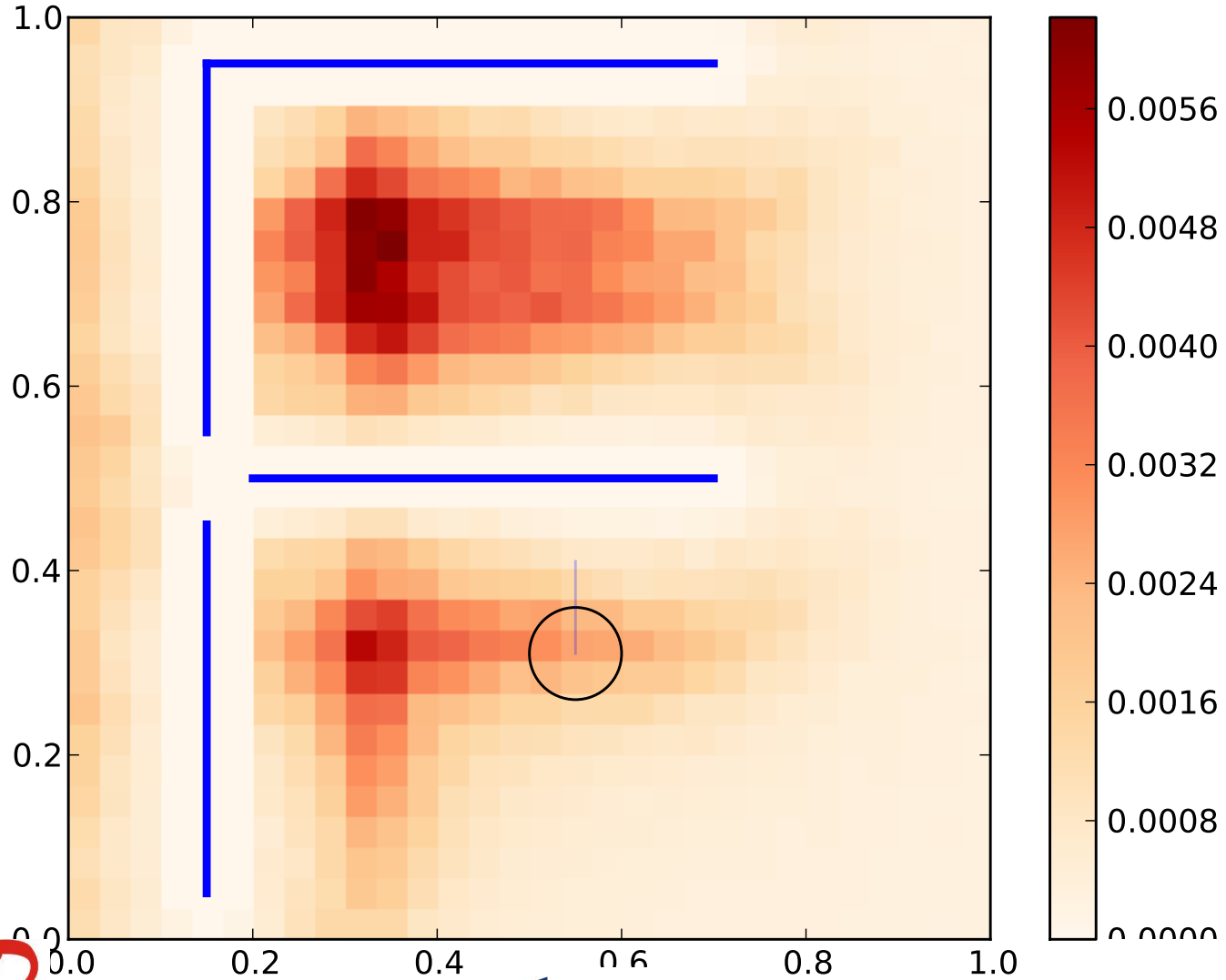
PhD STUDIES

KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

Example: robot localization problem



PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

Particle filters

- particle = sample + weight

$$(\mathbf{x}_i^t, w_i^t)$$

- A set of particles $\{(\mathbf{x}_i^t, w_i^t) : i = 1, \ldots, N\}$ approximates the distribution $P(\mathbf{X}^t)$ of states at time t
- Samples and weights are modified appropriately as we move from t to $t + 1$

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Particle filters – the basic algorithm

1. Initialization
   - $t \leftarrow 0$
   - draw N samples $\mathbf{x}_i$ from $P(\mathbf{X}^0)$
   - $w_i \leftarrow P(\mathbf{y}^0|\mathbf{x}_i^0)$ for $i = 1 \ldots N$
2. Draw new samples from $P(\mathbf{X}^{t+1}|\mathbf{x}_i^t)$

$$\mathbf{x}_i^{t+1} \sim P(\mathbf{X}^{t+1}|\mathbf{x}_i^t)$$

3. Update particle weights based on $\mathbf{y}^{t+1}$, the sensor readings at time $t+1$

$$w_i \leftarrow w_i \cdot P(\mathbf{y}^{t+1}|\mathbf{x}_i^{t+1})$$

4. Normalize weights

$$w_i \leftarrow \frac{w_i}{\sum_{i=1}^{N} w_i}$$

5. $t \leftarrow t+1$; **Goto** 2

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Particle filters – proof

- How do we know that a PF gives us the correct distribution?
- Let's rewrite

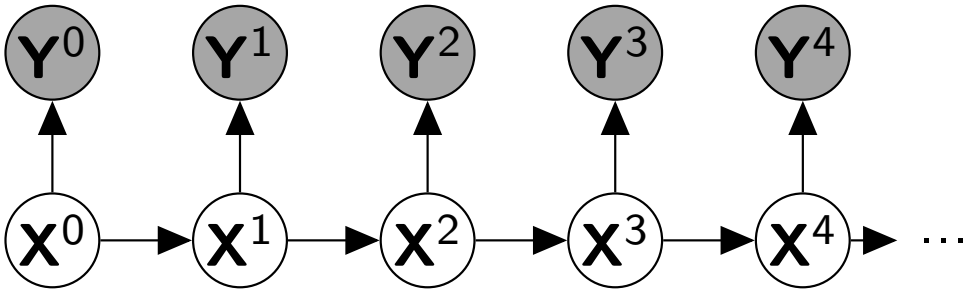
$$P(\mathbf{X}^{1..T}|\mathbf{y}^{1..T}) \propto P(\mathbf{X}^0)P(\mathbf{y}^0|\mathbf{X}^0)\prod_{t=1}^{T} P(\mathbf{X}^t|\mathbf{X}^{t-1})P(\mathbf{y}^t|\mathbf{X}^t)$$

$$= \underbrace{P(\mathbf{X}^0)\prod_{t=1}^{T} P(\mathbf{X}^t|\mathbf{X}^{t-1})}_{\text{samples}}\underbrace{P(\mathbf{y}^0|\mathbf{X}^0)\prod_{t=1}^{T} P(\mathbf{y}^t|\mathbf{X}^t)}_{\text{weights}}$$

- Note that samples are taken from a pure Markov Chain independent of outputs
- Weights are applied independently from sampling
- We've seen such a decomposition before...

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Particle filters – proof

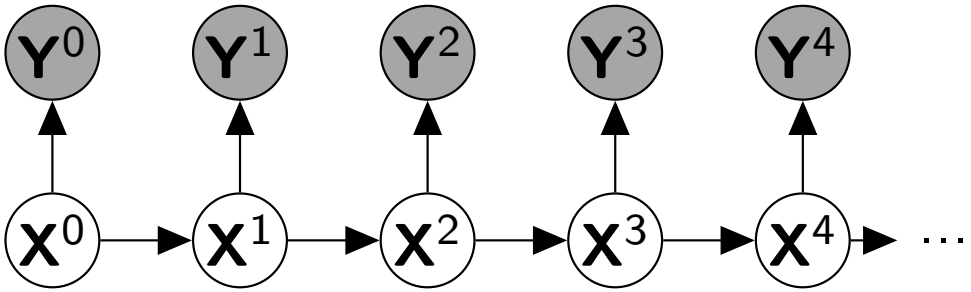


$$P(\mathbf{X}^{1..T}|\mathbf{y}^{1..T}) \propto \underbrace{P(\mathbf{X}^0)\prod_{t=1}^{T} P(\mathbf{X}^t|\mathbf{X}^{t-1})}_{\text{samples}}\ \underbrace{P(\mathbf{y}^0|\mathbf{X}^0)\prod_{t=1}^{T} P(\mathbf{y}^t|\mathbf{X}^t)}_{\text{weights}}$$

- This is an example of ????
- We have already proved it works

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Particle filters – proof



$$P(\mathbf{X}^{1..T}|\mathbf{y}^{1..T}) \propto \underbrace{P(\mathbf{X}^0)\prod_{t=1}^{T} P(\mathbf{X}^t|\mathbf{X}^{t-1})}_{\text{samples}}\ \underbrace{P(\mathbf{y}^0|\mathbf{X}^0)\prod_{t=1}^{T} P(\mathbf{y}^t|\mathbf{X}^t)}_{\text{weights}}$$

- This is an example of likelihood weighting
- We have already proved it works

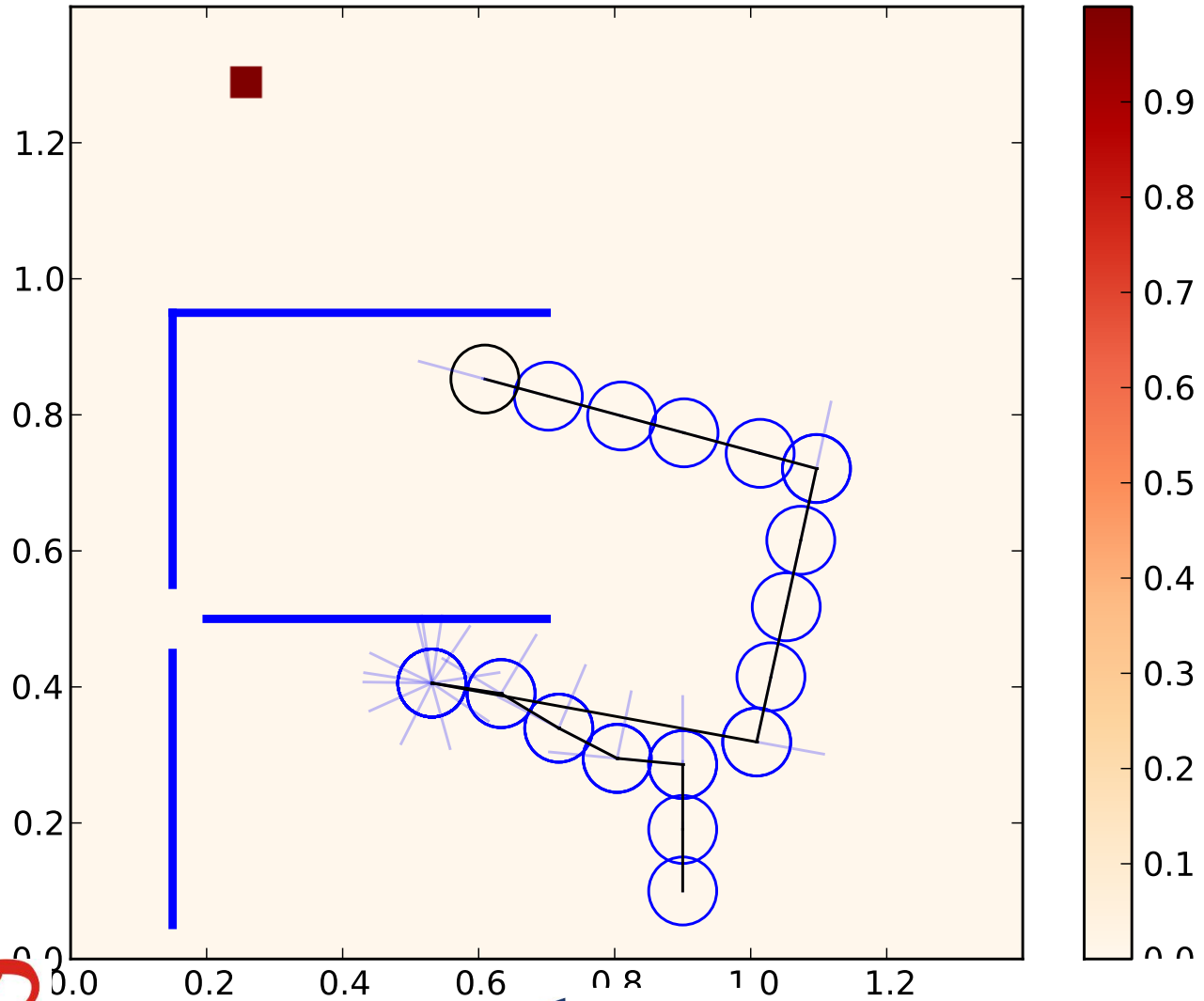
PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Sample degeneracy

- Problem with the basic particle filter
- Sample degeneracy
  - one of the particles (not necessarily closest to the true location) initially gives sensor readings most similar to true readings
  - this particle gets the largest weight
  - this is likely to remain so for a few steps
  - after a few steps all other particles have weights practically equal to zero
- Result: estimates based on a single sample, hugely inaccurate

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

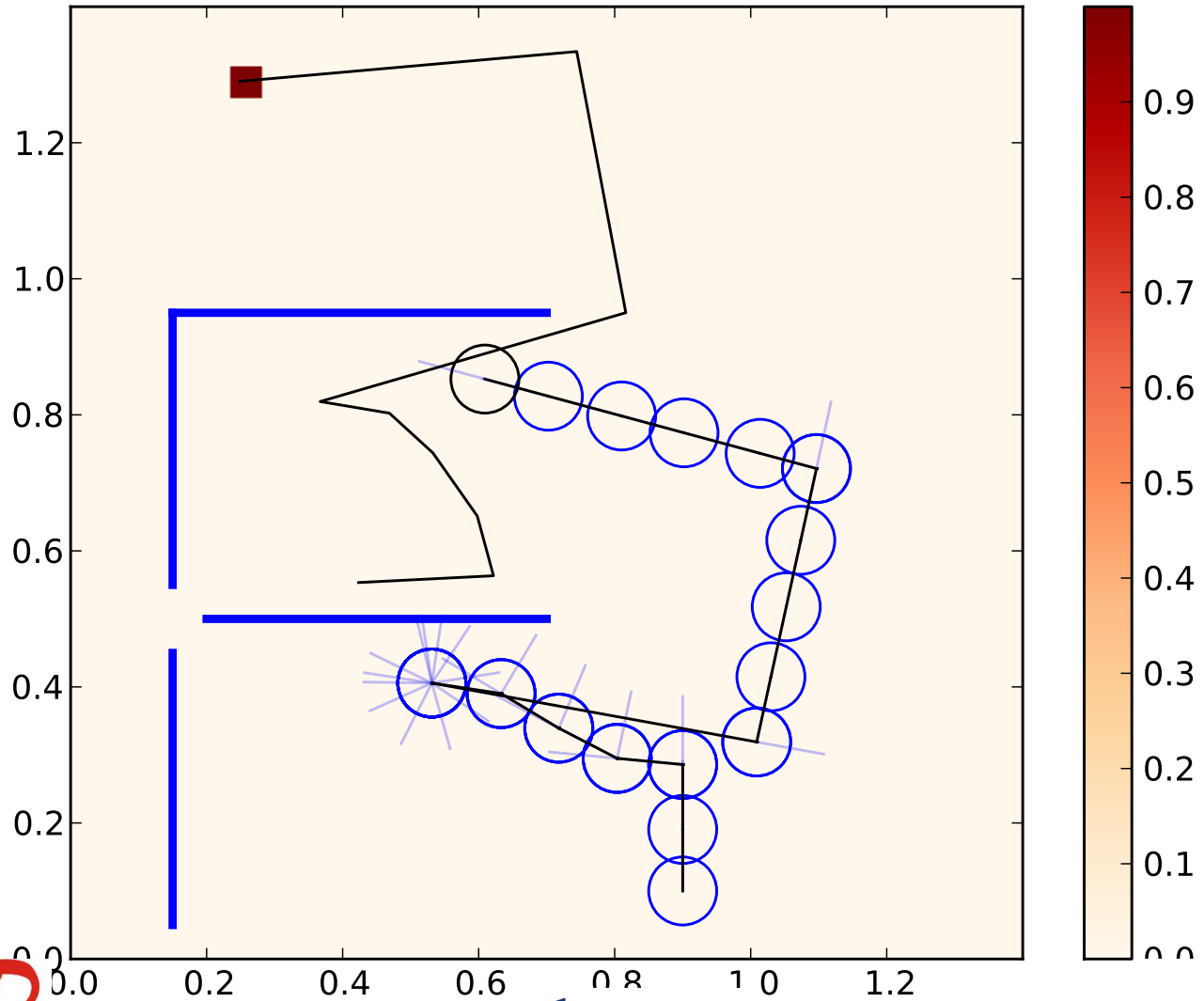
Sample degeneracy example (100000 samples)



PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

Sample degeneracy – trajectory of the best particle



PhD STUDIES

KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

Sample degeneracy – some theory

Definition

Effective sample size is defined as

$$N_{eff} = \frac{1}{\sum {w_i}^2}$$

- The definition is intuitively correct:
  - When all samples have equal weights ($\frac{1}{N}$), $N_{eff} = N$
  - When one sample weight is equal to 1 and all other to 0, $N_{eff} = 1$
- Proper derivation of N_{eff}
  - use Taylor expansion of weighted sample variance
- Here we will use a simplified derivation assuming the weights are constant (not r.v.'s)

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Effective sample size

- Let f be a function of $\mathbf{X}^t$. We want to estimate $\mathrm{E}_{P(\mathbf{X}^t)} f$
- If we had N true unweighted i.i.d. samples, the variance of the estimate would be $\frac{1}{N}\mathrm{Var}(f(\mathbf{X}^t))$
- Suppose we have a weighted estimate $\sum_{i=1}^{N} w_i f(\mathbf{x}_i^t)$
- Now

$$\mathrm{Var}\left[\sum_{i=1}^{N} w_i f(\mathbf{x}_i^t)\right] = \sum_{i=1}^{N} w_i^2 \mathrm{Var}(f(\mathbf{x}_i^t)) \approx \mathrm{Var}(f(\mathbf{X}^t)) \sum_{i=1}^{N} w_i^2$$

  the last assumption is necessary since $\mathbf{x}_i^t$ comes from the distribution ignoring output symbols
- How many unweighted i.i.d. samples N_{eff} do we need to get the same variance?

$$\frac{1}{N_{eff}}\mathrm{Var}(f(\mathbf{X}^t)) = \mathrm{Var}(f(\mathbf{X}^t)) \sum_{i=1}^{N} w_i^2 \quad \Rightarrow \quad N_{eff} = \frac{1}{\sum w_i^2}$$

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Effective sample size

Theorem

Effective sample size N_{eff}

- is never greater than the number of particles N
- can only decrease with each time step

- For details and proofs see
  - Kong, Liu, Wong 1994
  - Doucet, Godsill, Andrieu, 2000

- Avoiding sample degeneracy:
  - resampling
  - importance sampling

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Resampling

Main idea:

- From the current set of particles pick N samples with replacement
- Sampling probability is proportional to particle weight
- Set weight of each sample to $\frac{1}{N}$
- Replace the current set of particles with the new sample

Advantages:

- This way we will get a more varied sample
- If we resample early enough, new sample is based on several important particles which are all assigned high weights

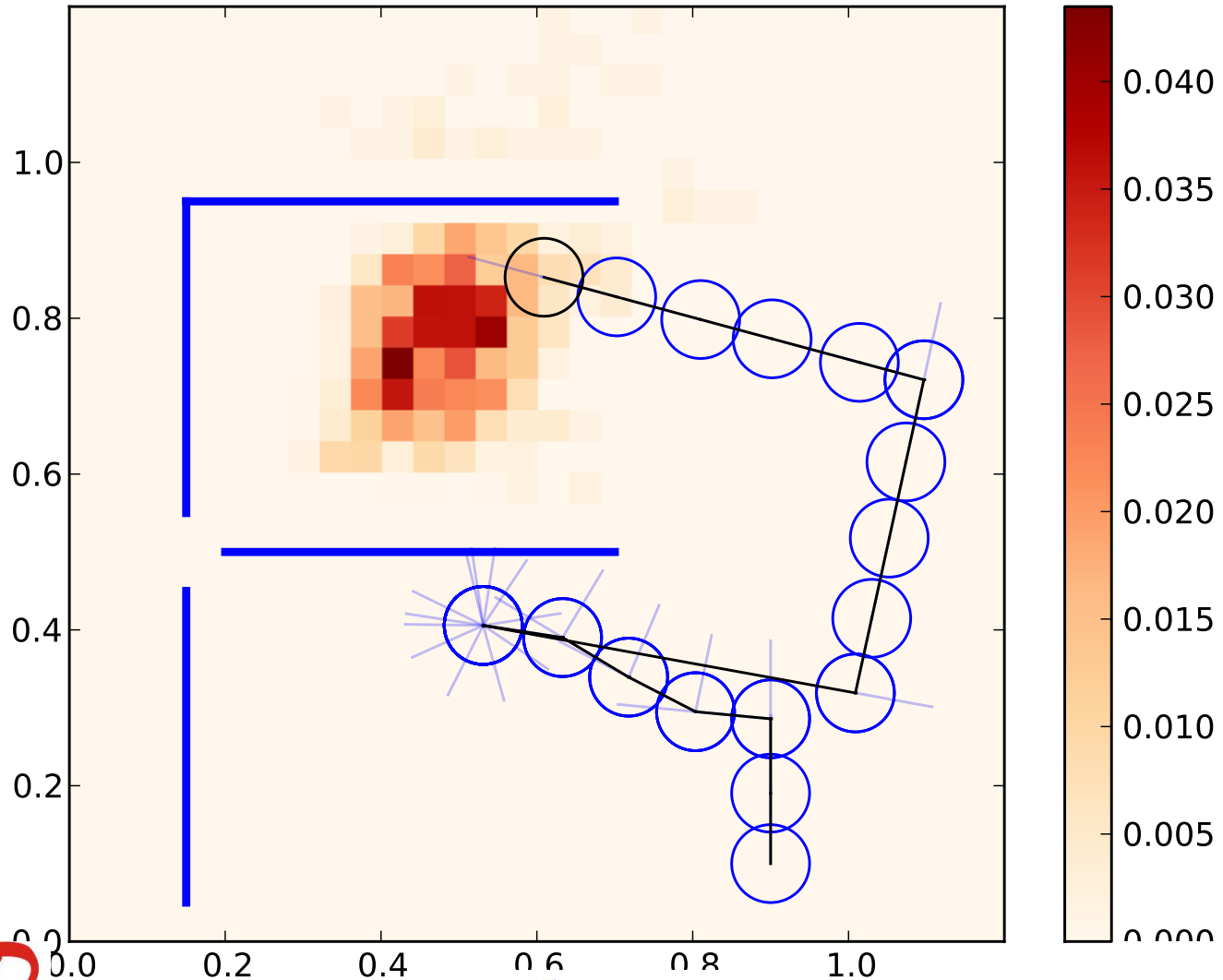
PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Particle filter with resampling

1. Initialization
   - $t \leftarrow 0$
   - draw N samples $\mathbf{x}_i$ from $P(\mathbf{X}^0)$
   - $w_i \leftarrow P(\mathbf{y}^0|\mathbf{x}_i^0)$ for $i = 1 \ldots N$
2. Draw new samples from $P(\mathbf{X}^{t+1}|\mathbf{x}_i^t)$
3. Update particle weights based on $\mathbf{y}^{t+1}$
4. Normalize particle weights
5. Compute effective sample size N_{eff}
6. **If** $N_{eff} < \epsilon N$
7. Resample
8. $t \leftarrow t + 1$; **Goto** 2

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

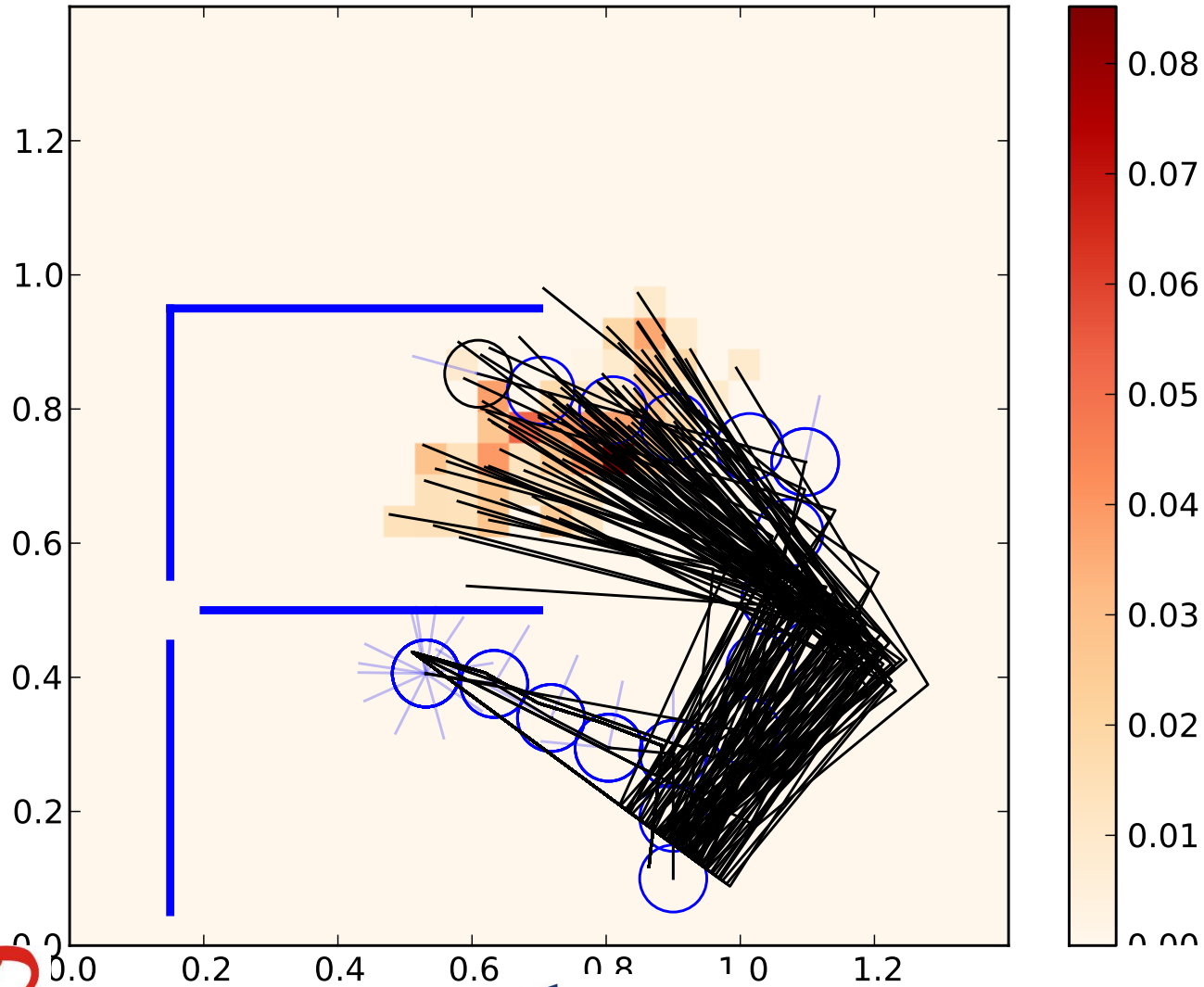
Example: resampling with 1000 samples



PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

Example: resampling with 100 samples



PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

Importance sampling

- The standard particle filter draws samples from

$$P(\mathbf{X}^{t+1}|\mathbf{x}_i^t)$$

- However, we also know $\mathbf{y}^{t+1}$ which could be used to get better estimates

- Idea: use importance sampling
  - instead of sampling from $P(\mathbf{X}^{t+1}|\mathbf{x}_i^t)$
  - sample $\mathbf{X}^{t+1}$ from some other proposal distribution

$$\pi(\mathbf{X}^{t+1}|\mathbf{x}^t, \mathbf{y}^{t+1})$$

  - use appropriate weights to accommodate the difference

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Particle filters – importance sampling

Only small changes to the main loop are necessary:

- Draw new samples from $\pi(\mathbf{X}^{t+1}|\mathbf{x}_i^t, \mathbf{y}^{t+1})$

$$\mathbf{x}_i^{t+1} \sim \pi(\mathbf{X}^{t+1}\ |\mathbf{x}_i^t, \mathbf{y}^{t+1})$$

- Update particle weights

$$w_i \leftarrow w_i \cdot P(\mathbf{y}^{t+1}|\mathbf{x}_i^{t+1}) \frac{P(\mathbf{x}_i^{t+1}|\mathbf{x}_i^t)}{\pi(\mathbf{x}_i^{t+1}|\mathbf{x}_i^t, \mathbf{y}^{t+1})}$$

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Importance sampling

- The optimal proposal distribution for sampling $\mathbf{X}^{t+1}$ is

$$P(\mathbf{X}^{t+1}|\mathbf{x}_i^t, \mathbf{y}^{t+1})$$

- Unfortunately
  - this distribution is hard to sample from
  - updating weights (see previous slide) requires computing

$$P(\mathbf{y}^{t+1}|\mathbf{x}^t) = \int P(\mathbf{y}^{t+1}|\mathbf{X}^{t+1})P(\mathbf{X}^{t+1}|\mathbf{x}^t)\, d\mathbf{X}^{t+1}$$

- Typically we have to use other approximations

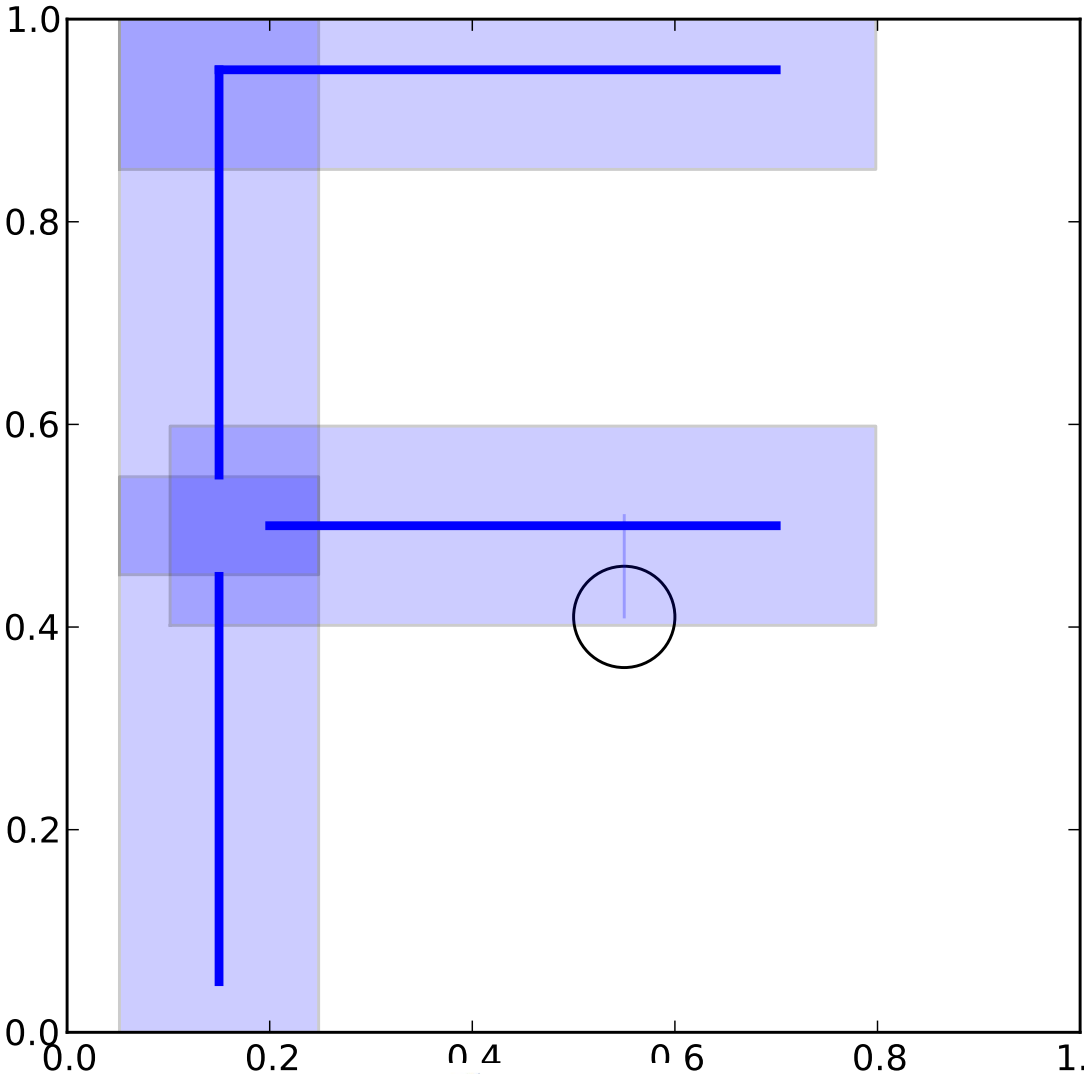
PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Importance sampling: robot localization problem

- We will only discuss the initial position
- Instead of sampling the initial position from $U([0,1] \times [0,1])$ we will take $\mathbf{y}^0$ into account
- High value of $\mathbf{y}^0$ means we are close to a wall
- Our importance distribution:
  - a mixture of uniform distributions on rectangles around walls
  - the higher $\mathbf{y}^0$, the narrower the rectangles
- Other proposals are of course possible e.g. Gaussians 'stretched' along the walls

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Importance sampling: proposal distribution



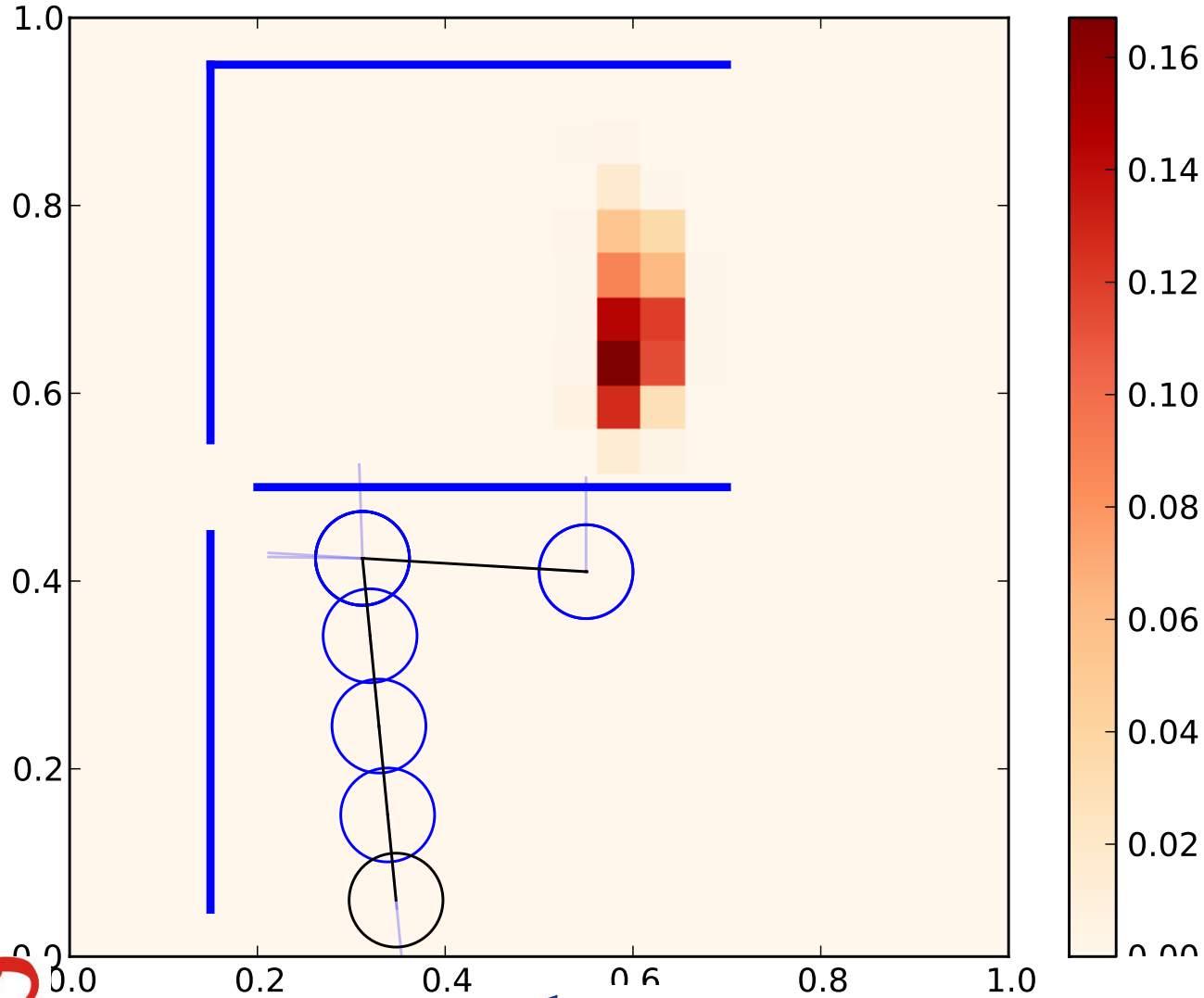
PhD STUDIES

KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

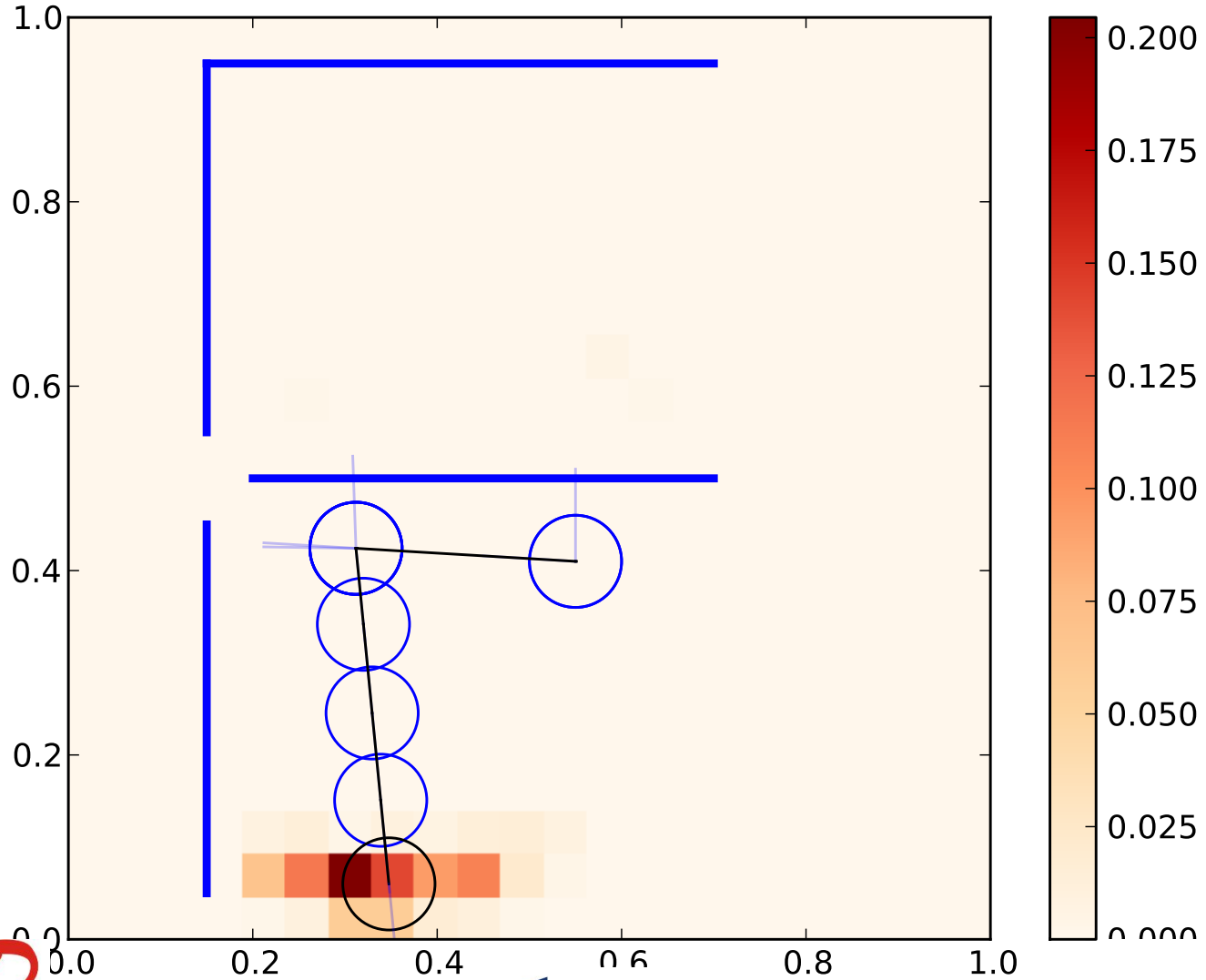
Example: no importance sampling



PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

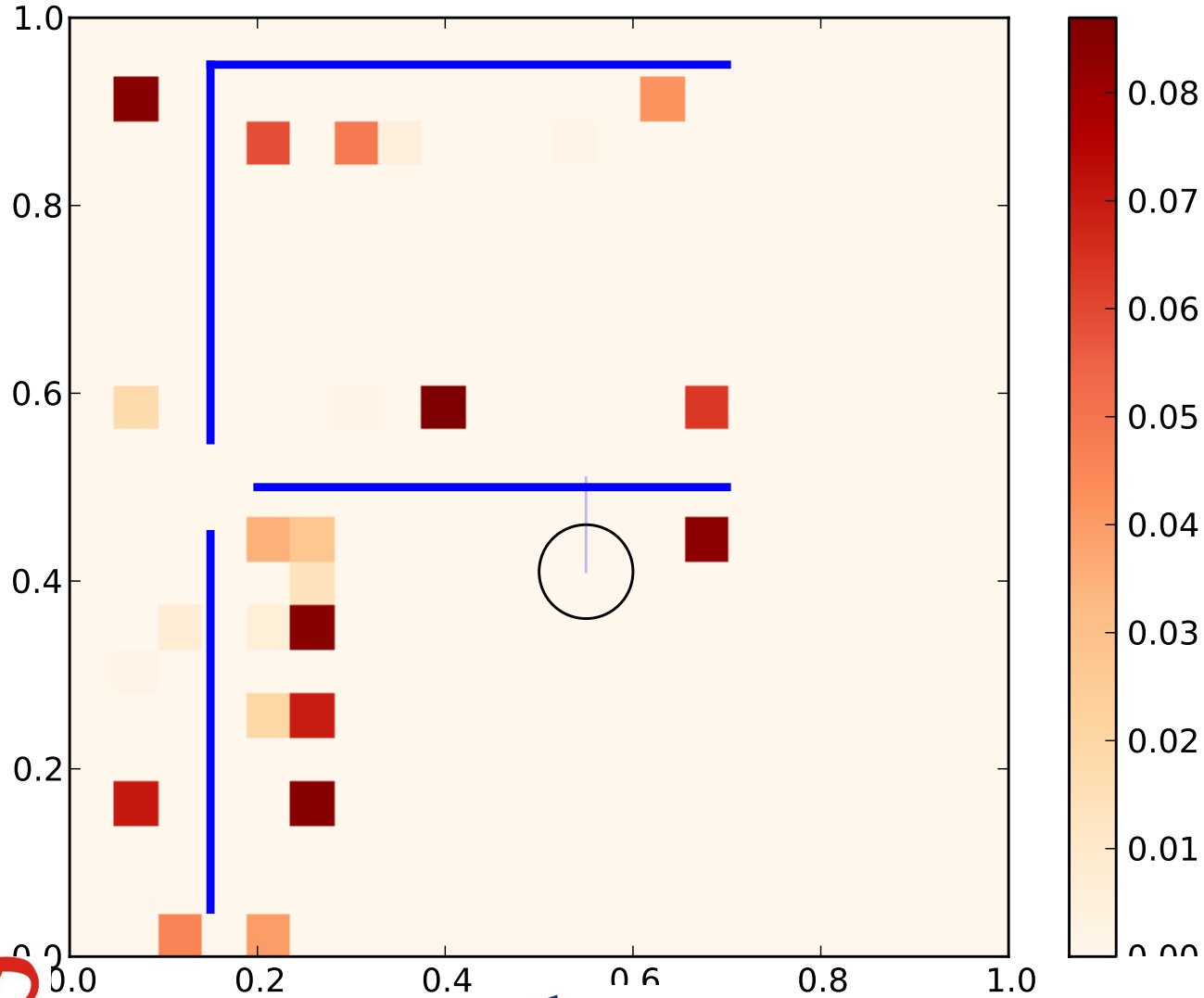
Example: with importance sampling



PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

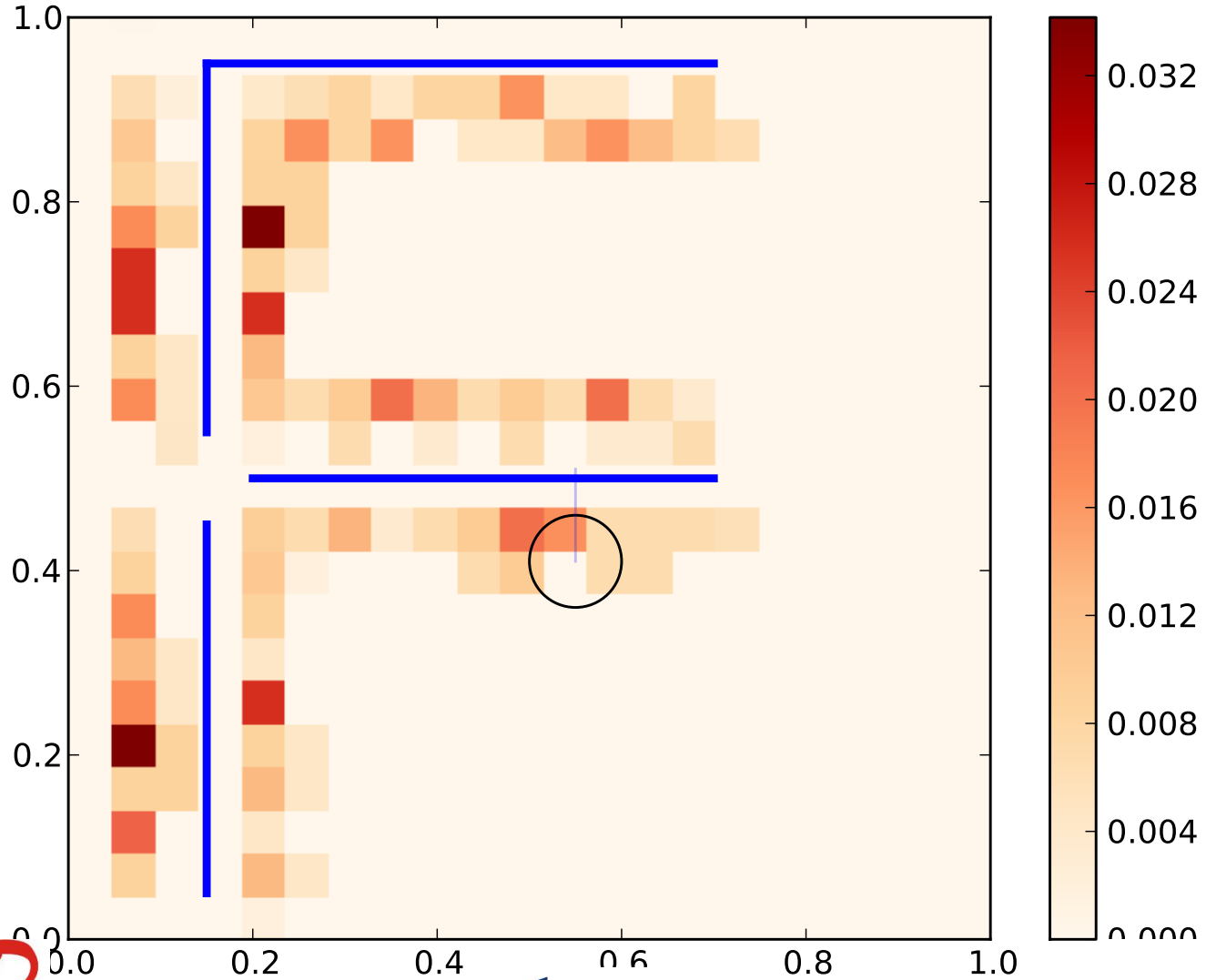
Example: initial position, no importance sampling



PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

Example: initial position, with importance sampling



PhD STUDIES

KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Project co-financed by European Union within the framework of European Social Fund

Particle filters – most popular variations

SIS (Sequential Importance Sampling) – importance sampling based version

SIR (Sequential Importance Resampling) – importance sampling based version with resampling

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund

Extentsions

- Problem specific methods of choosing proposal distributions
- 'Backward' estimation – correcting samples at previous time steps based on future observations
  - Auxiliary Particle Filter – we sample from the previous time step
  - resampling whole fragments of trajectories
- SLAM – Simultaneous Localization and Mapping
- Particle Swarm Optimization (PSO) – find the most probable trajectory by simulating a swarm of particles

PhD STUDIES
KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI
UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY
Project co-financed by European Union within the framework of European Social Fund