# Explicit formulas for the variance of the Creation Spectrum and the F- Monte Carlo network reliability estimation methods

Héctor Cancela[1] Leslie Murray[2] Gerardo Rubino[3]

[1]Facultad de Ingeniería
Universidad de la República
Montevideo, Uruguay

[2]Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Universidad Nacional de Rosario
Rosario, Argentina

[3]INRIA–Rennes
Bretagne Atlantique, Campus de Beaulieu
Rennes, France

# Aim of this talk

Present work-in-progress analyzing the variance of three Monte Carlo methods for network reliability estimation:

- Standard Monte Carlo
- F-Monte Carlo
  E. Canale, F. Robledo, P. Romero, P. Sartor *Monte Carlo methods in diameter-constrained reliability.* Optical Switching and Networking, 14, 134-148, 2014.
- Creation Spectrum
  I. B. Gertsbakh, Y. Shpungin, *Models of Network Reliability Analysis, Combinatorics, and Monte Carlo*, CRC Press, Boca Raton, 2012.

# Talk Outline

1. Static Network Reliability Model
2. Standard Monte Carlo
3. F–Monte Carlo
4. Creation Spectrum
5. Discussion and conclusions

# Static Network Reliability Model

- Network topology: graph $G = (V, E)$, with $|V| = n$ and $|E| = m$.
- Nodes are perfect.
- Links fail independently from each other.
- State of link $i$:
  $X_i = \begin{cases} 1 \text{ w.p. } p_i & i^{th} \text{ link operational or } \textit{up} \\ 0 \text{ w.p. } q_i = 1 - r_i & i^{th} \text{ link failed or } \textit{down} \end{cases}$
- Homogeneous network model: $p_i = p$ and $q_i = q = 1 - p$, $i = 1, \ldots, m$.
- Network state space $\Omega = \{0, 1\}^m$.
- State of the network: random vector $\mathbf{X} = (X_1, \ldots, X_m) \in \Omega$.
- Configuration $\mathbf{x} = (x_1, \ldots, x_m)$: any possible value of $\mathbf{X}$.

$$P(\mathbf{X} = \mathbf{x}) = \prod_i P(X_i = x_i).$$

# Static Network Reliability Model (2)

- $\Omega$ is the union of two disjoint subspaces, one of which associated to a network *up* condition, and the other one associated to a network *down* condition.
- Usually, network *up* means that a subset $K \subseteq V$ of nodes is connected by the *up* links, and network *down* means that $K$ of nodes is disconnected.
- Structure function $\phi(\mathbf{X}) \in \{0, 1\}$: $\phi(\mathbf{X}) = 0$ means that network is *down*, whereas $\phi(\mathbf{X}) = 1$ means that network is *up*.
- Network reliability, $R$, and network unreliability, $Q$:

$$R = \mathbb{P}\{\phi(\mathbf{X}) = 1\} \quad \text{and} \quad Q = \mathbb{P}\{\phi(\mathbf{X}) = 0\} = 1 - R.$$

- Computing network reliability is a well-known hard computational problem (in $\#P$ class).

# Network Reliability Polynomial

- Partition $\Omega = \{0,1\}^m$ in subsets:

$$\Omega_j = \left\{ \mathbf{x} \in \Omega \ : \ \sum_{i=1}^{m} x_i = j \right\} \quad j = 0, \ldots, m,$$

  $|\Omega_j| = \binom{m}{j}$.
- Per link state independency, for any $\mathbf{x} \in \Omega_j$, $P(\mathbf{X} = \mathbf{x}) = p^j \, q^{m-j}$, then

$$P(\mathbf{X} \in \Omega_j) = \binom{m}{j} p^j \, q^{m-j} = M_j.$$

- $\alpha_j$: fraction of configurations in every $\Omega_j$ such that $\phi(\mathbf{x}) = 1$.
- Network reliability polynomial:

$$R = \sum_{j=1}^{m} \alpha_j \binom{m}{j} p^j \, q^{m-j} = \sum_{j=1}^{m} \alpha_j \, M_j.$$

The classic standard Monte Carlo estimation is based on sampling $N$ independent copies of $\mathbf{X}$ from the original distribution $\mathbb{P}\{X_i = 1\} = p$, $\mathbb{P}\{X_i = 0\} = q = 1 - p$, $i = 1, \ldots, m$:

$$\widehat{R}_{\mathrm{S}} = \frac{1}{N} \sum_{i=1}^{N} \phi\left(\mathbf{X}^{(i)}\right)$$

Its variance is:

$$\begin{aligned}
\mathbb{V}\{\widehat{R}_{\mathrm{S}}\} &= \frac{1}{N^2}\, \mathbb{V}\left\{\sum_{i=1}^{N} \phi\left(\mathbf{X}^{(i)}\right)\right\} = \frac{1}{N}\, \mathbb{V}\left\{\phi\left(\mathbf{X}^{(i)}\right)\right\} \\
&= \frac{1}{N}\, R(1-R) = \frac{1}{N}\, R - R^2 \\
&= \frac{1}{N}\, \left[\sum_{j=1}^{m} \alpha_j\, Mj - \left(\sum_{j=1}^{m} \alpha_j M_j\right)^2\right]
\end{aligned}$$

# F–Monte Carlo

F–Monte Carlo is based on computing unbiased estimators $\widehat{\alpha}_j$ for every $\alpha_j$ in order to build an unbiased network reliability estimator:

$$\widehat{R}_{\mathsf{F}} = \sum_{j=1}^{m} \widehat{\alpha}_j Mj$$

Define random vector $\mathbf{Y}_j$ uniformly distributed in $\Omega_j$ (i.e, a random vector of size $m$, composed of $j$ ones and $m-j$ zeros):
Then applying Standard Monte Carlo for estimating $\alpha_j$:

$$\widehat{\alpha}_j = \frac{1}{N_j} \sum_{i=1}^{N_j} \phi\left(\mathbf{Y}_j^{(i)}\right) \quad j = 0, \ldots, m,$$

where $\mathbf{Y}_j^{(i)}$, $i = 1, \ldots, N_j$, is a set of independent samples of $\mathbf{Y}_j$.

# F–Monte Carlo

Replacing each $\widehat{\alpha}_j$ by its estimator, the F–Monte Carlo network unreliability estimator is:

$$\widehat{R}_{\mathrm{F}} = \sum_{j=1}^{m} \left( \frac{1}{N_j} \sum_{i=1}^{N_j} \phi\left(\mathbf{Y}_j^{(i)}\right)\right) M_j$$

and its variance is

$$\begin{aligned}
\mathbb{V}\{\widehat{R}_{\mathrm{F}}\} &= \mathbb{V}\left\{ \sum_{j=1}^{m} \left( \frac{1}{N_j} \sum_{i=1}^{N_j} \phi\left(\mathbf{Y}_j^{(i)}\right)\right) Mj \right\} \\
&= \sum_{j=1}^{m} \frac{Mj^2}{N_j^2} \mathbb{V}\left\{ \sum_{i=1}^{N_j} \phi\left(\mathbf{Y}_j^{(i)}\right) \right\} \\
&= \sum_{j=1}^{m} \frac{Mj^2}{N_j^2} N_j \, \mathbb{V}\left\{ \phi\left(\mathbf{Y}_j^{(i)}\right)\right\} = \sum_{j=1}^{m} \frac{Mj^2}{N_j} \; \alpha_j(1-\alpha_j)
\end{aligned}$$

# F-MC Variance (2)

Distributing $N$ Monte Carlo replications uniformly between the coefficient estimators ($N_j = N/m,\ j = 1, \ldots, m$), then:

$$\begin{aligned}
\mathbb{V}\{\widehat{R}_{\mathrm{F}}\} &= \sum_{j=1}^{m} \frac{m}{N} M_j^2 \, (\alpha_j - \alpha_j^2) \\
&= \frac{m}{N} \sum_{j=1}^{m} Mj^2(\alpha_j - \alpha_j^2) \\
&= \frac{m}{N} \left[ \sum_{j=1}^{m} \alpha_j Mj^2 - \sum_{j=1}^{m} \alpha_j^2 Mj^2 \right]
\end{aligned}$$

# Creation Spectrum Monte Carlo

Auxiliary dynamic model: initially all links are *down*; they are repaired sequentially (only taking into account the order, without considering times between repairs).
Reparation order is a *permutation* $\pi$ of the links in $E$:

$$\pi = \{e_{(1)}, e_{(2)}, \ldots, e_{(m)}\}$$

When all the links are *down*, the network is also *down*.
It becomes operational simultaneously with the repair of some $e_{(k)}$, $1 \leq k \leq m$. We call *anchor* of $\pi$, denoted $a(\pi)$, to the number $k$.
$\pi$ is a random variable on the state space $\Pi$.
$a(\pi)$ is a random variable on $\{1, 2, \ldots, m\}$; its PDF is called Creation Spectrum:

$$\mathrm{C}_S = (f_1, f_2, \ldots, f_m),$$

where $f_k = \mathbb{P}\{a(\pi) = k\}, \; k = 1, \ldots, m$.

The Cumulative Creation Spectrum is the corresponding CDF:

$$\mathrm{CC}_S = (F_1, F_2, \dots, F_m)$$

where $F_k = \sum_{i=1}^{k} f_i, \; k = 1, \dots, m.$
By a combinatorial argument, it can be shown that $F_k = \alpha_k$ (defined previously - fraction of up configurations in $\Omega_j$), and

$$R = \sum_{j=1}^{m} \alpha_j M_j$$

Then

$$R = \sum_{j=1}^{m} \left( \sum_{k=1}^{j} f_k \right) M_j = \sum_{j=1}^{m} f_j \left( \sum_{k=j}^{m} M_k \right)$$

# Creation Spectrum Monte Carlo (3)

Calling $U_j = \sum_{k=j}^{m} M_k$,

$$R = \sum_{j=1}^{m} f_j U_j$$

If we generate randomly $N$ *permutations*, $(\pi_1, \ldots, \pi_N)$, and for each one of them, we find its *anchor*, $a(\pi_k)$, if $A_k$ is the number of *permutations* $\pi$ for which $a(\pi) = k$, then $\widehat{f_k} = A_k/N$ is an estimator of $f_k$, $k = 1, \ldots, m$.
The Creation Spectrum Monte Carlo estimator is:

$$\widehat{R}_{CS} = \sum_{j=1}^{m} \widehat{f_j} U_j.$$

# Creation Spectrum Variance

Define the indicator random vector $\mathbf{B} = (B_1, \ldots, B_m)$:

$$B_j = \begin{cases} 1 & \text{if } a(\pi) = j \\ 0 & \text{otherwise,} \end{cases} \qquad j = 1, \ldots, m.$$

If we sample $N$ independent *permutations* $\pi^{(i)}$ and compute the corresponding samples of $\mathbf{B}$, then:

$$\widehat{f}_j = \frac{1}{N} \sum_{i=1}^{N} B_j^{(i)}$$

The Creation Spectrum estimator can be written as:

$$\widehat{R}_{\mathrm{CS}} = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{m} B_j^{(i)} U_j$$

# Creation Spectrum Variance (2)

Then

$$\begin{aligned}
\mathbb{V}\{\widehat{R}_{\mathrm{CS}}\} &= \mathbb{V}\left\{\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{m} B_j^{(i)}\, U_j\right\} \\
&= \frac{1}{N^2}\,\mathbb{V}\left\{\sum_{i=1}^{N}\sum_{j=1}^{m} B_j^{(i)}\, U_j\right\} \\
&= \frac{1}{N^2}\sum_{i=1}^{N}\mathbb{V}\underbrace{\left\{\sum_{j=1}^{m} B_j^{(i)}\, U_j\right\}}_{Z^{(i)}}
\end{aligned}$$

Here, $Z^{(i)}$, $i = 1, \ldots, N$, are independent samples from random variable $Z$, taking values $U_j$ with probabilities $f_j$, $j = 1, \ldots, m$.

# Creation Spectrum Variance (3)

$$\mathbb{V}\{Z\} = \mathbb{E}\{Z^2\} - \mathbb{E}\{Z\}^2 = \sum_{j=1}^{m} f_j\, U_j^2 \;-\; \left(\sum_{j=1}^{m} f_j\, U_j\right)^2$$

Finally:

$$\mathbb{V}\{\widehat{R}_{C_S}\} = \frac{1}{N^2} \sum_{i=1}^{N} \mathbb{V}\left\{Z^{(i)}\right\} = \frac{1}{N}\,\mathbb{V}\{Z\}$$

$$= \frac{1}{N}\left[\sum_{j=1}^{m} f_j\, U_j^2 \;-\; \left(\sum_{j=1}^{m} f_j\, U_j\right)^2\right]$$

where $f_1 = \alpha_1$, $f_j = \alpha_j - \alpha_{j-1}$, $j = 2, \ldots, m$, and $U_j = \sum_{i=j}^{m} M_i = \sum_{i=j}^{m} \binom{m}{i} p^i q^{m-i}$.

# Variance formulas for the three methods

$$\mathbb{V}\{\widehat{R}_{\mathrm{S}}\} = \frac{1}{N} \left[ \sum_{j=1}^{m} \alpha_j \; Mj \; - \; \left( \sum_{j=1}^{m} \alpha_j \; Mj \right)^2 \right]$$

$$\mathbb{V}\{\widehat{R}_{\mathrm{F}}\} = \frac{m}{N} \left[ \sum_{j=1}^{m} \alpha_j Mj^2 - \sum_{j=1}^{m} \alpha_j^2 Mj^2 \right]$$

$$\mathbb{V}\{\widehat{R}_{\mathrm{CS}}\} = \frac{1}{N} \left[ \sum_{j=1}^{m} f_j \; U_j^2 \; - \; \left( \sum_{j=1}^{m} f_j \; U_j \right)^2 \right]$$

$$M_j = \binom{m}{j} p^j \, q^{m-j}; \; U_j = \sum_{k=j}^{m} M_k; \; f_j = \alpha_j - \alpha_{j-1}$$

# Numerical example

Small Grid network:

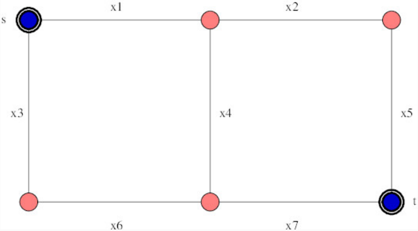


Table: Exact variances of Standard Monte Carlo, F–Monte Carlo, and Destruction Spectrum Monte Carlo, with sample size $N = 10^8$

| | $Q = 1 - R$ | $\mathbb{V}\{\widehat{Q}_{\mathrm{S}}\}$ | $\mathbb{V}\{\widehat{Q}_{\mathrm{F}}\}$ | $\mathbb{V}\{\widehat{Q}_{\mathrm{DS}}\}$ |
|---|---|---|---|---|
| $r = 0.9$ | 4.12E−02 | 3.95E−10 | 1.74E−10 | 2.88E−11 |
| $r = 0.99$ | 4.03E−04 | 4.03E−12 | 4.31E−14 | 6.24E−15 |
| $r = 0.999$ | 4.00E−06 | 4.00E−14 | 4.71E−18 | 6.74E−19 |

# Conclusions and Lines of future work

- We showed explicit formulas using $\alpha_j$ values.
- Computing $\alpha_j$ is also computationally expensive, so using the formulas to compute exact variances is only feasible for small networks, and in general is as difficult as computing the reliability.
- Nevertheless, formulas can be used to analitically compare methods' precisions.
- Lines of future work: use the formulas over the (unknown) $\alpha_j$ to provide a priori comparisons of the precision of the methods, to find better parameters or importance-sampling versions (in the case of F-Monte Carlo); and to study asymptotics.

Thanks!