Freie Universität Berlin

Proseminar Technische Informatik: Rare Event Simulation

Jan Sydow

Fachbereich Informatik

Institut für Mathematik/Informatik

Gliederung

- Motivation
- Methoden
  - Standardsimulation
  - Importance Sampling
  - Importance Splitting
- Beispiel

Freie Universität Berlin

Motivation

- Seltenes Ereignis
  - Auftrittswahrscheinlichkeit $<10^{-9}$
  - Möchte genaue Aussagen treffen
    - → bestimmte Anzahl von Auftritten nötig
    - → enormer Simulationsaufwand

Freie Universität Berlin

Beispiele

- Ruin in der Wirtschaftsmathematik
- Erdbeben in der Geologie
- Chemische Reaktionen
- Atomunfall (SuperGAU)
- Aussterben von Spezien in der Biologie

- Pufferüberläufe
- System-/Komponentenausfall
- Zellverlust in ATM-Netzwerken

Gliederung

- Motivation
- Methoden
  - Standardsimulation
  - Importance Sampling
  - Importance Splitting
- Beispiel

Standardsimulation

- γ Wahrscheinlichkeit des seltenen Ereignisses *B*
- *n* Simulationsläufe $Z_1, \ldots, Z_n$
- Schätzwert für γ ist dann:

$$\hat{\gamma} = \frac{\sum_{i=1}^{n} I(Z_i)}{n}$$

- Konfidenzintervall: $\left[\hat{\gamma} - z(1-\tfrac{\alpha}{2})\frac{\sqrt{\hat{\gamma}(1-\hat{\gamma})}}{\sqrt{n}} ; \hat{\gamma} + z(1-\tfrac{\alpha}{2})\frac{\sqrt{\hat{\gamma}(1-\hat{\gamma})}}{\sqrt{n}}\right]$
  - bei gegebener Genauigkeit: wenn $\gamma \to 0$, dann $n \to \infty$

Gliederung

- Motivation
- Methoden
  - Standardsimulation
  - Importance Sampling
  - Importance Splitting
- Beispiel

Importance Sampling

- Neues Wahrscheinlichkeitsmaß p�much einführen
- Ziel: geringere Varianz
  - Weniger Simulationsergebnisse für gleiche Genauigkeit
- Methode verfälscht Ergebnis
- Ausgleichsfaktor $L_{p'} = {}^{p}/_{p'}$ (*likelihood ratio*)
- Schätzwert aus:

$$\hat{\hat{\gamma}} = \frac{\sum_{i=1}^{n} I(Z_i)}{n} \cdot L_{p'}$$

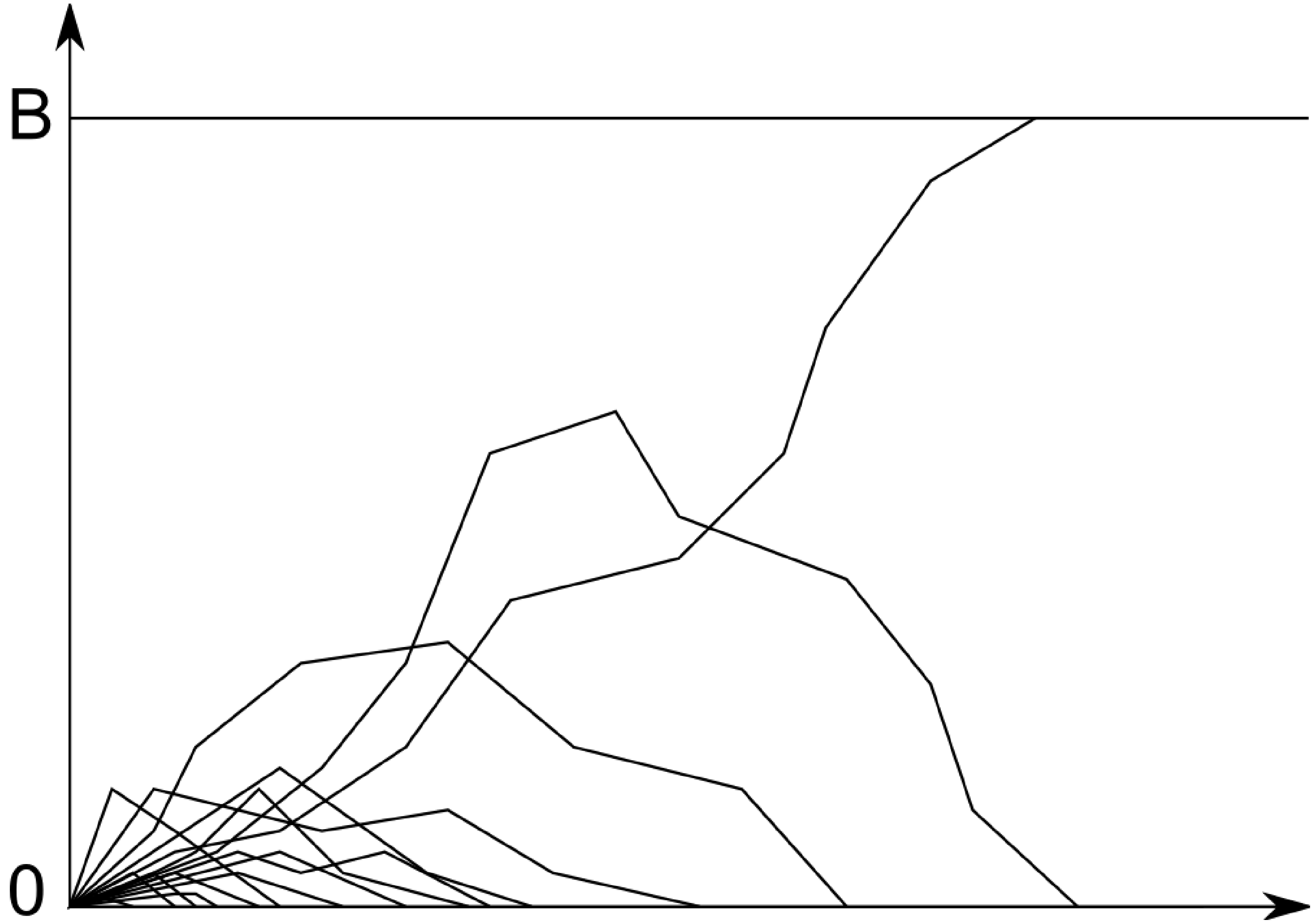
Varianzreduktion

- Zero-Variance
  - konstante Ergebnisse
  - $L_{p`} = \gamma$
  - Selten implementierbar, γ bekannt sein muss
- Bounded Relative Error (BRE)
  - System hat Eigenschaft, wenn gilt: $\limsup_{\gamma \to 0} \frac{\sigma_{p'}(Z_b)}{\gamma} \leq \infty$
  - *n* begrenzt, wenn $\gamma \to 0$
- Asymptotisch optimal
  - Schwächer als BRE
  - Für das System muss gelten: $\frac{\log(\mathrm{E}_{p'}(Z^2))}{\log \gamma} \leq 2$

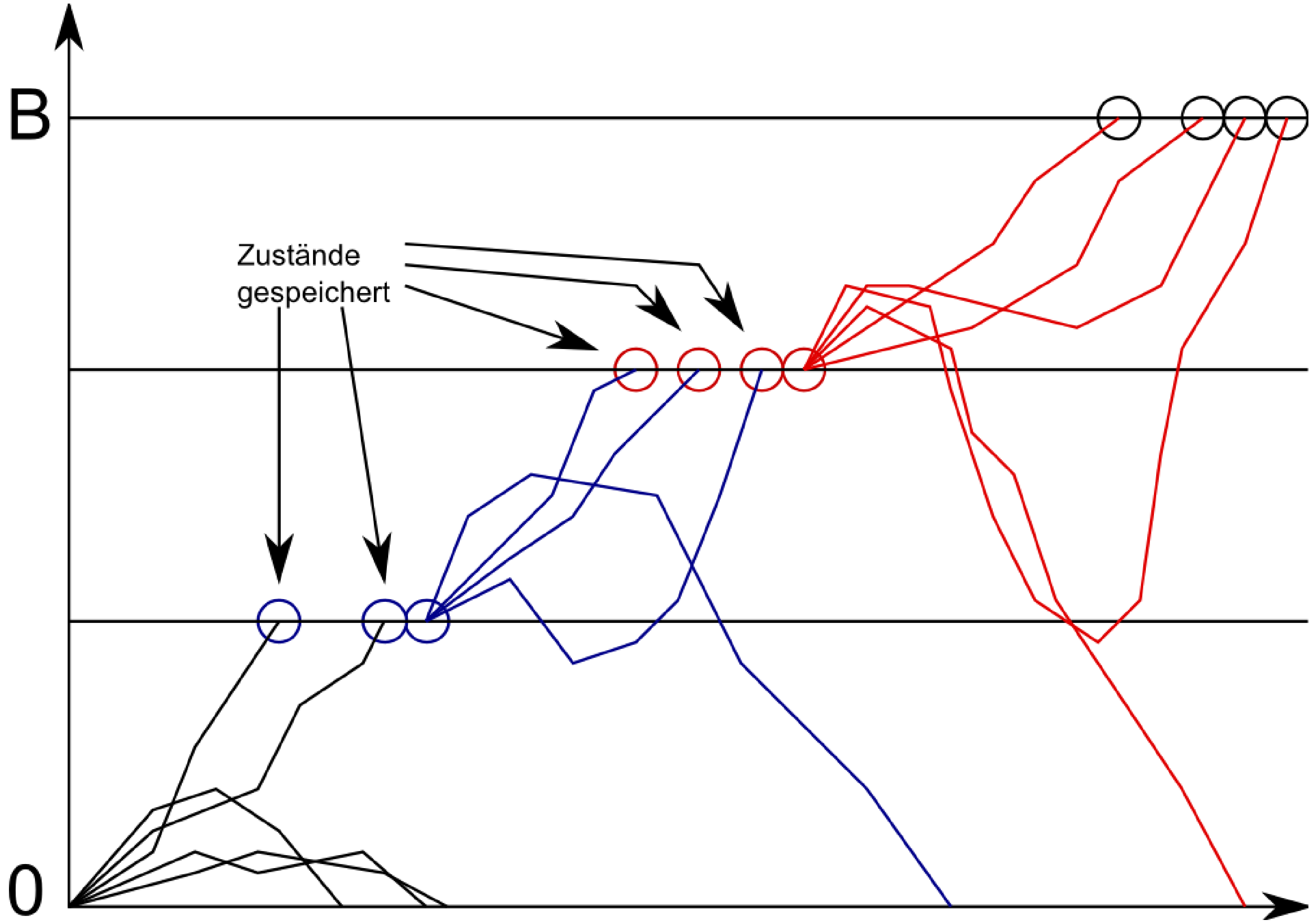
Gliederung

- Motivation
- Methoden
  - Standardsimulation
  - Importance Sampling
  - Importance Splitting
- Beispiel

Importance Splitting - Motivation



Importance Splitting - Idee

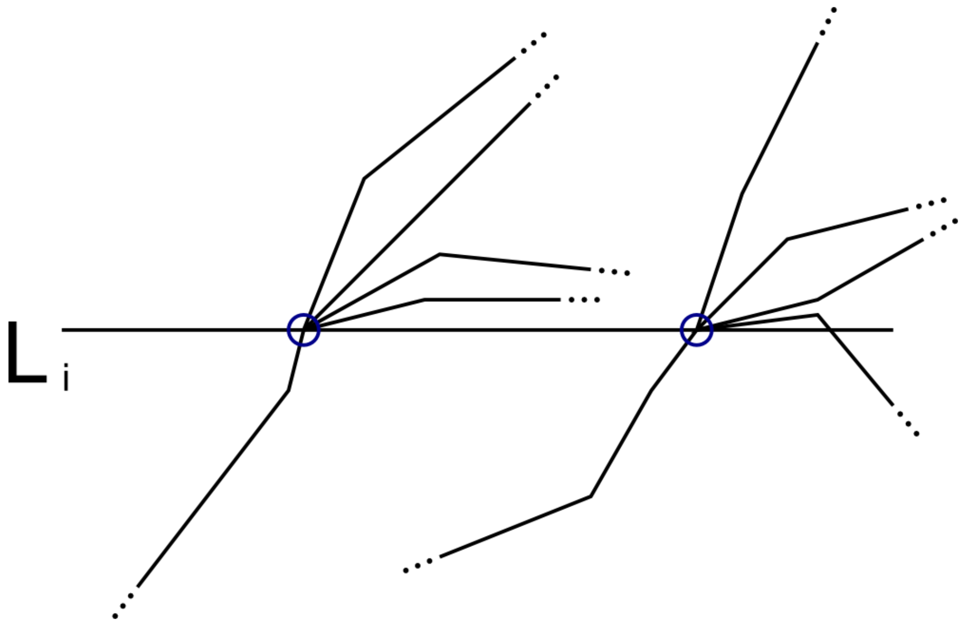


Importance Splitting

- Stufen $L_0 \supset L_1 \supset \ldots \supset L_n = B$ einführen (B: seltenes Ereignis)
- Wahrscheinlichkeit D_i, dass Pfad L_i vor Startzustand erreicht
- Stufenwahrscheinlichkeiten p_k bestimmen durch Simulation
  - $p_1 = P(D_1)$
  - $p_k = P(D_k | D_{k-1})$
- Wahrscheinlichkeit $\gamma = p_1 \, p_2 \, \ldots \, p_n$

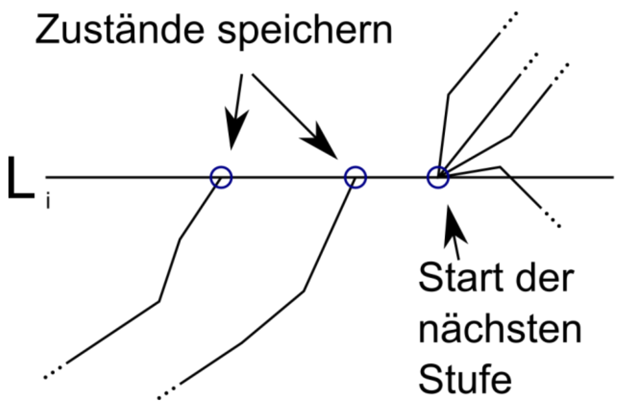
Implementierung

- Fixed Splitting
  - Jeden Pfad r mal aufsplitten
  - → Aussterben oder Explodieren der Pfade möglich
  - Als Tiefendurchlauf implementierbar (*Global Step approach*)
    - → weniger Speicherverbrauch



Implementierung

- Fixed Effort
  - r Pfade insgesamt pro Stufe
  - Verhindert Aussterben/Explodieren
  - Gespeicherte Pfade werden auf r Pfade verteilt
    - Random Assignment
    - Fixed Assignment
  - Nur Breitendurchlauf möglich (*Single Step approach*)
    - → höherer Speicherbedarf

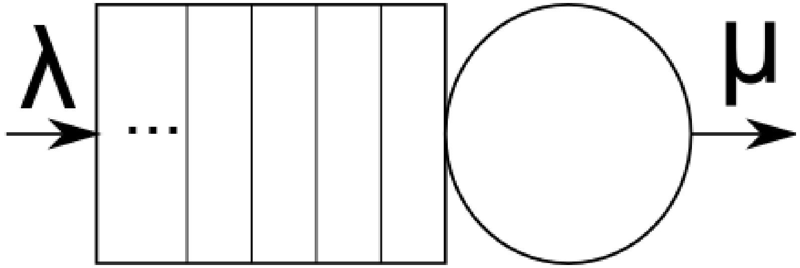


Gliederung

- Motivation
- Methoden
  - Standardsimulation
  - Importance Sampling
  - Importance Splitting
- Beispiel

Beispiel

- M/M/1 Queue
  - Ankunftsrate $\lambda=0{,}5$
  - Bedienungsrate $\mu=1$



M/M/1 Queue

- seltenes Ereignis: Warteschlange hat ≥k Kunden
- Gesucht:

$$\gamma = \Pr\{N \geq k\} = \rho^k = (\lambda/\mu)^k$$

- Gefordert: Konfidenzintervallweite ≤5% von γ

Beispiel

k	Konfidenz-niveau	γ	n
10 10	95% 98%	$9{,}7656 \cdot 10^{-4}$	2.508 2.982
20 20	95% 98%	$9{,}5367 \cdot 10^{-7}$	81.006 96.298
50 50	95% 98%	$8{,}8818 \cdot 10^{-16}$	263.066.513 312.727.028
100 100	95% 98%	$7{,}8886 \cdot 10^{-31}$	$8{,}827 \cdot 10^{15}$ $1{,}0493 \cdot 10^{16}$

Freie Universität Berlin

Beispiel

- Anwendung von Importance Sampling
  - λ verändern und damit ρ
  - Parameter: k=50 und Konfidenzniveau 95%

ρ`	**γ`**	**n**	$L_{\rho`}$
0,5	$8,8818 \cdot 10^{-16}$	263.066.513	1
0,7	$1,7985 \cdot 10^{-8}$	584.602	$4,9384 \cdot 10^{-8}$
0,9	$5,1538 \cdot 10^{-3}$	1.089	$1,7233 \cdot 10^{-13}$

Zusammenfassung

- Seltene Ereignisse reell nicht simulierbar
- Andere Simulationsmethoden werden benötigt
- Importance Sampling
  - System wird verändert, sodass weniger simuliert werden muss
  - Veränderung muss gegengerechnet werden
- Importance Splitting
  - Vielversprechende Pfade werden weiterverfolgt
  - Stufen werden eingeführt

Vielen Dank!

Fragen?