Ant Colony Optimized Importance Sampling: Principles, Applications and Challenges

Poul E. Heegaard
Department of Telematics
Norwegian University of Science and Technology (NTNU)

Werner Sandmann
Department Information Systems & Applied Computer Science
University of Bamberg

7th Int'l Workshop on Rare Event Simulation, Sept 24-26 2008, Rennes, France

UNIVERSITÄT BAMBERG

Simulation problem

- Strict QoS requirements need to be validated
  - Analytic models need (too) strict assumptions be be solved
  - Numerical solutions may suffer from state space explosion
  - Hard to simulate, e.g. loss ratio less than 10^{-7} takes forever
  - Importance sampling might work if parameters can be found

- This presentation
  - Model type handled
  - Speed-up simulation approach
  - **Swarm technique for adaptation of simulation parameters**
  - Numerical results
  - Inner workings of adaptive scheme

Model description

- D-dimensional discrete-state models (examples are Markovian)
- Finite or infinite restrictions in each dimension
- Transition affects at most two dimensions
- In paper described by **transition classes**

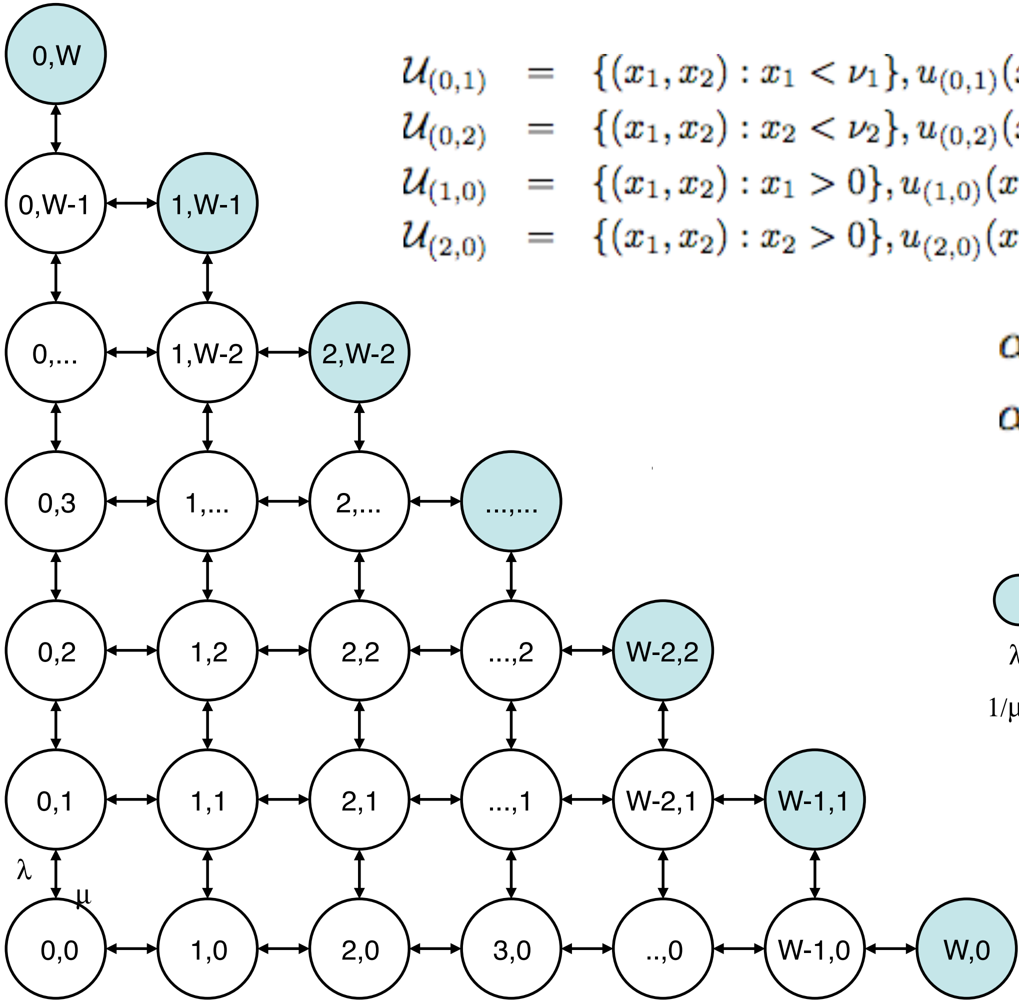
$$\tau = (\mathcal{U}, u, \alpha)$$

where

- Source state space $\mathcal{U} \subseteq \mathbb{N}^d$
- Destination state function $u : \mathcal{U} \to \mathbb{N}^d$
- Transition rate function $\alpha : \mathcal{U} \to \mathbb{R}$

- The model is applied in performance and dependability evaluation of Optical Packet Switched networks [other papers by the authors]

Model example



$$\mathcal{U}_{(0,1)} = \{(x_1,x_2) : x_1 < \nu_1\}, u_{(0,1)}(x_1,x_2) = (x_1+1, x_2),$$
$$\mathcal{U}_{(0,2)} = \{(x_1,x_2) : x_2 < \nu_2\}, u_{(0,2)}(x_1,x_2) = (x_1, x_2+1),$$
$$\mathcal{U}_{(1,0)} = \{(x_1,x_2) : x_1 > 0\}, u_{(1,0)}(x_1,x_2) = (x_1-1, x_2),$$
$$\mathcal{U}_{(2,0)} = \{(x_1,x_2) : x_2 > 0\}, u_{(2,0)}(x_1,x_2) = (x_1, x_2-1).$$

$$\alpha_{(0,i)}(x) = \lambda_i(x)$$
$$\alpha_{(i,0)}(x) = \mu_i(x)$$

State with loss

λ Packet arrival rate

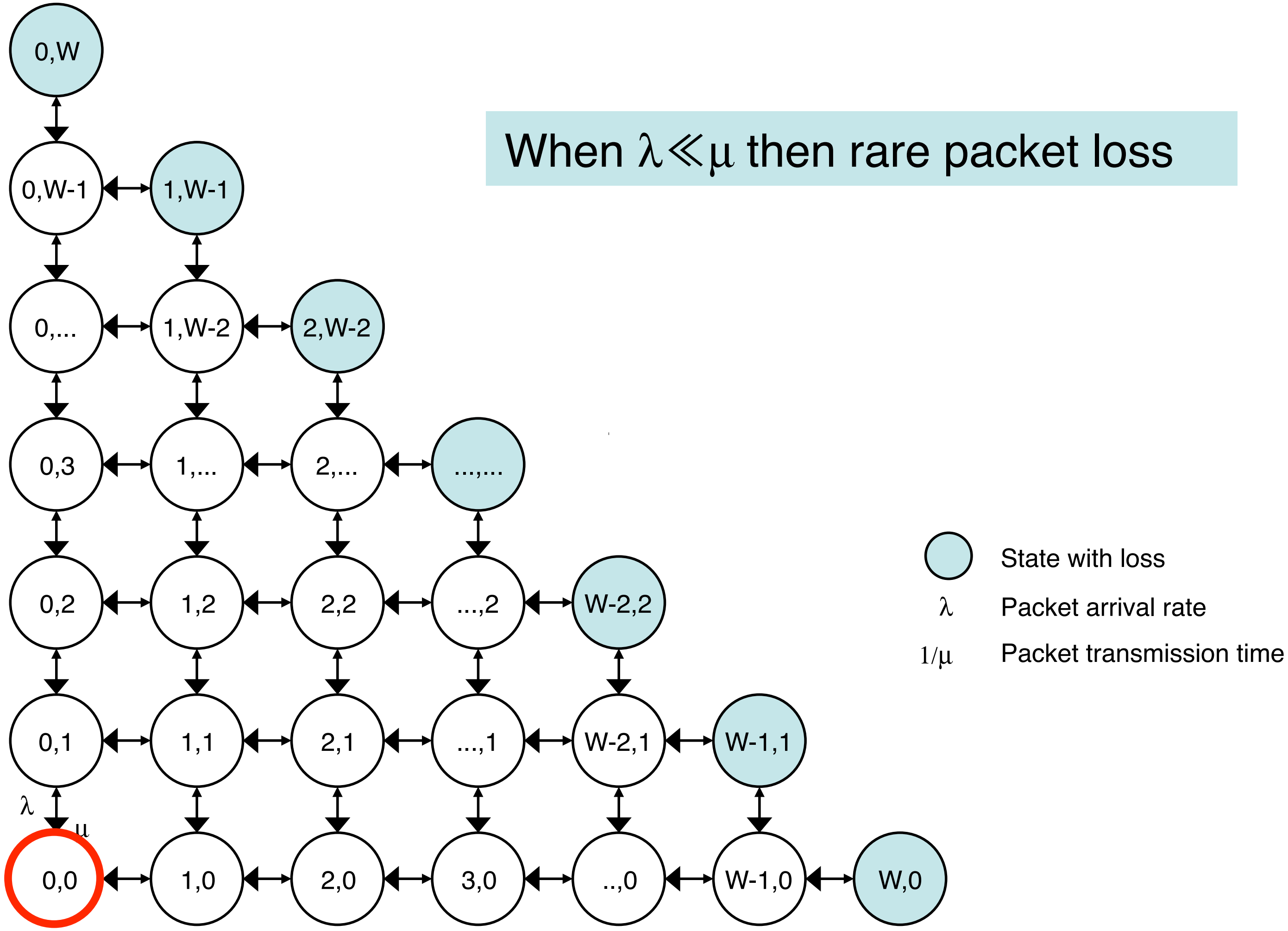
$1/\mu$ Packet transmission time

NTNU

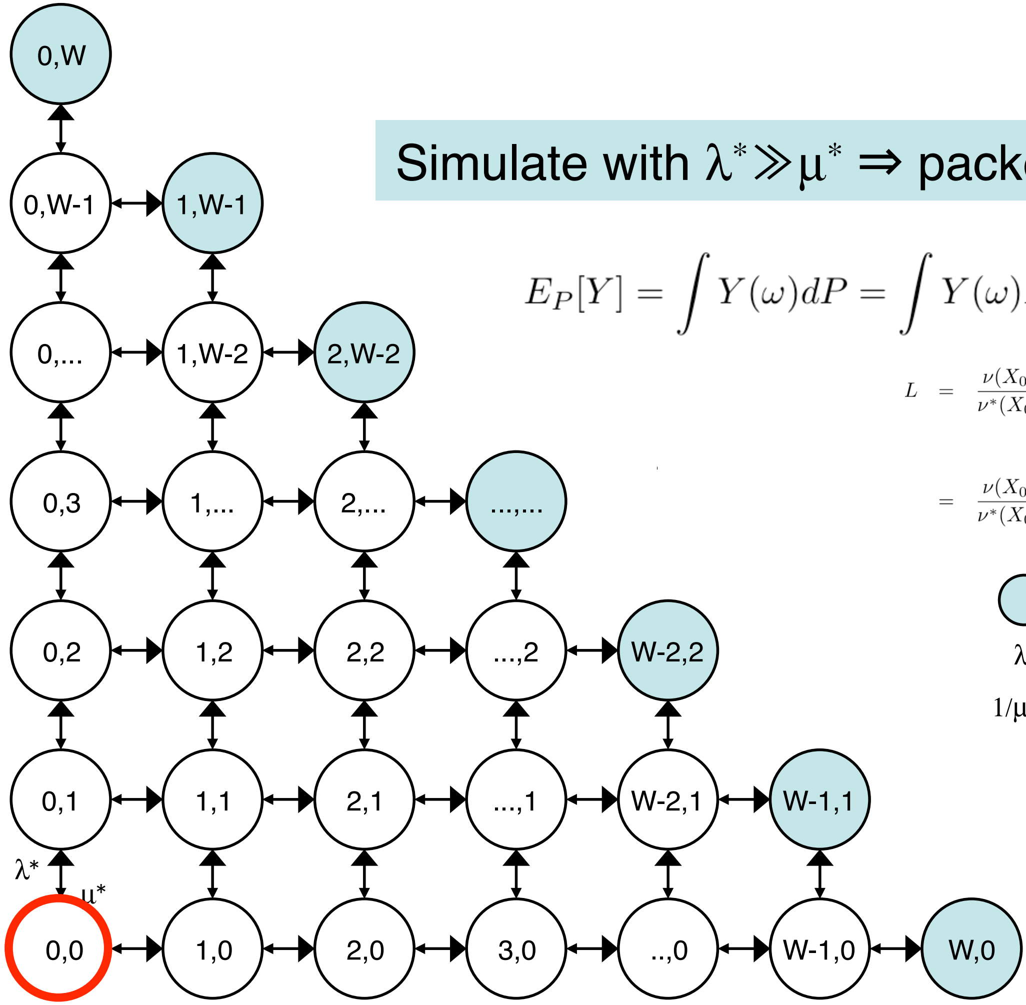
UNIVERSITÄT BAMBERG

Simulation problem

When $\lambda \ll \mu$ then rare packet loss



Simulation with importance sampling



Simulate with $\lambda^* \gg \mu^*$ ⇒ packet loss not rare

$$E_P[Y] = \int Y(\omega)dP = \int Y(\omega)L(\omega)dP^* = E_{P^*}[YL]$$

$$L = \frac{\nu(X_0)}{\nu^*(X_0)} \prod_{i=1}^{k} \frac{q_{X_{i-1}} \exp\left(-q_{X_{i-1}} \tau_{X_{i-1}}\right) p_{X_{i-1}x_i}}{q^*_{X_{i-1}} \exp\left(-q^*_{X_{i-1}} \tau_{X_{i-1}}\right) p^*_{X_{i-1}x_i}}$$

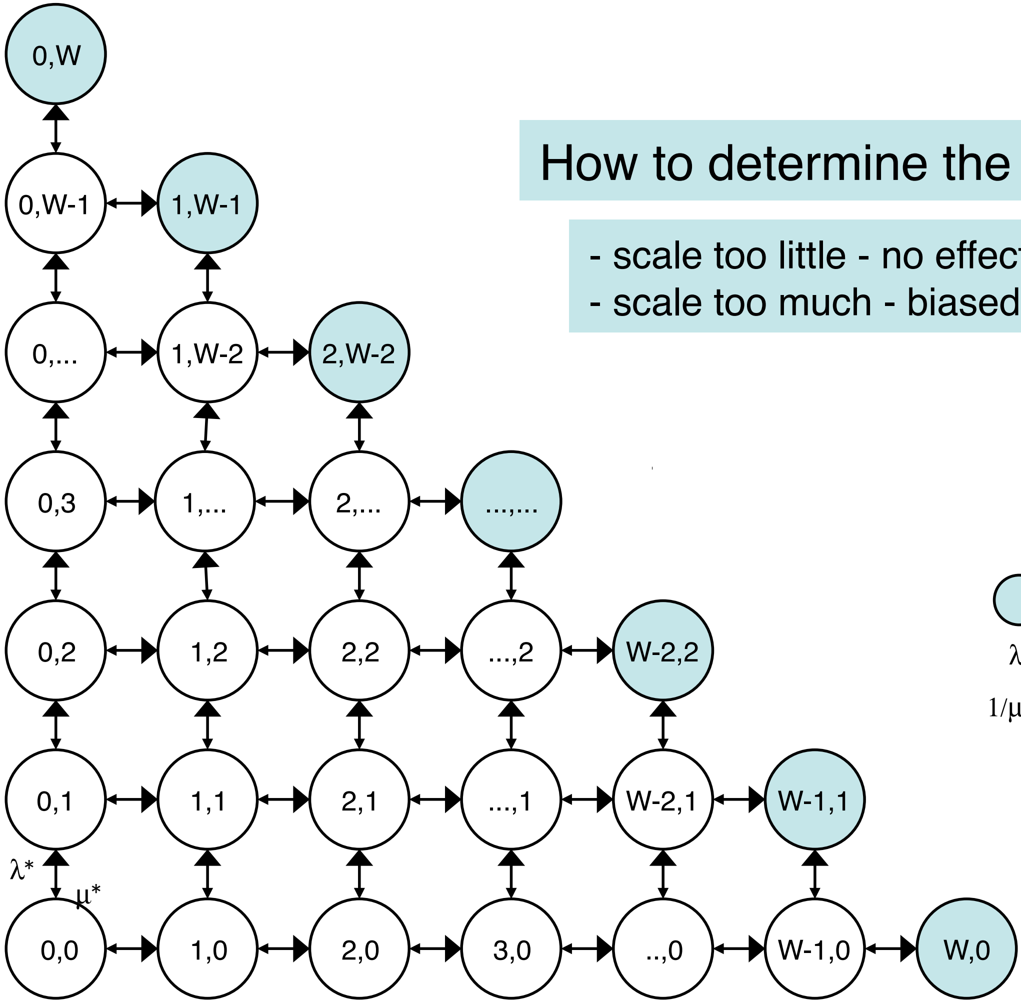
$$= \frac{\nu(X_0)}{\nu^*(X_0)} \prod_{i=1}^{k} \frac{q_{X_{i-1}x_i} \exp\left(-q_{X_{i-1}} \tau_{X_{i-1}}\right)}{q^*_{X_{i-1}x_i} \exp\left(-q^*_{X_{i-1}} \tau_{X_{i-1}}\right)},$$

State with loss

λ^* Packet arrival rate

$1/\mu^*$ Packet transmission time

The problem with importance sampling



How to determine the λ^* and μ^*?

- scale too little - no effect
- scale too much - biased estimates

State with loss

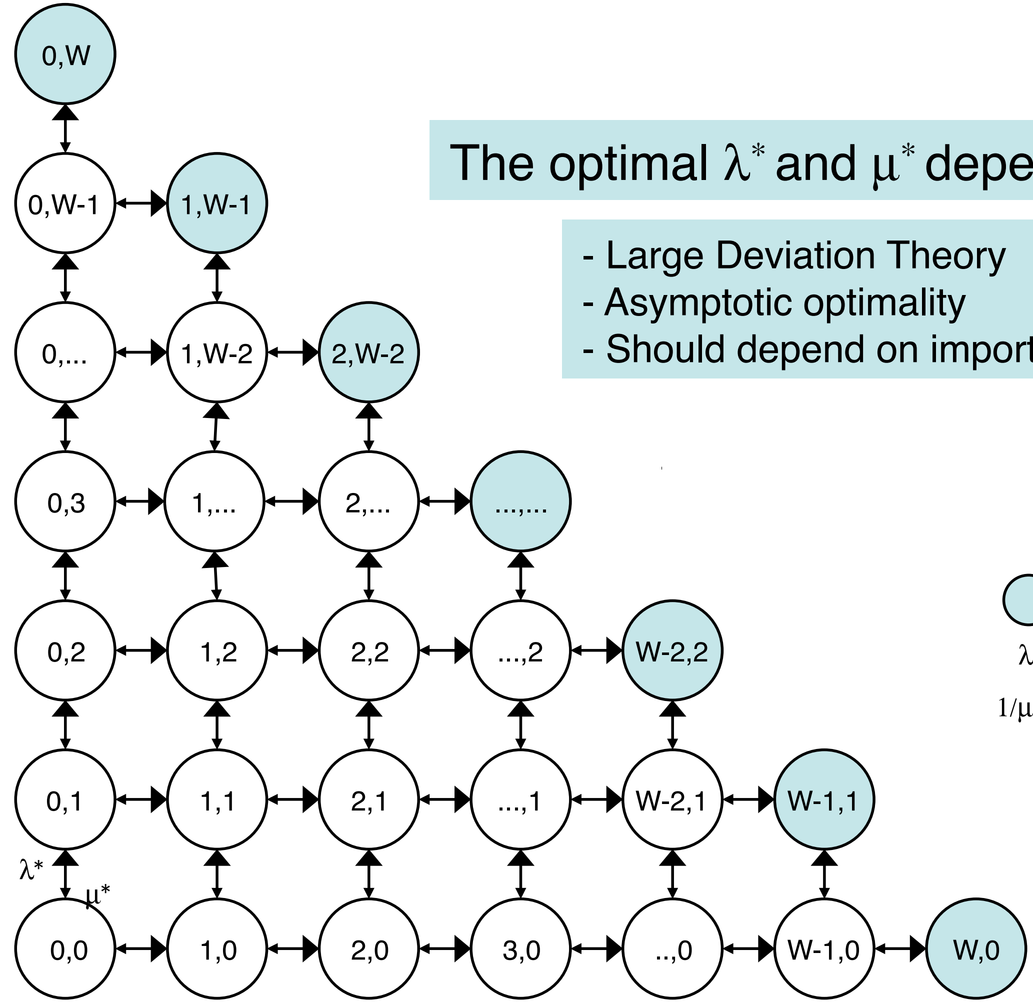
λ^* Packet arrival rate

$1/\mu^*$ Packet transmission time

Adaptive change of measure in IS

The optimal λ^* and μ^* depend on the state

- Large Deviation Theory
- Asymptotic optimality
- Should depend on importance of state

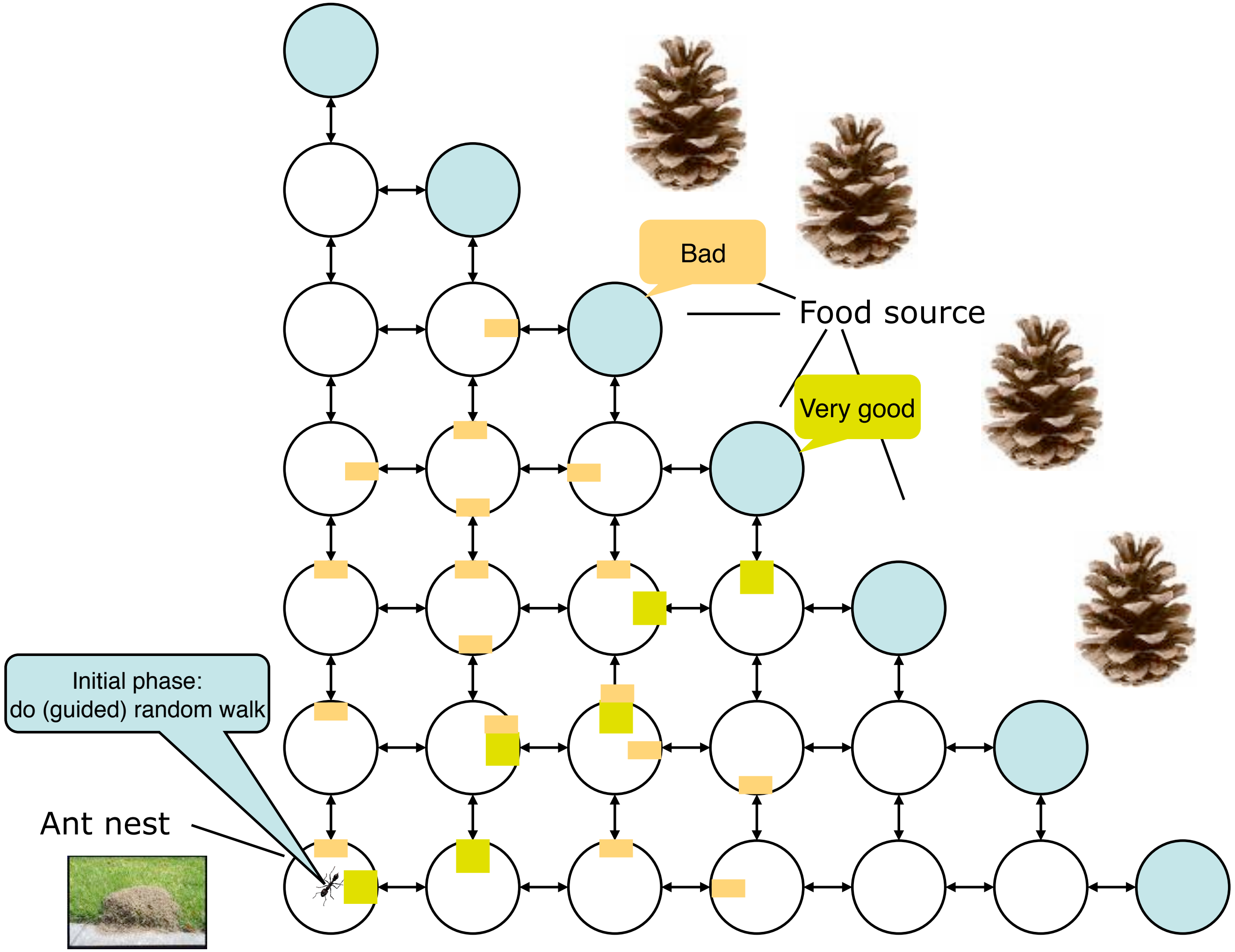


State with loss

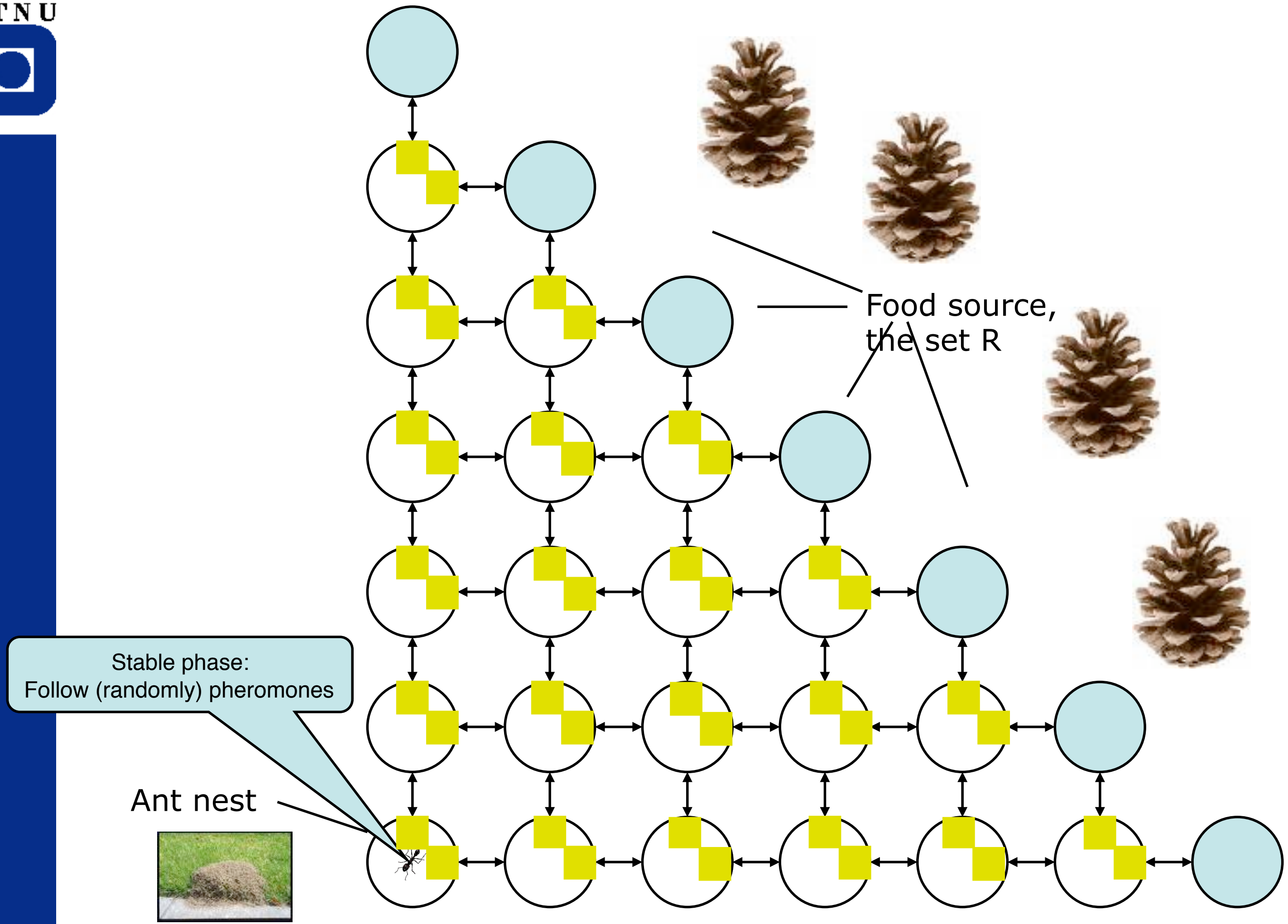
λ^* Packet arrival rate

$1/\mu^*$ Packet transmission time

Swarm intelligence (ex. ants)

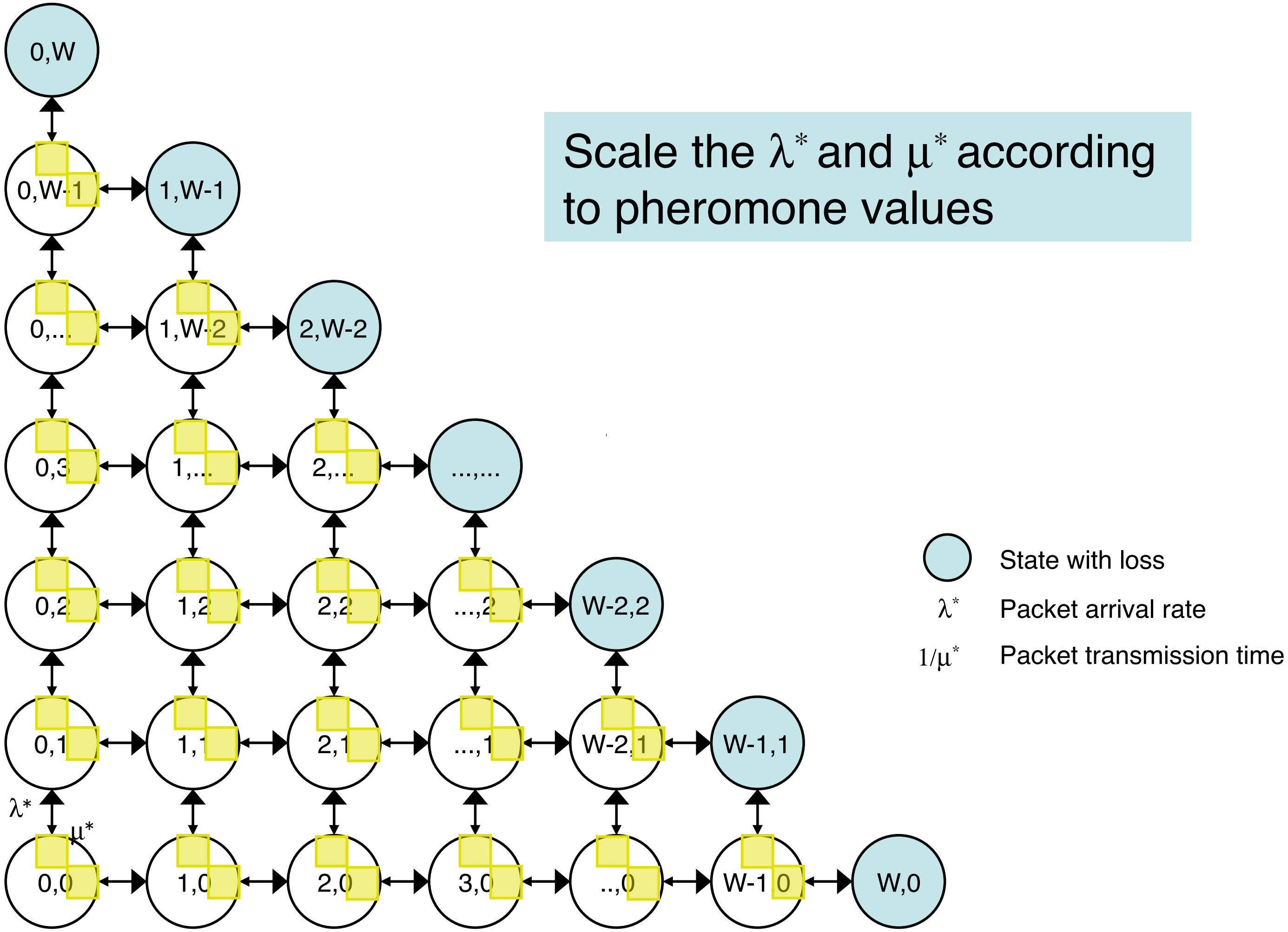


Swarm intelligence (ex. ants)



How ants guide IS parameters

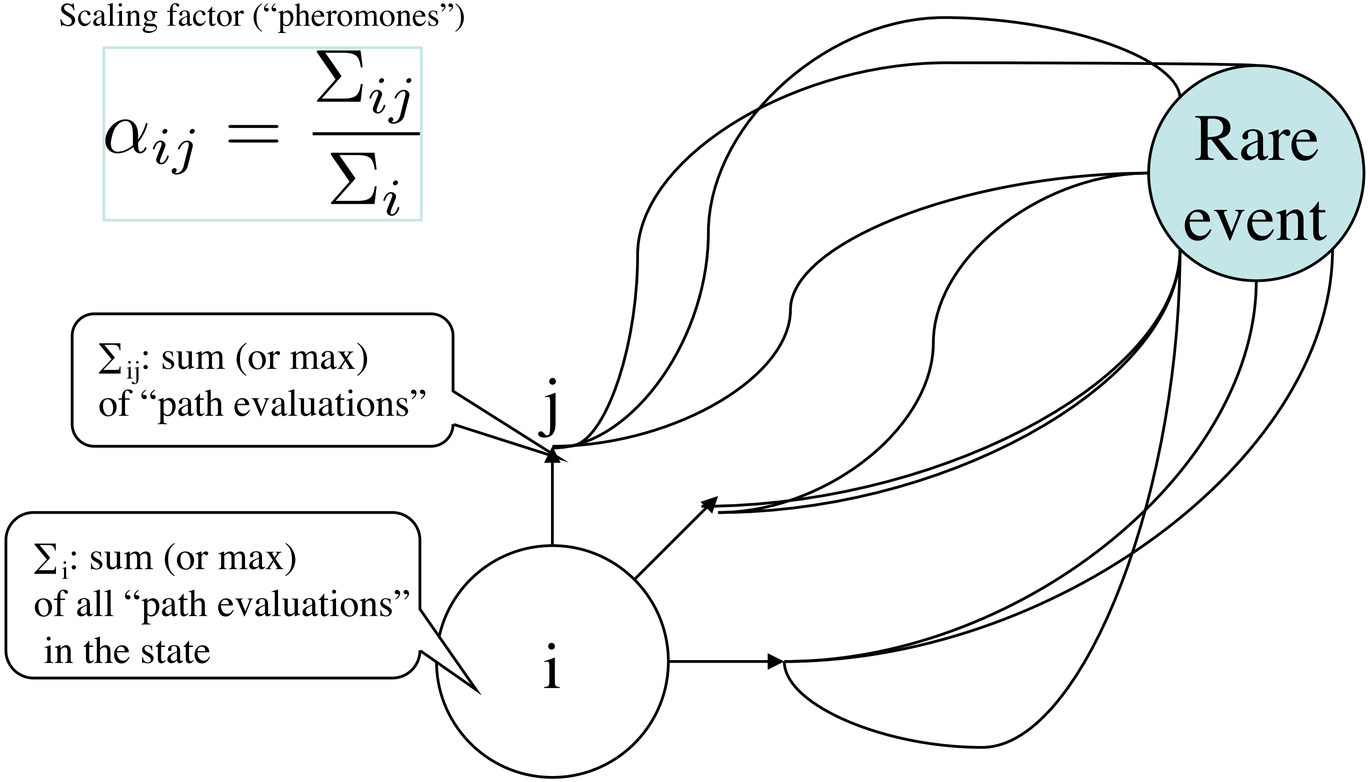
Scale the λ^* and μ^* according to pheromone values



How ants guide IS parameters

Scaling factor ("pheromones")

$$\alpha_{ij} = \frac{\Sigma_{ij}}{\Sigma_i}$$



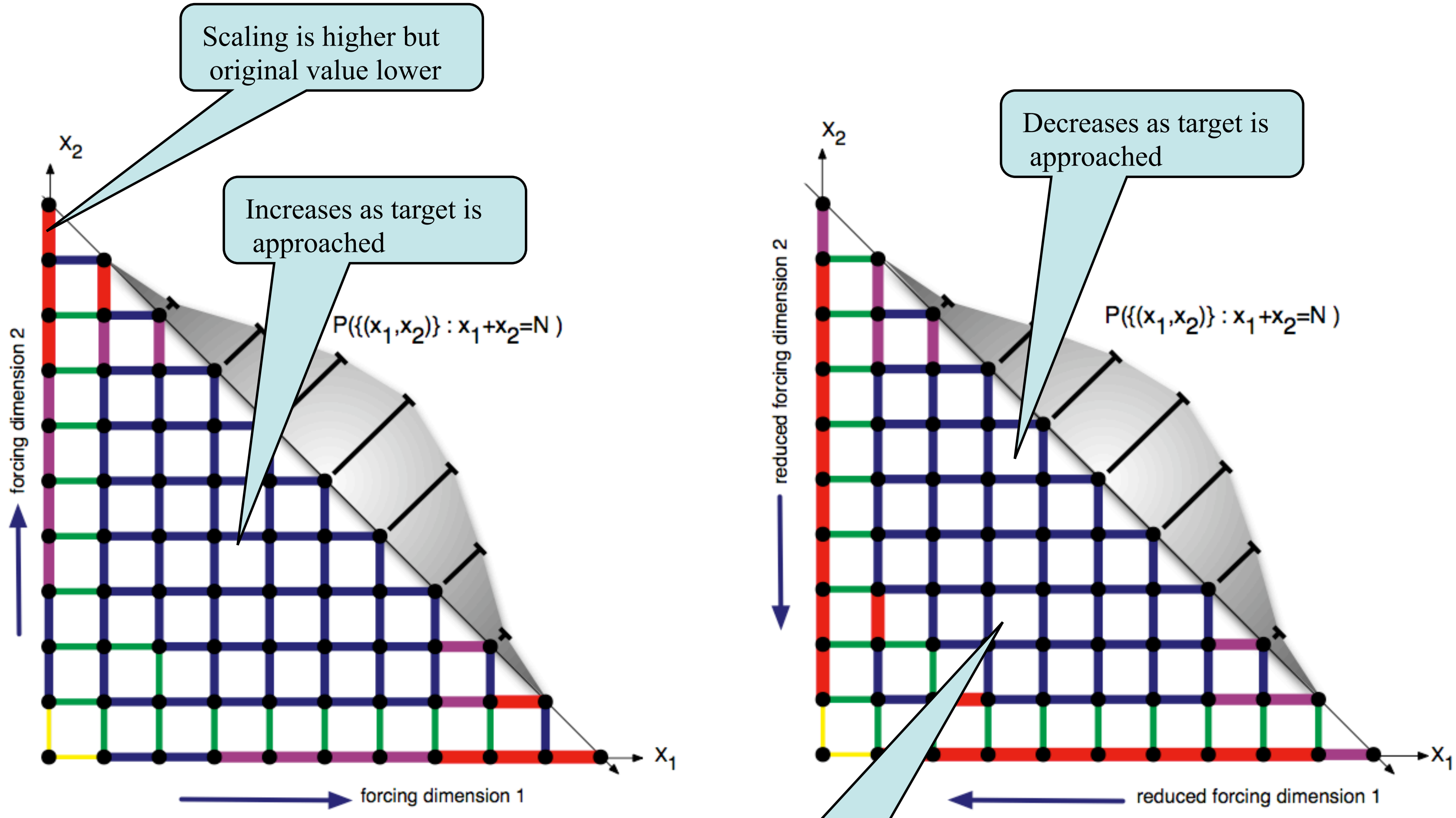
UNIVERSITÄT BAMBERG

How ants guide IS parameters

- ACO-IS approach

$$\begin{aligned} \lambda_{ij}^* &= \lambda_{ij} + \alpha_{ij}(\mu_{ji} - \lambda_{ij}) \\ \mu_{ji}^* &= \mu_{ji} + \alpha_{ij}(\lambda_{ij} - \mu_{ji}) \end{aligned}$$

Inner workings



Simulation cases

Table 1: Parameter and rare event state space

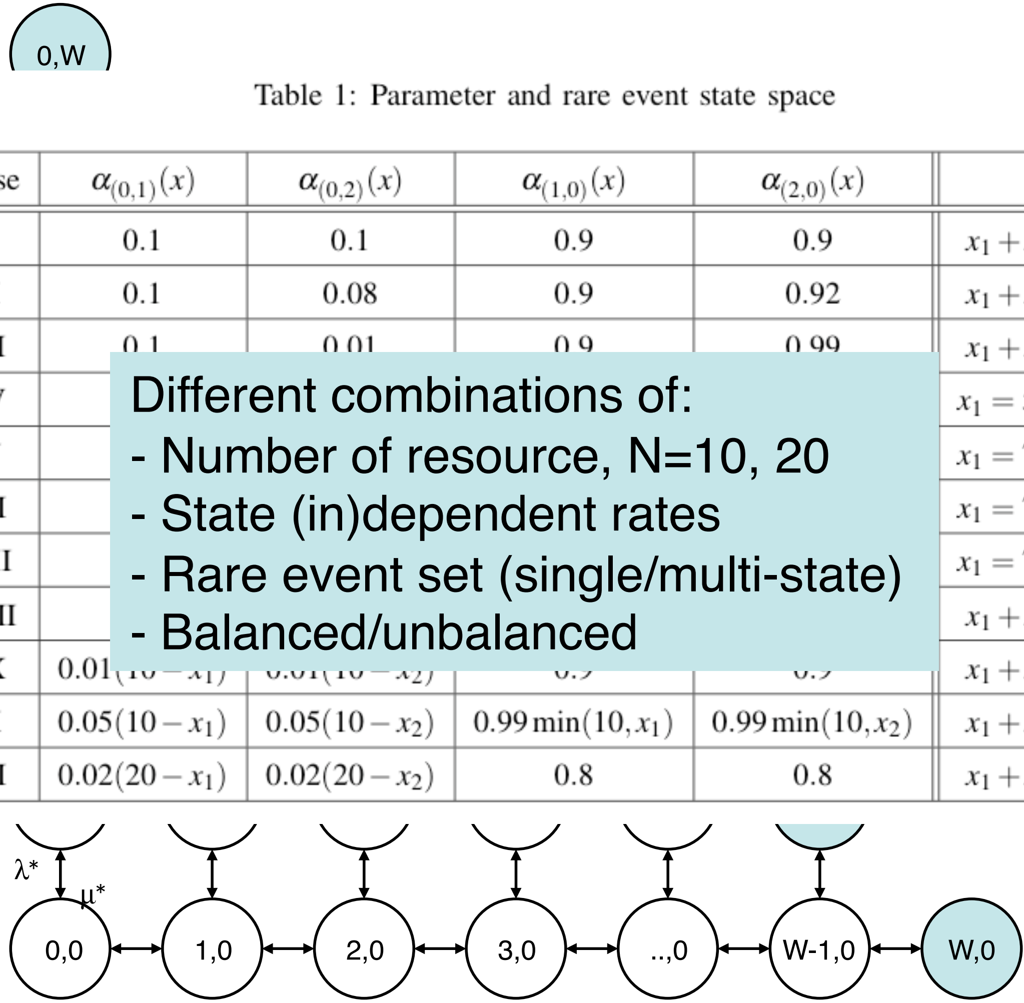
Case	$\alpha_{(0,1)}(x)$	$\alpha_{(0,2)}(x)$	$\alpha_{(1,0)}(x)$	$\alpha_{(2,0)}(x)$	$\mathscr{R}$
I	0.1	0.1	0.9	0.9	$x_1 + x_2 = 10$
II	0.1	0.08	0.9	0.92	$x_1 + x_2 = 10$
III	0.1	0.01	0.9	0.99	$x_1 + x_2 = 10$
IV	[illegible]	[illegible]	[illegible]	[illegible]	$x_1 = 5, x_2 = 5$
V	[illegible]	[illegible]	[illegible]	[illegible]	$x_1 = 7, x_2 = 3$
VI	[illegible]	[illegible]	[illegible]	[illegible]	$x_1 = 7, x_2 = 3$
VII	[illegible]	[illegible]	[illegible]	[illegible]	$x_1 = 7, x_2 = 3$
VIII	[illegible]	[illegible]	[illegible]	[illegible]	$x_1 + x_2 = 10$
IX	$0.01(10 - x_1)$	$0.01(10 - x_2)$	0.9	0.9	$x_1 + x_2 = 10$
X	$0.05(10 - x_1)$	$0.05(10 - x_2)$	$0.99 \min(10, x_1)$	$0.99 \min(10, x_2)$	$x_1 + x_2 = 10$
XI	$0.02(20 - x_1)$	$0.02(20 - x_2)$	0.8	0.8	$x_1 + x_2 = 20$

Different combinations of:
- Number of resource, N=10, 20
- State (in)dependent rates
- Rare event set (single/multi-state)
- Balanced/unbalanced

ith loss

arrival rate

transmission time



UNIVERSITÄT BAMBERG

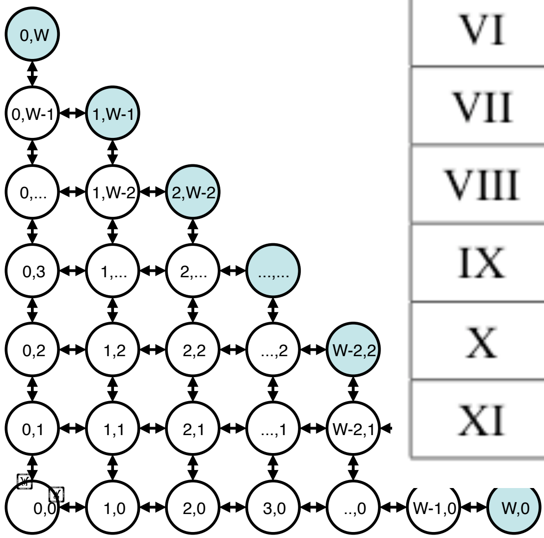
NTNU

Numerical results

Case	exact	$\bar{X}$	$S_{\bar{X}}/\bar{X}$
I	2.49×10^{-09}	2.52×10^{-09}	0.025
II	9.98×10^{-10}	9.97×10^{-10}	0.024
III	2.77×10^{-10}	2.81×10^{-10}	0.008
IV	2.27×10^{-10}	2.13×10^{-10}	0.052
V	2.27×10^{-10}	2.24×10^{-10}	0.045
VI	1.12×10^{-10}	1.10×10^{-10}	0.050
VII	1.24×10^{-10}	1.17×10^{-10}	0.045
VIII	8.84×10^{-10}	8.76×10^{-10}	0.044
IX	8.45×10^{-11}	8.80×10^{-11}	0.054
X	7.45×10^{-09}	7.41×10^{-09}	0.067
XI	6.31×10^{-09}	6.79×10^{-09}	0.062

High accuracy

Low relative error

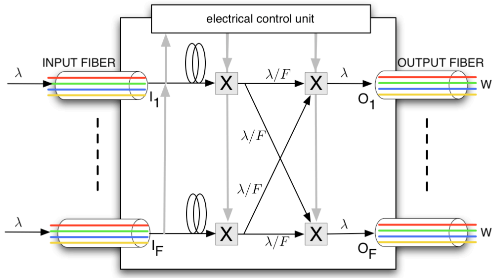


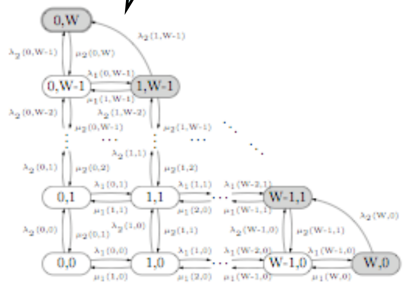
UNIVERSITÄT BAMBERG

High accuracy and low relative error observed for all cases

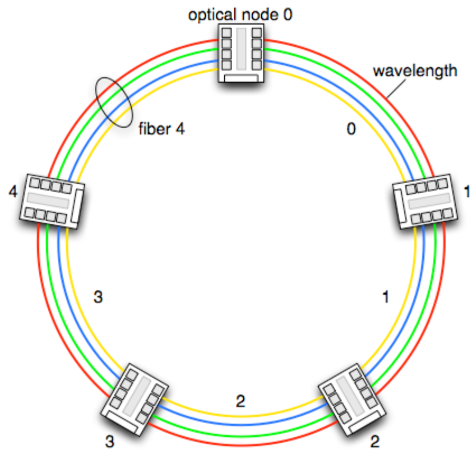
Other simulation cases

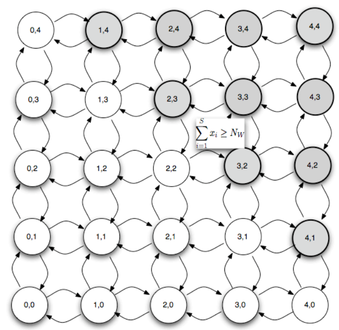
- Optical Burst Switch networks
  - Multiple service classes





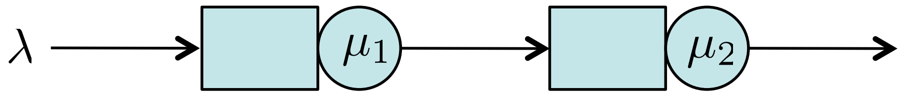
  - Multiple service classes and preemptive priorities
  - Node and link failures



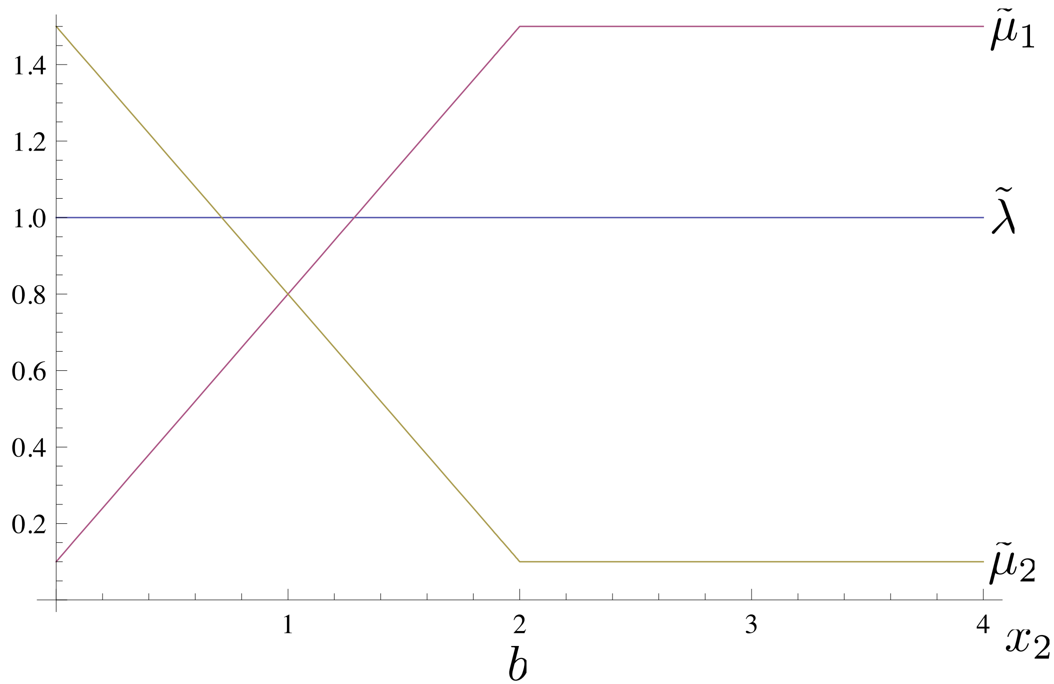


UNIVERSITÄT BAMBERG

Tandem queues



V.F. Nicola, T.S. Zaburnenko: *Importance Sampling Simulation of Population Overflow in Two-node Tandem Networks*. QEST 2005: 220-229



Tandem queues

- ACO-IS approach

$$
\begin{aligned}
\lambda_i^* &= \lambda_i + \alpha_{i,i+1}(\min(\mu_{1,i}, \mu_{2,i}) - \lambda_i) \\
\mu_{1,i}^* &= \mu_{1,i} + \alpha_{ii}(\max(\mu_{1,i}, \mu_{2,i}) - \mu_{1,i}) \\
\mu_{2,i}^* &= \mu_{2,i} + (\lambda_i + \mu_{1,i}) - (\lambda_i^* + \mu_{1,i}^*)
\end{aligned}
$$

Concluding remarks and further work

- Speed-up simulations
  - Importance sampling
  - Adaptive parameters by ACO meta heuristics
  - No a priori system knowledge required
- Promising results
- Simulated rare packet loss in OPN
  - Buffer-less
  - Multiple service classes
- Further work
  - Asymptotic behaviour
  - Non-exponential distribution
  - More complex system models
  - Detailed studies of the inner working of the Ants+IS methods

Challenges

- Initial phase: what are the consequences of biased sampling in initial phase?
- Inner workings of the ACO-IS? What is the result of ACO -IS biasing?
- Does it work for non-exponential distributions? Phase type distribution is the first to be checked?
- Other models structures? Tandem queue example is the next to be checked