STOCHASTIC PETRI NETS FOR DISCRETE-EVENT SIMULATION

Peter J. Haas

IBM Almaden Research Center
San Jose, CA

Petri Nets 2007
Part I

Introduction
My Background

- Mid 1980’s: PhD student studying discrete-event simulation
  - Under Donald Iglehart (Stanford) & Gerald Shedler (IBM)
  - “Performance analysis using stochastic Petri nets”
- Wrote PNPM85 simulation paper with Gerry Shedler
  - “Regenerative simulation of stochastic Petri nets”
- Kept working (in between Info. Mgmt. research) . . .
  - Modelling power for simulation [HS88]
  - Prototypes: SPSIM, “Labelled” SPN simulator [JS89, HS90]
  - Delays [HS93a,b]
  - Standardized time series [Haa97,99a,99b]
  - Transience and recurrence [GH06, GH07]
- Gave this seminar at 2004 Winter Simulation Conference
Complex Systems

Concurrency

Synchronization

Precedence

Priority

Randomness
(non-Markovian stochastics)
Simulation and SPNs

- Assessment of system performance is difficult
  - Even modelling the system is hard!
  - Model is usually analytically and numerically intractable
    - Huge state space and/or non-Markovian
  - Simulation is often the only available numerical method
    - But can’t simulate blindly

- SPNs can help
  - An attractive graphically-oriented modelling framework
  - Well suited to sample-path generation on a computer
  - Solid mathematical foundation
Simulation theory for SPNs

- SPNs as a modelling framework for discrete-event systems
- Sample path generation for SPNs
- Steady-state output analysis: theory and methods
Sources for This Tutorial


Outline

- Simulation basics
  - Discrete-event systems
  - The simulation process
- Modelling with SPNs
  - Building blocks
  - Modeling power for simulation
- Sample-path generation
  - The marking process
  - Efficiency issues, parallelism
- Steady-state estimation for SPNs
  - Conditions for long-run stability (recurrence, limit theorems)
  - Output-analysis methods and their validity
Goals

- Illustrate the rich behavior of non-Markovian SPNs
- Introduce you to some basic simulation methodology
- Explore foundational issues in modelling and analysis
- Connect modeling practice and simulation theory
- Stimulate your interest in SPNs as a simulation framework
Part II

Simulation Basics
What We Simulate: Discrete-Event Stochastic Systems

- System changes state when events occur
  - Stochastic changes at random times
- Underlying stochastic process \( \{ X(t): t \geq 0 \} \)
  - \( X(t) \) = state of system at time \( t \) (a random variable)
  - Piecewise-constant sample paths
  - Typically not a continuous-time Markov chain
- Modelling challenge: defining appropriate system state
  - Compact for efficiency reasons
  - Enough info to compute performance measures
  - Enough info to determine evolution
Why We Simulate: Performance Evaluation

- Steady-state performance measures
  - Time-average limits:
    \[ \alpha = \lim_{t \to \infty} \frac{1}{t} \int_{0}^{t} f(X(u)) \, du \]
  - Steady-state means:
    \[ \alpha = E[f(X)], \text{ where } X(t) \Rightarrow X \]
    - I.e., \( P_\mu \{ X(t) = s \} \rightarrow P \{ X = s \} \) as \( t \to \infty \)
  - Want point estimate \( \hat{\alpha}(t) \)
    - Unbiased: \( E_\mu [\hat{\alpha}(t)] = \alpha \)
    - Strongly consistent: \( P_\mu \{ \lim_{t \to \infty} \hat{\alpha}(t) = \alpha \} = 1 \)
  - Want asymptotic 100\% confidence interval
    - \( I(t) = [\hat{\alpha}(t) - H(t), \hat{\alpha}(t) + H(t)] \)
    - \( P_\mu \{ I(t) \ni \alpha \} \approx p \) for large \( t \)
    - CI width indicates precision of point estimate
Challenges in Performance Evaluation

- Is steady-state quantity $\alpha$ well-defined?
  - Ex: steady-state number in $M/M/1$ queue with $\rho > 1$

- Is steady-state quantity independent of startup condition $\mu$?
  - Ex: reducible Markov chain

- Statistical challenges
  - Autocorrelation problem
  - Initialization bias problem

- How to handle Delays?
  \[
  \lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} f(D_j)
  \]
The Simulation Process

Real-world system

- Decide
  - Expected idle time
  - Long-run avg. cost
  - Long-run throughput
  - Expected delay

- Model
  - Robot arm
  - Conveyor
  - Exp(u) arrival rate

Estimates

- Analyze

Simulation model

- Specify
  - \{X(t) : t \geq 0\}
  - \{S_n : n \geq 0\}
  - \{D_n : n \geq 0\}

Sample paths

- Program & Execute

Stochastic process(es)
How Modelling Frameworks Can Help

- But challenges, also:
  - Immediate transitions and markings
Part III

Modelling with SPNs
The SPN Graph

$D = \text{finite set of places}$

$E = \text{finite set of transitions (timed and immediate)}$

marking = assignment of token counts to places

$s = (2, 1, 1)$
The marking changes when an enabled transition fires.

Removes 1 token per place from random subset of input places.

Deposits 1 token per place in random subset of output places.
Clocks (Event Scheduling)

- One clock per transition: records remaining time until firing
- Enabled transitions compete to trigger marking change
  - The clock that runs down to 0 first is the “winner”
  - Can have simultaneous transition firing: $p(s'; s, E^*)$
  - Numerical priorities: specify simultaneous-firing behavior
- At a marking change: three kinds of transitions
  - New transitions: Use clock-setting distribution function
  - Old transitions: Clocks continue to run down
  - Newly-disabled transitions: Clock readings are discarded
Clocks, Continued

- Clock-setting distribution depends on:
  - Old marking, new marking, trigger set
- Clocks run down at marking-dependent speeds $r(s, e)$
  - Processor sharing
  - Zero speeds: preempt-resume behavior
Timed and Immediate Markings

- **Immediate marking**: \( \geq 1 \) immediate transition is enabled
- An immediate marking vanishes as soon as it is attained
- Otherwise, marking is timed
Example: Cyclic Queues with Feedback

position: 1 2 3 4 5
Bottom-Up and Top-Down Modeling
Other Modeling Features

Concurrency:

Synchronization:

Precedence:

Priority:
Why This SPN Model?

- **Conciseness:** small set of building blocks
- **Generality:** subsumes GSPNs, etc.
  - Theory carries over
- **Modelling power:** captures many discrete-event systems
Modeling Power of SPNs

- Compare to Generalized semi-Markov processes (GSMPs)
  - Arbitrary state definition \((s)\)
  - Set \(E(s)\) of active events is a building block
  - No restrictions on \(p(s'; s, E^*)\)
  - No “immediate events”

- Strong mimicry
  - Define \(X(t)\) = state of system at time \(t\)
  - Define \((S_n, C_n)\) = (state, clocks) after \(n\)th state transition
  - \(\{ X(t) : t \geq 0 \}\) processes have same dist’n (under mapping)
  - \(\{ (S_n, C_n) : n \geq 0 \}\) have same dist’n (under mapping)

- Theorem: SPNs and GSMPs have same modeling power
  - Establishes SPNs as framework for discrete-event simulation
  - Allows application of GSMP theory to SPNs
  - Methodology allows other comparisons (e.g., inhibitor arcs)
Part IV

Sample-Path Generation
The Marking Process

- **Marking process**: \( \{ X(t) : t \geq 0 \} \)
  - \( X(t) = \) the marking at time \( t \)
  - A very complicated process

- **Defined in terms of Markov chain** \( \{(S_n, C_n) : n \geq 0\} \)
  - System observed after the \( n \)th marking change
  - \( S_n = (S_{n,1}, \ldots, S_{n,L}) = \) the marking
  - \( C_n = (C_{n,1}, \ldots, C_{n,M}) = \) the clock-reading vector
  - Chain defined via SPN building blocks
Definition of the Marking Process

\[ X(t) = S_{N(t)} \]
Generation of the GSSMC \( \{(S_n, C_n): n \geq 0\} \)

1. [Initialization] Set \( n = 0 \). Select marking \( S_0 \) and clock readings \( C_{0,i} \) for \( e_i \in E(S_0) \); set \( C_{0,i} = -1 \) for \( e_i \not\in E(S_0) \).

2. Determine holding time \( t^*(S_n, C_n) \) and firing set \( E_n^* \).

3. Generate new marking \( S_{n+1} \) according to \( p(\cdot; S_n, E_n^*) \).

4. Set clock-reading \( C_{n+1,i} \) for each new transition \( e_i \) according to \( F(\cdot; S_{n+1}, e_i, S_n, E_n^*) \).

5. Set clock-reading \( C_{n+1,i} \) for each old transition \( e_i \) as \( C_{n+1,i} = C_{n,i} - t^*(S_n, C_n)r(S_n, e_i) \).

6. Set clock-reading \( C_{n+1,i} \) equal to \(-1\) for each newly disabled transition \( e_i \).

7. Set \( n \leftarrow n + 1 \) and go to Step 2.

Can compute GSMP \( \{X(t): t \geq 0\} \) from GSSMC.
Implementation Considerations for Large-Scale SPNs

- Use event lists (e.g., heaps) to determine $E^*$
  - $O(1)$ computation of $E^*$
  - $O(\log m)$ update time, where $m = \#$ of enabled transitions
- Updating the state is often simpler in an SPN
- Efficient techniques for event scheduling [Chiola91]
  - Encode transitions potentially affected by firing of $e_i$
- Parallel simulation of subnets
  - E.g., Adaptive Time Warp [Ferscha & Richter PNPM97]
  - Guardedly optimistic
  - Slows down local firings based on history of rollbacks
Part V

Stability Theory for SPNs
Stability and Simulation

- Focus on time-average limits:

\[ r(f) = \lim_{t \to \infty} \frac{1}{t} \int_0^t f(X(u)) \, du \quad \tilde{r}(\tilde{f}) = \lim_{n \to \infty} \frac{1}{n-1} \sum_{i=0}^{n-1} \tilde{f}(S_n, C_n) \]

- Ex: long-run cost, availability, utilization
- Extensions:
  - Functions (e.g. ratios) of such limits
  - Cumulative rewards (impulse/continuous/mixed), gradients
  - Steady-state means
- Key questions:
  - When do such limits exist?
  - When do various estimation methods apply?
  - Can get weird behavior: \( \lim_n E [\zeta_n - \zeta_{n-1}] = \infty \) but explodes!
- Approach: analyze the chain \( \{(S_n, C_n) : n \geq 0\} \)
Harris Recurrence: A Basic Form of Stability

- Definition for general chain \( \{ Z_n : n \geq 0 \} \) with state space \( \Gamma \)

\[
P_Z \{ Z_n \in A \text{ i.o.} \} = 1, \quad z \in \Gamma \quad \text{whenever} \quad \phi(A) > 0
\]

- \( \phi \) is a recurrence measure (often “Lebesgue-like”)
- Every “dense enough” set is hit infinitely often w.p. 1
- No “wandering off to \( \infty \)”

- Positive Harris recurrence:
  - Chain admits invariant probability measure \( \pi \)
  - \( P_\pi \{ Z_1 \in A \} = \pi(A) \)
  - Implies stationarity when initial dist’n is \( \pi \)

- When is \( \{ (S_n, C_n) : n \geq 0 \} \) (positive) Harris recurrent?
  - Fundamental question for steady-state estimation
Some Stability Conditions

- Density component $g$ of a cdf $F$: $F(t) \geq \int_0^t g(u) \, du$
- $s \rightarrow s'$ iff $p(s'; s, e) > 0$ for some $e$
- $s \sim s'$: either $s \rightarrow s'$ or $s \rightarrow s^{(1)} \rightarrow \ldots \rightarrow s^{(n)} \rightarrow s'$
- Assumption PD($q$):
  - Marking set $G$ is finite
  - SPN is irreducible: $s \sim s'$ for all $s, s' \in G$
  - All speeds are positive
  - There exists $\bar{x} \in (0, \infty)$ s.t. all clock-setting dist'n functions
    - Have finite $q$th moment
    - Have density component positive on $[0, \bar{x}]$
- Assumption PDE: replace finite $q$th moment requirement by
  \[
  \int_0^\infty \int_0^{a_F} e^{ux} dF(x) < \infty \quad \text{for} \quad u \in [0, a_F]
  \]
Harris Recurrence in SPNs

- **Embedded chain:** \( \{ (S_n, C_n) : n \geq 0 \} \) observed only at transitions to timed markings
- \( \bar{\phi}(\{s\} \times A) = \text{Lebesgue measure of } A \cap [0, \bar{x}]^M \)
- **Theorem:** If Assumption PD(1) holds, then the embedded chain is positive Harris recurrent with recurrence measure \( \bar{\phi} \)
- Implies \( P_\mu \{ S_n = s \text{ i.o.} \} = 1 \) for all \( s \in S \)
- **Proof:**
  - First assume no immediate transitions
  - Show that embedded chain is “\( \bar{\phi} \)-irreducible”
  - Establish Lyapunov drift condition and apply MC machinery
  - Extend to case of immediate transitions using strong mimicry
- **Alternate approach to recurrence:** geometric-trials arguments
  - Can drop positive-density assumption
  - Use detailed analysis of specific SPN structure
A Surprising Recurrence Result [Glynn and Haas 2007]

- $S_n =$ marking just after $n$th marking change
- Conjecture: $P \{ S_n = s \text{ i.o.} \} = 1$ for each $s$ if
  - Marking set $S$ is finite
  - SPN is irreducible
  - $\exists \bar{x} > 0$ s.t. each $F(\cdot; e)$ has positive density on $(0, \bar{x})$
- CONJECTURE IS FALSE!
  - In the presence of heavy-tailed clock-setting dist’ns
The Counterexample

- \( S = \{(2, 1, 1), (1, 2, 1), (1, 1, 2)\} \)
- \( p(s'; s, e^*) = 0 \) or \( 1 \)
  (see schematic diagram)
- Clock-setting distributions:
  - \( F(t; e_1) = 1 - (1 + t)^{-\alpha} \)
  - \( F(t; e_2) = 1 - (1 + t)^{-\beta} \)
  - \( F(\cdot; e_3) \) is Uniform[0, \( a \)]

  with \( \beta > 1/2 \) and \( \alpha + \beta < 1 \)
- SPN hits marking \( s = (1, 2, 1) \) only if:
  - \( e_1 \) occurs and then \( e_2 \) occurs
  - No intervening occurrence of \( e_3 \)
- Theorem: \( P\{S_n = (1, 2, 1) \) i.o. \} = 0
Another Type of Stability: Limit Theorems

- **Theorem (SLLN):** If Assumption PD(1) holds, then for any \( f \)
  \[
  \lim_{t \to \infty} \frac{1}{t} \int_{0}^{t} f(X(u)) \, du = r(f) \text{ a.s.}
  \]

- **Theorem (FCLT):** If Assumption PD(2) holds, then for any \( f \)
  \[
  U_\nu(f) \Rightarrow \sigma(f) W \quad \text{as } \nu \to \infty
  \]
  \[
  U_\nu(f)(t) = \nu^{-1/2} \int_{0}^{\nu t} \left( f(X(u)) - r(f) \right) \, du
  \]
  \[
  \Rightarrow \text{ denotes weak convergence on } C[0, \infty)
  \]
  \[
  W = \text{standard Brownian motion on } [0, \infty)
  \]
  \[
  \text{“Functional” form of CLT (ordinary CLT is a special case)}
  \]

- **Note:** \( r(f) \) and \( \sigma(f) \) are independent of initial conditions

- **Follows from general result in [Glynn and Haas 2006]**
  - Uses results for Harris recurrent MCs
FCLT Example: Donsker’s Theorem

\[ S_n = \sum_{i=0}^{n} X_i \]
Part VI

Steady-State Simulation
A regenerative process can be decomposed into i.i.d. cycles

System “probabilistically restarts” at each $T_i$
  - Ex: successive arrival times to an empty GI/G/1 queue

Analogous definition for discrete-time process $\{X_n : n \geq 0\}$

Extension: one-dependent cycles
  - Harris recurrent chains are od-regenerative (basis for previous SLLN and FCLT)
Regenerative Simulation: The Ratio Formula

Let

\[ Y_i = \int_{T_{i-1}}^{T_i} f(X(u)) \, du \quad \text{and} \quad \tau_i = T_i - T_{i-1} \]

\((Y_1, \tau_1), (Y_2, \tau_2), \ldots \) are i.i.d. pairs

It follows that

\[ \frac{1}{T_n} \int_0^{T_n} f(X(u)) \, du = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} \tau_i} = \frac{\bar{Y}_n}{\bar{\tau}_n} \to \frac{E[Y_1]}{E[\tau_1]} \overset{\text{def}}{=} r \]

almost surely as \( n \to \infty \) (need \( E[\tau_1] < \infty \))

Can show that

\[ \frac{1}{t} \int_0^{t} f(X(u)) \, du \to r \text{ a.s. as } t \to \infty \]

If \( \tau_1 \) is “aperiodic”, then \( X(t) \Rightarrow X \) and \( E[f(X)] = r \)
Regenerative Simulation: The Regenerative Method

▶ **Point estimate** (biased): \( \hat{r}_n = \bar{Y}_n / \bar{\tau}_n \):
  - \( \hat{r}_n \to r \text{ a.s. as } n \to \infty \) (strong consistency)

▶ **Confidence interval**
  - Set \( Z_i = Y_i - r \tau_i \)
  - \( Z_1, Z_2, \ldots \) i.i.d. with \( E[Z_i] = 0 \) and \( \text{Var}[Z_1] = \sigma^2 \)
  - Apply Central Limit Theorem (CLT) for i.i.d. random variables:

\[
\frac{\sqrt{n}(\hat{r}_n - r)}{\sigma / E[\tau_1]} \Rightarrow N(0, 1) \quad \text{and} \quad \frac{\sqrt{n}(\hat{r}_n - r)}{s_n / \bar{\tau}_n} \Rightarrow N(0, 1)
\]

as \( n \to \infty \), where \( s_n \) estimates \( \sigma \) (we assume \( \sigma^2 < \infty \))

▶ 100\(p\)% asymptotic confidence interval:

\[
\left[ \hat{r}_n - \frac{z_p s_n}{\bar{\tau}_n \sqrt{n}}, \hat{r}_n + \frac{z_p s_n}{\bar{\tau}_n \sqrt{n}} \right],
\]

where \( P \{ -z_p \leq N(0, 1) \leq z_p \} = p \), i.e., \( (1 + p)/2 \) quantile

▶ **Many extensions**: bias reduction, fixed-time or fixed-precision, generalized \( Y \) and \( \tau \), estimate \( \alpha = g(E[Y], E[\tau]) \), ...
Regenerative Simulation of SPNs

- A marking $\bar{s}$ is a **single state** if $E(\bar{s}) = \{ \bar{e} \}$
- Define $\theta(k) = k$th marking change at which $\bar{e}$ fires in $\bar{s}$
- Set $T_k = \zeta_{\theta(k)}$ and $\tau_k = T_k - T_{k-1}$
- **Theorem**: Suppose Assumption PD(2) holds and SPN has a single state $\bar{s}$
  - Random times $\{ T_k : k \geq 0 \}$ form sequence of regeneration points for marking process
  - Finite expected cycle length: $E_{\mu}[\tau_1] < \infty$
  - Finite variance constant for any $f$:
    $$\sigma^2(f) = \text{Var}_{\mu} \left[ \int_{T_0}^{T_1} f(X(u)) \, du - r\tau_1 \right] < \infty$$
- Can therefore apply standard regenerative method
- **Variant theorems are available**
  - Variants of single state (e.g., memoryless property)
  - Other recurrence conditions (geometric trials)
  - Discrete-time results
The Method of Batch Means

- Simulate system to (large) time $t = mv$ (where $10 \leq m \leq 20$)
- Divide into $m$ batches of length $v$ and compute batch means:

$$
\bar{Y}_i = \frac{1}{v} \int_{(i-1)v}^{iv} f(X(u)) \, du
$$

- Treat $\bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_m$ as i.i.d., $N(\mu, \sigma^2)$:
  - Point estimate: $\hat{r}_t = \frac{1}{m} \sum_{i=1}^{m} \bar{Y}_i$
  - $100p\%$ confidence interval:

$$
\left[ \hat{r}_t - \frac{t_{p,m-1} s_m}{\sqrt{m}}, \hat{r}_t + \frac{t_{p,m-1} s_m}{\sqrt{m}} \right],
$$

where $t_{p,m-1} = (1 + p)/2$ quantile of Student’s $T$ dist’n
Batch Means, Continued

Why might batch means work?
Formally, want to show
- Consistency of $\hat{r}_t$ and validity of CI as $t \to \infty$
- For $m$ fixed (standard batch means)
- What if $m = m(t)$? Overlapping batches?

Special case of standardized-time-series methods
Standardized Time Series

- Consider a mapping $\xi : C[0, 1] \mapsto \mathbb{R}$ such that
  - $\xi(ax) = a\xi(x)$ and $\xi(x - be) = \xi(x)$, where $e(t) = t$
  - $P\{\xi(W) > 0\} = 1$ and $P\{W \in D(\xi)\} = 0$
- Set $\bar{Y}_\nu(t) = (1/\nu) \int_0^t f(X(u)) \, du$ and $\hat{r}_\nu = \bar{Y}_\nu(1)$
- Theorem: If Assumption PD(2) holds, then $r$ exists and

$$
\frac{\hat{r}_\nu - r}{\xi(\bar{Y}_\nu)} = \frac{\sqrt{\nu}(\hat{r}_\nu - r)}{\xi(\sqrt{\nu}(\bar{Y}_\nu - re))} \Rightarrow \frac{\sigma W(1)}{\sigma \xi(W)} = \frac{W(1)}{\xi(W)},
$$

so that an asymptotic 100$p\%$ confidence interval for $r$ is

$$
[\hat{r}_\nu - \xi(\bar{Y}_\nu)z_p, \hat{r}_\nu + \xi(\bar{Y}_\nu)z_p],
$$

where $P\{-z_p \leq W(1)/\xi(W) \leq z_p\} = p$
- Different choices of $\xi$ yield different estimation methods
  - batch means (fixed # of batches)
  - STS area method, STS maximum method
Consistent-Estimation Methods (Discrete Time)

- Set $\hat{r}_n = (1/n) \sum_{j=0}^{n-1} \tilde{f}(S_j, C_j)$ and suppose that

$$\lim_{n \to \infty} \hat{r}_n = \tilde{r} \text{ a.s. and } \frac{\sqrt{n}(\hat{r}_n - \tilde{r})}{\tilde{\sigma}} \Rightarrow N(0, 1)$$

- If we can find a consistent estimator $V_n \Rightarrow \tilde{\sigma}^2$, then

$$\frac{\sqrt{n}(\hat{r}_n - \tilde{r})}{V_n^{1/2}} \Rightarrow N(0, 1)$$

- Then an asymptotic 100$p\%$ confidence interval for $\tilde{r}$ is

$$\left[\hat{r}_n - \frac{z_p}{\sqrt{n}} V_n^{1/2}, \hat{r}_n + \frac{z_p}{\sqrt{n}} V_n^{1/2}\right],$$

where $z_p = (1 + p)/2$ quantile of $N(0, 1)$

- Narrower asymptotic confidence intervals than STS methods
Consistent-Estimation Methods for SPNs

- Look at polynomially dominated functions:
  \[
  \tilde{f}(s, c) = O(1 + \max_{1 \leq i \leq M} c_i^q) \quad \text{for some } q \geq 0
  \]

- Require aperiodicity: no partition of marking set \( G \) s.t.
  \( G_1 \rightarrow G_2 \rightarrow \cdots \rightarrow G_d \rightarrow G_1 \rightarrow G_2 \rightarrow \cdots \)

- Focus on “localized quadratic-form variance estimators”
  - Quadratic-form:
    \[
    V_n = \sum_{i=0}^{n} \sum_{j=0}^{n} \tilde{f}(S_i, C_i)\tilde{f}(S_j, C_j)q_{i,j}^{(n)}
    \]
  - Localized:
    \[
    |q_{i,j}^{(n)}| \leq \begin{cases} 
      a_1/n & \text{if } |i - j| \leq m(n); \\
      a_2(n)/n & \text{if } |i - j| > m(n)
    \end{cases}
    \]
    where \( a_2(n) \rightarrow 0 \) and \( m(n)/n \rightarrow 0 \)
Theorem: For an aperiodic SPN, suppose that

- Assumption PDE holds (\( \exists \) invariant distribution \( \pi \))
- \( \{ \tilde{f}(S_n, C_n) : n \geq 0 \} \) obeys a CLT with variance constant \( \tilde{\sigma}^2 \)
- \( V_n \) is a localized quadratic-form estimator of \( \tilde{\sigma}^2 \)
- \( V_n \Rightarrow \tilde{\sigma}^2 \) when initial distribution = \( \pi \)

Then \( V_n \Rightarrow \tilde{\sigma}^2 \) for any initial distribution

Proof:

- \( \{ (S_n, C_n) : n \geq 0 \} \) couples with stationary version
- Localization: difference between \( V_n \) versions becomes negligible

Consequence: can exploit existing consistency results for stationary output
Coupling Harris-Ergodic Markov Chains

\[ Z_n(\lambda) \]

\[ Z_n(\mu) \]
Application to Specific Variance Estimators

- **Variable batch means estimator of** $\tilde{\sigma}^2$:
  - $b(n)$ batches of $m(n)$ observations each
  - VBM estimator is consistent if Assumption PDE holds, $\tilde{f}$ is polynomially dominated, $b(n) \to \infty$, and $m(n) \to \infty$.

- **Spectral estimator of** $\tilde{\sigma}^2$:
  - Form of estimator: $V_n^{(S)} = \sum_{h=-\infty}^{m-1} \lambda(h/m) \hat{R}_h$
  - $\hat{R}_h =$ sample lag-$h$ autocorrelation of $\{ \tilde{f}(S_n, C_n): n \geq 0 \}$
  - $\lambda(\cdot) =$ “regular” window function (Bartlett, Hanning, Parzen)
  - $m = m(n) =$ spectral window length
  - Spectral estimator is consistent if Assumption PDE holds, $\tilde{f}$ is polynomially dominated, $m(n) \to \infty$, and $m(n)/n^{1/2} \to 0$

- **Overlapping batch means**: asymp. equivalent to spectral

- Can extend results to **continuous time** (and drop aperiodicity)
Want to estimate \( \lim_{n \to \infty} \left( \frac{1}{n} \sum_{j=0}^{n-1} f(D_j) \right) \)

Delays \( D_0, D_1, \ldots \) “determined by marking changes of the net”

Specified as \( D_j = B_j - A_j \)

- **Starts:** \( \{ A_j = \zeta_{\alpha(j)} : j \geq 0 \} \) nondecreasing
- **Terminations:** \( \{ B_j = \zeta_{\beta(j)} : j \geq 0 \} \)
- Determined by \( \{ (S_n, C_n) : n \geq 0 \} \)

Measuring lengths of delay intervals is nontrivial

- Must link starts and terminations
- Multiple ongoing delays
- Overtaking: delays need not terminate in start order
- Can avoid for limiting average delay \( \lim_{n \to \infty} \left( \frac{1}{n} \sum_{j=0}^{n-1} D_j \right) \)

Measurement methods: tagging and start vectors
Tagging
Start Vectors

- Assume \# of ongoing delays = \( \psi(s) \) when marking is \( s \)
- \( V_n \) records starts for all ongoing delays at \( \zeta_n \)
- Positions of starts = position of entities in system (usually)
- Use -1 as placeholder
- At each marking change:
  - Insert current time according to \( i_{\alpha}(s'; s, E^*) \)
  - Delete components according to \( i_{\beta}(s'; s, E^*) \)
  - Permute components according to \( i_{\pi}(s'; s, E^*) \)
  - Subtract deleted components from current time to compute delays (ignore -1’s)
Start Vector Example

\[ T = 2.9 \]

\[ V_5 = (2.9, 2.4, 0) \]

\[ D = 2.9 - 1.2 = 1.7 \]
Regenerative Methods: The Easy Case

▶ Assume SPN has single state and “well behaved” cycles
▶ Use standard regenerative method
Regenerative Methods: The Hard Case

- Assume SPN has single state and “well behaved” cycles
- Decompose delays into one-dependent cycles
- Use extended regenerative method or multiple-runs method
Limiting Average Delay

- Under appropriate regularity conditions

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} D_j = \frac{E_\mu[Z_1]}{E_\mu[\delta_1]} \text{ a.s.}$$

- $\delta_1 =$ # of starts in regenerative cycle
- $Z_1 = \int_{\text{cycle}} \psi(X(t)) \, dt$
- $\psi(s) =$ # of ongoing delays when marking is $s$
- $(Z_1, \delta_1), (Z_2, \delta_2), \ldots$ are i.i.d.

- Can use standard regenerative method
- No need to measure individual delays
- One proof of this result uses Little’s Law
STS Methods for Delays

- Focus on “regular” start-vector mechanism
- Use polynomially-dominated functions $f: \mathbb{R}_+ \mapsto \mathbb{R}$:
  \[ |f(x)| = O(1 + x^q) \text{ for some } q \geq 0 \]
- **Theorem**: If Assumption PDE holds, then
  \[
  \lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} f(D_j) = r \text{ a.s.} \quad \text{and} \quad U_n(f) \Rightarrow \sigma(f)W
  \]
  where $U_n(f)(t) = n^{-1/2} \int_0^t (f(D_{\lfloor u \rfloor}) - r) \, du$
- **Proof**: Identify one-dependent cycles
- Apply limit theorems for od-regenerative processes
Colored SPNs

- Tokens have color and transitions fire “in a color”
- Yields more concise graphs
- “Symmetry with respect to color”
  - Captures variety of system symmetries
  - Can exploit to improve simulation efficiency
    - Shorter regenerative cycle lengths
    - Shorter CIs for delays
      Ex: delay for port 1 in symmetric token ring
Part VII

Conclusion
Summary

- **SPNs are an attractive framework for simulation**
  - User-friendly graphical orientation
  - Powerful and flexible modeling tool
  - Solid mathematical basis
- **Efficiency in sample-path generation**
- **Simulation theory: building-block conditions for**
  - Stability (recurrence, limit theorems)
  - Validity of simulation methods
- **Simulation methods:**
  - Regenerative
  - Standardized time series (batch means)
  - Consistent-estimation methods (spectral and VBM)
- **Further resources**
  - INFORMS College on Simulation (http://www.informs-cs.org)
  - www.almaden.ibm.com/cs/people/peterh