Some applications of importance sampling to dependability analysis

Bruno Tuffin
(based on joint works with
H. Cancela, M. El Khadiri, P. L'Ecuyer, G. Rubino, S. Saggadi)

INRIA Rennes - Centre Bretagne Atlantique

Paris
2016

Outline

1. Rare events, Static reliability estimation
2. Graph reductions to decrease the work-normalized variance
3. An adaptive ZVIS approximtation
4. Combination with Recursive Variance Reduction
   - Recursive Variance Reduction (RVR) algorithm
   - Zero-variance Approximation RVR
5. Conclusions

Outline

1 Rare events, Static reliability estimation

2 Graph reductions to decrease the work-normalized variance

3 An adaptive ZVIS approximtation

4 Combination with Recursive Variance Reduction
- Recursive Variance Reduction (RVR) algorithm
- Zero-variance Approximation RVR

5 Conclusions

Introduction: rare events and dependability

- in *telecommunication networks*: loss probability of a small unit of information (a packet, or a cell in ATM networks), connectivity of a set of nodes,
- in *dependability analysis*: probability that a system is failed at a given time, availability, mean-time-to-failure,
- in *air control systems*: probability of collision of two aircrafts,
- in *particle transport*: probability of penetration of a nuclear shield,
- in *biology*: probability of some molecular reactions,
- in *insurance*: probability of ruin of a company,
- in *finance*: value at risk (maximal loss with a given probability in a predefined time),
- ...

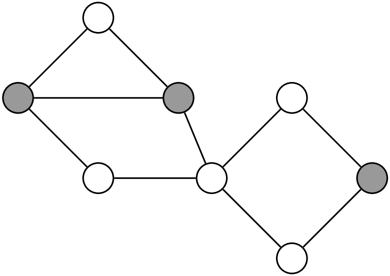
Robustness properties

- In rare-event simulation models, we often parameterize with a rarity parameter $\epsilon > 0$ such that $\mu = \mathbb{E}[X(\epsilon)] \to 0$ as $\epsilon \to 0$.
- An estimator $X(\epsilon)$ is said to have *bounded relative variance* (or *bounded relative error*) if $\sigma^2(X(\epsilon))/\mu^2(\epsilon)$ is bounded uniformly in ϵ.
- Interpretation: estimating $\mu(\epsilon)$ with a given relative accuracy can be achieved with a bounded number of replications even if $\epsilon \to 0$.
- Weaker property: *asymptotic optimality* (or *logarithmic efficiency*) if $\lim_{\epsilon\to 0} \ln(\mathbb{E}[X^2(\epsilon)])/\ln(\mu(\epsilon)) = 2$.
- Stronger property: *vanishing relative variance*: $\sigma^2(X(\epsilon))/\mu^2(\epsilon) \to 0$ as $\epsilon \to 0$. Asymptotically, we get the zero-variance estimator.
- Other robustness measures exist (based on higher degree moments, on the Normal approximation, on simulation time...).

L'Ecuyer, Blanchet, T., Glynn, ACM ToMaCS 2010

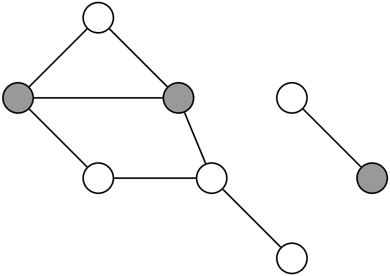
Graph model

- M links can fail independently, *elementary unreliability* $q_e = 1 - r_e$ for edge e.
- What is the probability that the set $\mathcal{K}$ of (grey) nodes is connected (in the underlying random partial graph of $\mathcal{G}$)?
- $X = (X_1, \ldots, X_M)$ (random) *configuration* with $X_e = 1$ if edge e works, 0 otherwise.
- state of the system: $\phi(X)$, where $\phi(X) = 1$ iff $\mathcal{K}$ not connected.
- $u = \mathbb{E}[\phi(X)] = \sum_{x\in\{0,1\}^M} \phi(x)\mathbb{P}[X = x]$.



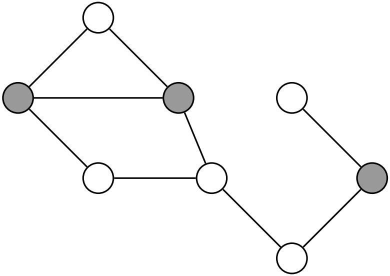
Graph model

- M links can fail independently, *elementary unreliability* $q_e = 1 - r_e$ for edge e.
- What is the probability that the set $\mathcal{K}$ of (grey) nodes is connected (in the underlying random partial graph of $\mathcal{G}$)?
- $X = (X_1, \ldots, X_M)$ (random) *configuration* with $X_e = 1$ if edge e works, 0 otherwise.
- state of the system: $\phi(X)$, where $\phi(X) = 1$ iff $\mathcal{K}$ not connected.
- $u = \mathbb{E}[\phi(X)] = \sum_{x \in \{0,1\}^M} \phi(x)\mathbb{P}[X = x]$.



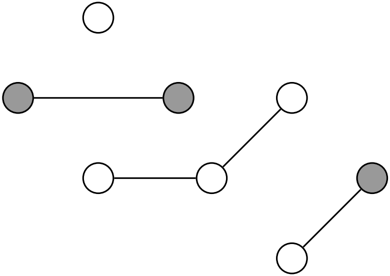
Graph model

- M links can fail independently, *elementary unreliability* $q_e = 1 - r_e$ for edge e.
- What is the probability that the set $\mathcal{K}$ of (grey) nodes is connected (in the underlying random partial graph of $\mathcal{G}$)?
- $X = (X_1, \ldots, X_M)$ (random) *configuration* with $X_e = 1$ if edge e works, 0 otherwise.
- state of the system: $\phi(X)$, where $\phi(X) = 1$ iff $\mathcal{K}$ not connected.
- $u = \mathbb{E}[\phi(X)] = \sum_{x \in \{0,1\}^M} \phi(x)\mathbb{P}[X = x]$.



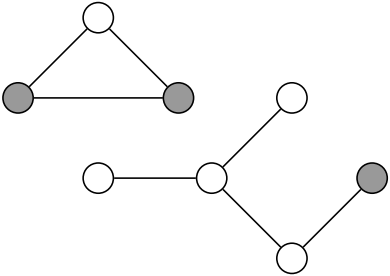
Graph model

- M links can fail independently, *elementary unreliability* $q_e = 1 - r_e$ for edge e.
- What is the probability that the set $\mathcal{K}$ of (grey) nodes is connected (in the underlying random partial graph of $\mathcal{G}$)?
- $X = (X_1, \ldots, X_M)$ (random) *configuration* with $X_e = 1$ if edge e works, 0 otherwise.
- state of the system: $\phi(X)$, where $\phi(X) = 1$ iff $\mathcal{K}$ not connected.
- $u = \mathbb{E}[\phi(X)] = \sum_{x \in \{0,1\}^M} \phi(x)\mathbb{P}[X = x]$.



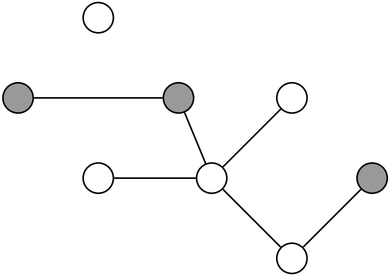
Graph model

- M links can fail independently, *elementary unreliability* $q_e = 1 - r_e$ for edge e.
- What is the probability that the set $\mathcal{K}$ of (grey) nodes is connected (in the underlying random partial graph of $\mathcal{G}$)?
- $X = (X_1, \ldots, X_M)$ (random) *configuration* with $X_e = 1$ if edge e works, 0 otherwise.
- state of the system: $\phi(X)$, where $\phi(X) = 1$ iff $\mathcal{K}$ not connected.
- $u = \mathbb{E}[\phi(X)] = \sum_{x \in \{0,1\}^M} \phi(x) \mathbb{P}[X = x]$.



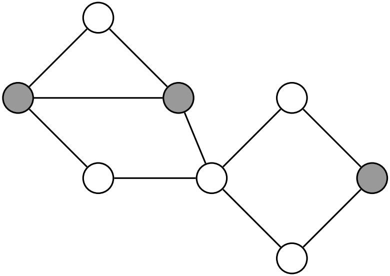
Graph model

- M links can fail independently, *elementary unreliability* $q_e = 1 - r_e$ for edge e.
- What is the probability that the set $\mathcal{K}$ of (grey) nodes is connected (in the underlying random partial graph of $\mathcal{G}$)?
- $X = (X_1, \ldots, X_M)$ (random) *configuration* with $X_e = 1$ if edge e works, 0 otherwise.
- state of the system: $\phi(X)$, where $\phi(X) = 1$ iff $\mathcal{K}$ not connected.
- $u = \mathbb{E}[\phi(X)] = \sum_{x \in \{0,1\}^M} \phi(x) \mathbb{P}[X = x]$.



Graph model

- M links can fail independently, *elementary unreliability* $q_e = 1 - r_e$ for edge e.
- What is the probability that the set $\mathcal{K}$ of (grey) nodes is connected (in the underlying random partial graph of $\mathcal{G}$)?
- $X = (X_1, \ldots, X_M)$ (random) *configuration* with $X_e = 1$ if edge e works, 0 otherwise.
- state of the system: $\phi(X)$, where $\phi(X) = 1$ iff $\mathcal{K}$ not connected.
- $u = \mathbb{E}[\phi(X)] = \sum_{x \in \{0,1\}^M} \phi(x) \mathbb{P}[X = x]$.



- We have to sum over the 2^M configurations.

Crude simulation

- Consider n independent copies $X^{(i)} = (X_1^{(i)}, \ldots, X_m^{(i)})$ of X, and compute $Y^{(i)} = \phi(X^{(i)})$.
- The crude estimator of q is then

$$\hat{Y}_n = \frac{1}{n} \sum_{i=1}^{n} Y^{(i)}.$$

- Confidence interval built from the central limit theorem.
- Rarity issue:
  - We assume $q_e \to 0 \ \forall e$, so that $u \to 0$.
  - The *relative error* is proportional to

$$\frac{\sqrt{\mathrm{Var}[\hat{Y}_n]}}{\mathbb{E}[Y]} = \frac{\sqrt{u(1-u)}}{u\sqrt{n-1}} \to \infty$$

    as $u \to 0$.
  - As a consequence, more and more paths are required to get a specified relative error as $u \to 0$.

Zero-variance est. L'Ecuyer, Rubino, Saggadi & T., IEEE Trans. Rel. 2011

- Idea: sample the links one after the other, with an IS probability that *depends on the state of previously sampled links*.
- Let $u_m(x_1, \cdots, x_{m-1})$, with $x_i \in \{0, 1\}$, be the unreliability of the graph G given the states of the links 1 to $m-1$: if $x_i = 1$ the link i is operational, otherwise it is failed.
- Then $u = u_1()$.
- Sample state of link m, giving 1 with probability:

$$\tilde{q}_m = u'_m(x_1, \cdots, x_{m-1}) = \frac{q_m u_{m+1}(x_1, \cdots, x_{m-1}, 0)}{(1-q_m)u_{m+1}(x_1, \cdots, x_{m-1}, 1) + q_m u_{m+1}(x_1, \cdots, x_{m-1}, 0)}.$$

- Remark (by conditionning) that

$$u_m(x_1, \cdots, x_{m-1}) = (1-q_m)u_{m+1}(x_1, \cdots, x_{m-1}, 1) + q_m u_{m+1}(x_1, \cdots, x_{m-1}, 0).$$

- The resulting unbiased estimator is $\phi(X)L(X)$, with

$$L(x) = \prod_{i=1}^{\ell} L_i(x_i) = \prod_{i=1}^{\ell}\left(x_i \frac{1-q_i}{1-\tilde{q}_i} + (1-x_i)\frac{q_i}{\tilde{q}_i}\right).$$

Where does it come from?

- From the zero-variance IS for a DTMC $(Y_j)_j$ trying to compute

$$\mu(Y_0) = \sum_{j=1}^{\tau} c(Y_{j-1}, Y_j)$$

- Use change of probability transitions

$$\begin{aligned}\tilde{P}(y,z) &= \frac{P(y,z)(c(y,z)+\mu(z))}{\sum_w P(y,w)(c(y,w)+\mu(w))} \\ &= \frac{P(y,z)(c(y,z)+\mu(z))}{\mu(y)}\end{aligned}$$

- This yields the *unique* Markov chain implementation of the zero-variance estimator.

Zero-variance estimation and approximation

Proposition

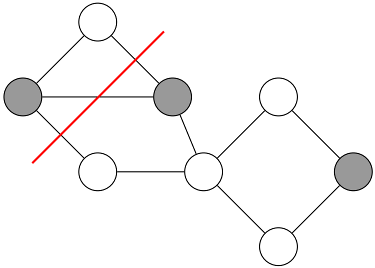
Using this IS, the estimator has zero variance (always yields u).

- Problem: the $u_m(\cdot)$ are not known, otherwise no need to simulate.
- Principle: approach $u_m(\cdot)$ by some $\hat{u}_m(\cdot)$ and use

$$\tilde{q}_m = \frac{q_m \hat{u}_{m+1}(x_1, \cdots, x_{m-1}, 0)}{q_m \hat{u}_{m+1}(x_1, \cdots, x_{m-1}, 0) + (1 - q_m)\hat{u}_{m+1}(x_1, \cdots, x_{m-1}, 1)}.$$

Approximation of the zero-variance estimator

- Our proposal: $\hat{u}_m(x_1, \cdots, x_{m-1})$ is the probability of a mincut of the graph with highest probability, given the state of links 1 to $m-1$.
  - A *cut* (or $\mathcal{K}$-cut) is a set of edges such that, if we remove them, the nodes in $\mathcal{K}$ are not in the same connected component.
  - A *mincut* (minimal cut) is a cut that contains no other cut than itself.



- Intuition: the unreliability is the probability of union of all cuts, the most crucial one(s) being the mincut(s) with highest probability.
- Cuts can be obtained in polynomial time.

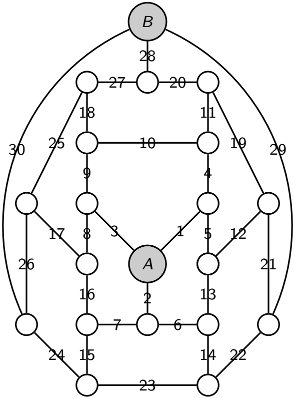
Results

- In a given state $(x_1, \cdots, x_{m-1})$, we need to determine $\hat{u}_{m+1}(x_1, \cdots, x_{m-1}, 1)$ and $\hat{u}_{m+1}(x_1, \cdots, x_{m-1}, 0)$.
- This adds some computational burden, but should substantially reduce the variance.

Proposition

Bounded relative error proved in general,
Vanishing relative error under identified conditions.

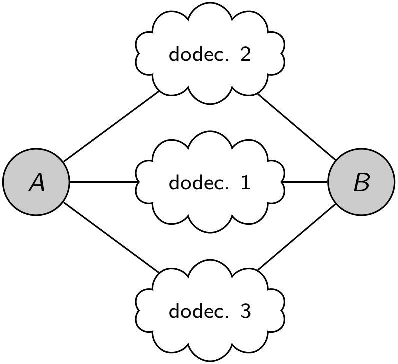
Ex: dodecahedron topology, all links with unreliability ϵ



$q_e = \epsilon$	Estimation	Confidence interval	Std deviation	Relative error
10^{-1}	$2.8960\ 10^{-3}$	$(2.8276\ 10^{-3}, 2.9645\ 10^{-3})$	$3.49\ 10^{-3}$	1.2
10^{-2}	$2.0678\ 10^{-6}$	$(2.0611\ 10^{-6}, 2.0744\ 10^{-6})$	$3.42\ 10^{-7}$	0.17
10^{-3}	$2.0076\ 10^{-9}$	$(2.0053\ 10^{-9}, 2.0099\ 10^{-9})$	$1.14\ 10^{-10}$	0.057
10^{-4}	2.0007^{-12}	$(2.0000\ 10^{-12}, 2.0014\ 10^{-12})$	$3.46\ 10^{-14}$	0.017

With respect to crude MC, a computational time increase of 16.

Larger networks: 3 dodecahedrons in parallel



$q_e = \epsilon$	Estimate	95% confidence interval	std dev.	Relative Error
10^{-1}	2.3573×10^{-8}	$(2.2496 \times 10^{-8},\ 2.4650 \times 10^{-8})$	5.49×10^{-8}	2.3
5×10^{-2}	2.5732×10^{-11}	$(2.5138 \times 10^{-11},\ 2.6327 \times 10^{-11})$	3.03×10^{-11}	1.2
10^{-2}	8.7655×10^{-18}	$(8.7145 \times 10^{-18},\ 8.8165 \times 10^{-18})$	2.60×10^{-18}	0.30

- Vanishing relative error observed
- For 3 dodecahedron in series, Bounded relative error observed
- Works very well for such topologies with close to 100 links, and larger.

Outline

1 Rare events, Static reliability estimation

2 Graph reductions to decrease the work-normalized variance

3 An adaptive ZVIS approximtation

4 Combination with Recursive Variance Reduction
- Recursive Variance Reduction (RVR) algorithm
- Zero-variance Approximation RVR

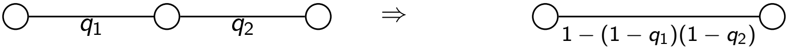
5 Conclusions

Improving ZVIS by applying graph reductions when sampling links

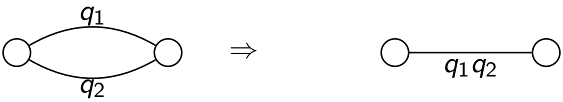
- Each time a link state is generated by the ZVIS algorithm, the graph evolves according to these rules: at step i $(1 \leq i \leq \ell)$,
  - either $X_i = 0$ which means that the link is removed,
  - or $X_1 = 1$ which means that the link is fixed, and can then be removed by merging the two nodes it links.
- At each step, we can therefore search if graph reductions can be applied, in order to simplify the topology, and potentially gain in terms of
  - variance
  - computational time (because the size of the graph is smaller).

Considered graph reductions

- Series reduction:
  - If node $s \in \mathcal{N}$ has only two incident links, l_1 and l_2, connecting it to nodes s_1 and s_2 respectively
  - If $s \notin \mathcal{K}$, node s can be removed and links l_1 and l_2 merged into a single one, with unreliability $q = 1 - (1 - q_{l_1})(1 - q_{l_2})$.



  - The case $s \in \mathcal{K}$ can hardly be treated without further topology information.
- Parallel reduction:
  - if there are two (parallel) links l_1 and l_2 both connecting nodes s_1 and s_2
  - those two links merged into a single one, with unreliability $q = q_{l_1} q_{l_2}$.



Two possible combinations with ZVIS

- *Posterior reduction* (PR)
  - link i sampled with failed probability

$$\hat{q}_i^{(1)} = \frac{q_i \hat{u}_{i+1}(\mathcal{G}_i^r, 0)}{q_i \hat{u}_{i+1}(\mathcal{G}_i^r, 0) + (1 - q_i)\hat{u}_{i+1}(\mathcal{G}_i^r, 1)},$$

    where $\mathcal{G}_i^r$ graph resulting from previous link samplings and reductions
  - link i is removed if $X_i = 0$ and compressed if $X_i = 1$
  - new reductions are searched , leading to a new graph $\mathcal{G}_{i+1}^r$.
- *Look-ahead reduction* (LAR)
  - the probability that i is failed:

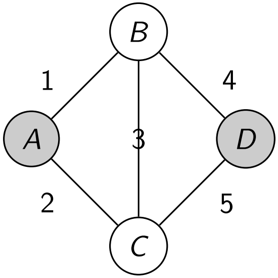
$$\hat{q}_i^{(2)} = \frac{q_i \hat{u}_{i+1}(\mathcal{G}_{i,0}^r)}{q_i \hat{u}_{i+1}(\mathcal{G}_{i,0}^r) + (1 - q_i)\hat{u}_{i+1}(\mathcal{G}_{i,1}^r)},$$

    where $\mathcal{G}_{i,k}^r$ for $k \in \{0, 1\}$ is the graph *reduced after* setting $X_i = k$
  - This requires to make two copies of the graph, setting $X_i = 0$ for the first and $X_i = 1$ for the other,
  - those two resulting graphs being reduced according to the above rules
  - When link i effectively sampled, we choose the appropriate *already* reduced graph.

Expected gain

- Computational time:
  - Time for graph reduction searches and making copies of the graph
  - but it decreases the number of links to sample and the number of mincut-maxprob approximations to be computed.
- Variance:
  - better mincut-maxprob approximation of the graph unreliabilities at the different steps, *usually* resulting in smaller variance.
- Comparing the two implementations:
  - LAR requires additional time to make copies of the graph and to perform twice more reductions at any given step
  - but computing the mincut-maxprob on an already reduced graph takes a shorter time than before proceeding to a reduction.
  - Moreover we usually get a better approximation of the zero-variance IS with this procedure.

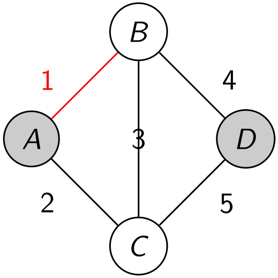
Toy example with cascading reductions



First sample link 1.

- If $X_1 = 1$,

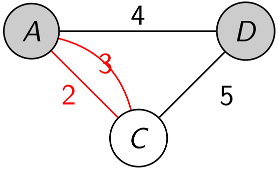
Toy example with cascading reductions



First sample link 1.

- If $X_1 = 1$,
  - the graph can then be reduced by compressing link 1, merging nodes A and B,

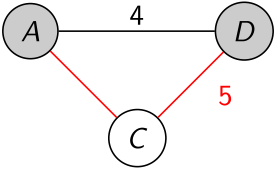
Toy example with cascading reductions



First sample link 1.

- If $X_1 = 1$,
  - the graph can then be reduced by compressing link 1, merging nodes A and B,
  - then a parallel reduction of links 2 and 3 can be applied.

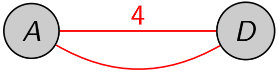
Toy example with cascading reductions



First sample link 1.

- If $X_1 = 1$,
  - the graph can then be reduced by compressing link 1, merging nodes A and B,
  - then a parallel reduction of links 2 and 3 can be applied.
  - This new link is then in series with link 5, leading to a reduction.

Toy example with cascading reductions



First sample link 1.

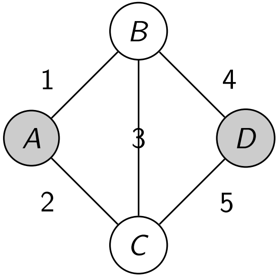
- If $X_1 = 1$,
  - the graph can then be reduced by compressing link 1, merging nodes A and B,
  - then a parallel reduction of links 2 and 3 can be applied.
  - This new link is then in series with link 5, leading to a reduction.
  - The resulting graph is then just made of two parallel links which can therefore be reduced.

Toy example with cascading reductions

First sample link 1.

- If $X_1 = 1$,
  - the graph can then be reduced by compressing link 1, merging nodes A and B,
  - then a parallel reduction of links 2 and 3 can be applied.
  - This new link is then in series with link 5, leading to a reduction.
  - The resulting graph is then just made of two parallel links which can therefore be reduced.
  - By IS, the link is necessarily considered failed. Just one link sampled!

Toy example with cascading reductions



- If $X_1 = 0$, proceeding similarly,
  - Link 1 is removed.
  - Links 3 and 4 are then in series and can be reduced,
  - the resulting link becomes a parallel link with link 5, reduced
  - to a link in series with link 2, which can be reduced to lead to a single link, necessarily failed under IS. Just one link sampled too!

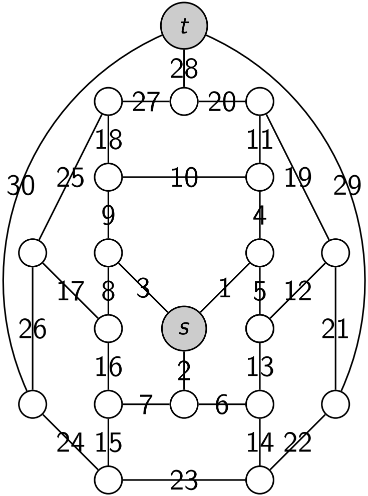
In terms of variance

The algorithms have the following robustness properties, as failures of individual links go to zero:

- On our toy example:
  - With PR, VRE is obtained
  - While with LAR, zero variance is obtained (perfect approximation of unreliabilities).
- With full generality,

Proposition

Our algorithms satisfy BRE.



$q_i = \epsilon \; \forall i$

Met.	ϵ	Variance	RE.	Time
ZVIS	10^{-1}	1.1048×10^{-5}	1.1733	15.18
ZVIS	10^{-2}	1.1670×10^{-13}	0.1652	14.35
ZVIS	10^{-3}	1.2714×10^{-20}	0.0561	14.88
PR	10^{-1}	5.5452×10^{-6}	0.8190	12.14
PR	10^{-2}	9.8889×10^{-14}	0.1522	15.33
PR	10^{-3}	9.5548×10^{-21}	0.0487	13.87
LAR	10^{-1}	3.9203×10^{-6}	0.6880	10.29
LAR	10^{-2}	4.4955×10^{-14}	0.1028	7.48
LAR	10^{-3}	2.4094×10^{-21}	0.0244	7.55

Outline

1 Rare events, Static reliability estimation

2 Graph reductions to decrease the work-normalized variance

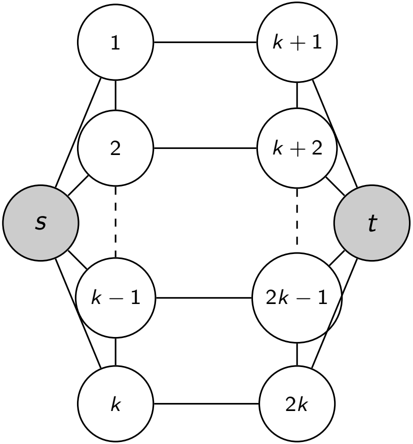
3 An adaptive ZVIS approximtation

4 Combination with Recursive Variance Reduction
- Recursive Variance Reduction (RVR) algorithm
- Zero-variance Approximation RVR

5 Conclusions

Limits of the above ZVIS approximation

- Shown to be very efficient for very low link unreliabilities
- But system failure rarity may come from other reasons. Ex: large number of possible paths.

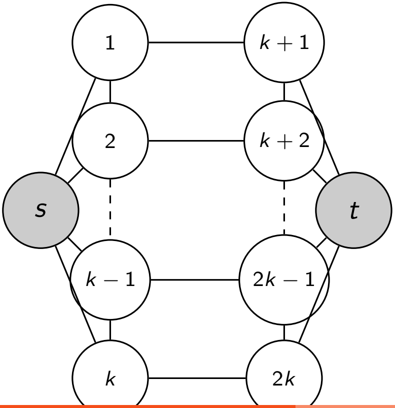


Increasing k but keeping the same overall unreliability

k	q_e	$10^8\,\hat{u}$	$\hat{\mathrm{RE}}$	$\hat{u}_1^{\mathrm{mc}}(\emptyset) = q^r$
2	7×10^{-5}	1.46	0.33	4.9×10^{-9}
5	0.02	1.06	0.46	3.2×10^{-9}
10	0.1245	1.11	1.8	8.9×10^{-10}
30	0.371	1.14	7.9	1.2×10^{-13}
40	0.427	1.05	9.9	1.6×10^{-15}
50	0.4665	1.08	31	2.7×10^{-17}
70	0.521	1.35	22	1.5×10^{-20}
100	0.575	1.48	40	9.2×10^{-25}
200	0.655	0.48	44	1.8×10^{-37}

Minpath-based approximation

- *Path*: set P of links such that when up, the nodes in $\mathcal{K}$ are connected.
- *Minpath*: path with no strict subset that is a path.
- *Minpath-maxprob approximation*: max probability of a minpath, $\hat{u}^{\text{mp}}(\mathcal{G}) = 1 - \max_{P \in \mathcal{F}_{\mathcal{G}}} p(P)$.
- Computed thanks to Dijkstra algorithm.
- Replacing the mincut-maxprob approximation in ZVIS



k	q_e	$10^8\hat{u}$	$\hat{\text{RE}}$	$\hat{u}_1^{\text{mp}}(\emptyset)$
2	0.00007	1.68	66	0.0002
5	0.02	3.18	160	0.058
10	0.1245	1.15	110	0.32
30	0.371	1.36	75	0.75
40	0.427	1.20	36	0.81
50	0.4665	0.98	26	0.84
70	0.521	1.58	17	0.89
90	0.559	1.19	6.6	0.91
100	0.575	1.52	9.8	0.92
200	0.655	1.13	3.9	0.95

What if we combine both approximations?

- Indeed, $\hat{u}^{\mathrm{mc}} \leq u \leq \hat{u}^{\mathrm{mp}}$.
- Take at each step

$$\hat{u}_{i+1}(x_1, \ldots, x_i) = \alpha \hat{u}^{\mathrm{mc}}_{i+1}(x_1, \ldots, x_i) + (1-\alpha)\hat{u}^{\mathrm{mp}}_{i+1}(x_1, \ldots, x_i).$$

- Should always be closer to the unreliability.
- How to determine the best α?
- First heuristic:
  - Compute a rough estimate $\widehat{u}_{n_0}(\mathcal{G})$ of u
  - Take

$$\alpha = \alpha_{\mathrm{tot}} \stackrel{\mathrm{def}}{=} \frac{\hat{u}^{\mathrm{mp}}(\emptyset) - \widehat{u}_{n_0}(\mathcal{G})}{\hat{u}^{\mathrm{mp}}(\emptyset) - \hat{u}^{\mathrm{mc}}(\emptyset)},$$

  the α leading to the above equality with this rough estimate. for the full network unreliability.

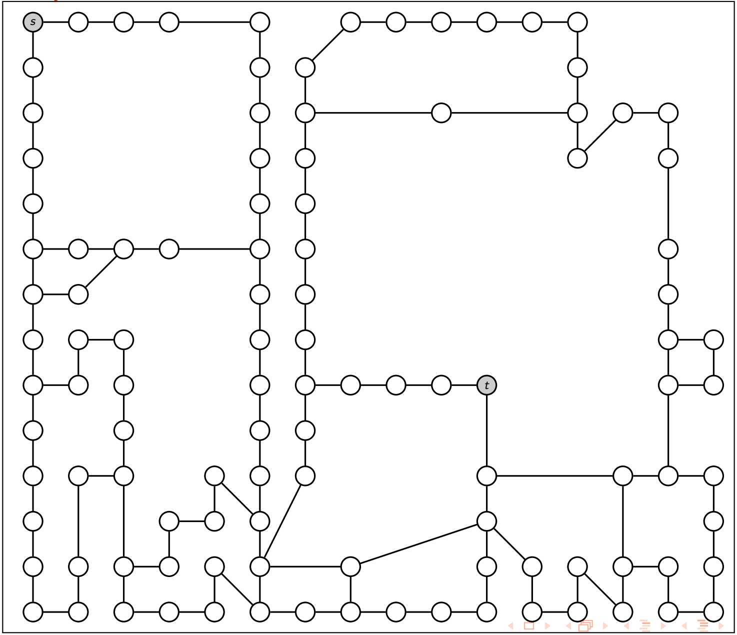
Learning through a Robbins-Monro algorithm

- Goal: compute the α minimizing the variance, i.e., st $V'(\alpha) = 0$.

  1. $\ell = 0$, start with a α_0 (the one from the heuristic)
     1. $\ell = \ell + 1$
     2. estimate $\hat{V}'(\alpha_\ell)$
     3. Stop when it seems to have converged, or update again.
  2. Launch the real simulation with the last α_ℓ.

- I skip the computation of the derivative and choice of parameters (paper available on requests).

Ex: transport network of ANTEL



method	q	$\hat{u}$	$\hat{\mathrm{RE}}$	$\hat{\alpha}$	$\hat{u}^{\mathrm{mc}}(\emptyset)$	$\hat{u}^{\mathrm{mp}}(\emptyset)$
MC	10^{-2}	1.22×10^{-2}	108			
	10^{-3}	2.11×10^{-4}	103			
	10^{-4}	1.92×10^{-6}	109			
MP	10^{-2}	7.58×10^{-3}	2.5			
	10^{-3}	7.48×10^{-5}	8.4			
	10^{-4}	6.74×10^{-7}	25			
heuristic	10^{-2}	7.49×10^{-3}	2.7	0.92873	10^{-4}	0.18209
	10^{-3}	7.37×10^{-5}	11	0.977225	10^{-6}	0.019811
	10^{-4}	7.25×10^{-7}	7.2	0.99770	10^{-8}	1.998×10^{-3}
SA	10^{-2}	7.54×10^{-3}	1.9	0.59887		
	10^{-3}	7.27×10^{-5}	2.8	0.99838		
	10^{-4}	7.26×10^{-7}	2.8	0.999843		

Outline

1 Rare events, Static reliability estimation

2 Graph reductions to decrease the work-normalized variance

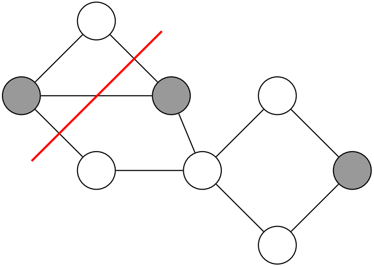
3 An adaptive ZVIS approximtation

4 Combination with Recursive Variance Reduction
- Recursive Variance Reduction (RVR) algorithm
- Zero-variance Approximation RVR

5 Conclusions

Recursive Variance Reduction (RVR)

- Principle: select a $\mathcal{K}$-cutset, i.e., a set $\mathcal{C}$ of links whose failure ensures the system failure.



- If all links in $\mathcal{C}$ are failed (probability $q_{\mathcal{C}}$), the system is failed. Clearly, $q_{\mathcal{C}} \leq q$.
- B_j="the $j-1$ first links of $\mathcal{C}$ are down, but the j-th is up"
- $\mathbb{P}[B_j] = \left(\prod_{k=1}^{j-1} q_k\right) r_j$
- Define $p_j = \mathbb{P}[B_j \,|\, \text{at least one link is working}] = \mathbb{P}[B_j]/(1 - q_{\mathcal{C}})$

Recursive Variance Reduction (RVR)

The RVR estimator:

- Select a cut, and compute $q_{\mathcal{C}}$ and the p_js.
- Pick an edge at random in $\mathcal{C}$ according to the probability distribution $(p_j)_{j=1,\cdots,|\mathcal{C}|}$
- Let the chosen edge be the jth. Call $\mathcal{G}_j$ the graph obtained from $\mathcal{G}$ by deleting the first $j-1$ edges of $\mathcal{C}$ and by contracting the jth.
- The value y_{RVR} returned by the RVR estimator of $q(\mathcal{G})$, the unreliability of $\mathcal{G}$, is recursively defined as

$$y_{RVR}(\mathcal{G}) = q_{\mathcal{C}} + (1 - q_{\mathcal{C}}) y_{RVR}(\mathcal{G}_j).$$

RVR estimator

Formally, the RVR estimator of $q(\mathcal{G})$ is the random variable

$$Y_{RVR} = q_{\mathcal{C}} + (1 - q_{\mathcal{C}}) \sum_{j=1}^{|\mathcal{C}|} \frac{\mathbf{1}_{B_j}}{1 - q_{\mathcal{C}}} Y_{RVR}(\mathcal{G}_j).$$

Theorem

The estimator is unbiased: $\mathbb{E}[Y_{RVR}] = q(\mathcal{G}) = q$.
Second moment computed as

$$\begin{aligned} \mathbb{E}[Y_{RVR}^2] &= q_{\mathcal{C}}^2 + 2q_{\mathcal{C}}(1 - q_{\mathcal{C}}) \left(\sum_{j=1}^{|\mathcal{C}|} \frac{\mathbb{P}[B_j]}{1 - q_{\mathcal{C}}} \mathbb{E}[Y_{RVR}(\mathcal{G}_j)] \right) \\ &\quad + (1 - q_{\mathcal{C}})^2 \left(\sum_{j=1}^{|\mathcal{C}|} \frac{\mathbb{P}[B_j]}{1 - q_{\mathcal{C}}} \mathbb{E}[Y_{RVR}^2(\mathcal{G}_j)] \right). \end{aligned}$$

But no BRE as $\epsilon \to 0$.

Zero-variance Approximation RVR

- *Zero-variance change of measure*: chooses the appropriate (ideally the best) IS for the first working link on the cut:
- choose B'_j with probability $\tilde{p}_j$ in the IS estimator, with

$$\tilde{p}_j = \frac{\mathbb{P}[B_j]q(\mathcal{G}_j)}{\sum_{k=1}^{|\mathcal{C}|}\mathbb{P}[B_k]q(\mathcal{G}_k)} \tag{1}$$

- Resulting estimator:

$$Y_{ZRVR} = q_{\mathcal{C}} + \left(\sum_{k=1}^{|\mathcal{C}|}\mathbb{P}[B_k]q(\mathcal{G}_k)\right)\sum_{j=1}^{|\mathcal{C}|}\mathbf{1}_{B'_j(\mathcal{G})}\frac{1}{q(\mathcal{G}_j)}Y_{ZRVR}(\mathcal{G}_j).$$

Theorem

Y_{ZRVR} has variance $\mathrm{Var}[Y_{ZRVR}] = 0$.

- Implementing it requires the knowledge of the $q(\mathcal{G}_i)$, but in that case, no need to simulate!

Zero Variance Approximation

- Instead, use some approximation $\hat{q}(\mathcal{G}_i)$ of $q(\mathcal{G}_i)$ plugged into (1).

$$Y_{AZRVR} = q_{\mathcal{C}} + \left(\sum_{k=1}^{|\mathcal{C}|} \mathbb{P}[B_k]\hat{q}(\mathcal{G}_k)\right) \sum_{j=1}^{|\mathcal{C}|} \mathbf{1}_{B'_j(\mathcal{G})} \frac{1}{\hat{q}(\mathcal{G}_j)} Y_{AZRVR}(\mathcal{G}_j).$$

Proposition

If $\forall 1 \leq j \leq |\mathcal{C}|$, $\hat{q}(\mathcal{G}_j) = \Theta(q(\mathcal{G}_j))$ as $\epsilon \to 0$, Y_{AZRVR} verifies BRE property.

- Define the *mincut-maxprob* approximation $\hat{q}(\mathcal{G})$ of $q(\mathcal{G})$ as maximal probability of a mincut of graph $\mathcal{G}$ (computed in polynomial time).

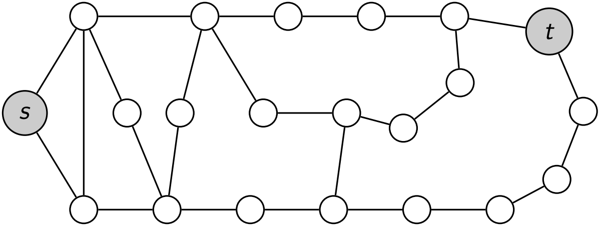
Proposition

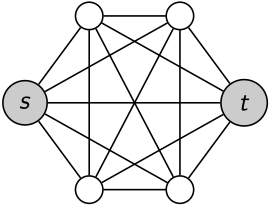
With the mincut-maxprob approximation, $\hat{q}(\mathcal{G}_j) = \Theta(q(\mathcal{G}_j))$ as $\epsilon \to 0$, therefore BRE property is obtained.

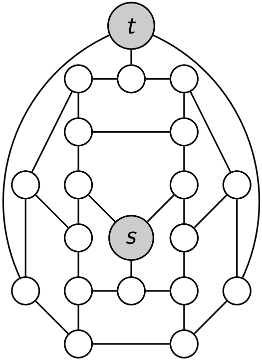
Proposition

If, $\hat{q}(\mathcal{G}_j) = q(\mathcal{G}_j) + o(q(\mathcal{G}_j))$ as $\epsilon \to 0$ for all $1 \leq j \leq |\mathcal{C}|$, the Vanishing relative (VRE) property (the RE tends to 0, stronger than just being bounded) is verified.

Three topologies: arpanet, C6, dodecahedron







Network (q_e)	$Q(G)$	$N \times Var(SMC)$	$N \times Var(RVR)$	$N \times Var(AZV)$
Arp (5.00 $e-01$)	9.63989 e-01	3.47133 e-02	3.71795 e-03	1.69321 e-01
Arp(3.00 e-01)	6.81507 e-01	2.17055 e-01	4.74801 e-02	8.45549 e-01
Arp (1.00 e-01)	9.54229 e-02	8.63174 e-02	1.46865 e-02	9.55806 e-02
Arp (1.00 e-02)	6.54074 e-04	6.53646 e-04	1.63753 e-05	3.06912 e-06
Arp (1.00 e-03)	6.05581 e-06	6.05577 e-06	1.60407 e-08	3.43246 e-11
Arp (1.00 e-04)	6.00560 e-08	6.00560 e-08	1.60041 e-11	3.47090 e-16
Arp (1.00 e-05)	6.00056 e-10	6.00056 e-10	1.60004 e-14	3.47477 e-21
Arp (1.00 e-06)	6.00006 e-12	6.00006 e-12	1.60000 e-17	3.47512 e-26
C6 (5.00 e-01)	7.64160 e-02	7.05766 e-02	7.72612 e-05	7.27858 e-05
C6 (3.00 e-01)	5.26728 e-03	5.23953 e-03	2.56429 e-07	2.27577 e-07
C6 (1.00 e-01)	2.00766 e-05	2.00762 e-05	1.28070 e-13	1.17223 e-13
C6 (1.00 e-02)	2.00001 e-10	2.00001 e-10	1.01244 e-26	1.00225 e-26
C6 (1.00 e-03)	2.00000 e-15	2.00000 e-15	1.00102 e-39	1.00002 e-039
C6 (1.00 e-04)	2.00000 e-20	2.00000 e-20	1.00000 e-52	1.00000 e-52
C6 (1.00 e-05)	2.00000 e-25	2.00000 e-25	1.42434 e-65	1.42434 e-65
Dod (5.00 e-01)	7.09745 e-01	2.06007 e-01	1.57246 e-02	1.34634 e-01
Dod (3.00 e-01)	1.68518 e-01	1.40120 e-01	9.22721 e-03	1.68222 e-02
Dod (1.00 e-01)	2.87960 e-03	2.87131 e-03	5.80985 e-06	6.32871 e-07
Dod (1.00 e-02)	2.06189 e-06	2.06189 e-06	2.17456 e-12	1.12133 e-14
Dod (1.00 e-03)	2.00602 e-09	2.00602 e-09	2.01614 e-18	1.01110 e-21
Dod (1.00 e-04)	2.00060 e-12	2.00060 e-12	2.00160 e-24	1.00110 e-28
Dod (1.00 e-05)	2.00006 e-15	2.00006 e-15	2.00016 e-30	1.00011 e-35
Dod (1.00 e-06)	2.00001 e-18	2.00001 e-18	2.00002 e-36	1.00001 e-42

Outline

1. Rare events, Static reliability estimation
2. Graph reductions to decrease the work-normalized variance
3. An adaptive ZVIS approximtation
4. Combination with Recursive Variance Reduction
   - Recursive Variance Reduction (RVR) algorithm
   - Zero-variance Approximation RVR
5. Conclusions

A summary of best existing methods for static reliability estimation on the dodecahedron

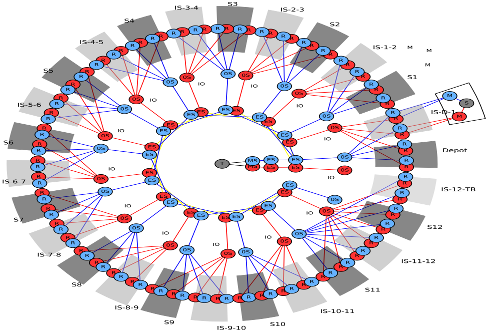
Without presenting all implementations.

(Normalized) relative error $\frac{\sqrt{n}\times \mathrm{RE}}{c_\alpha}$ for various methods and unreliabilities ϵ of links on the dodecahedron topology

Method	$\epsilon = 0.1$	$\epsilon = 10^{-2}$	$\epsilon = 10^{-3}$	$\epsilon = 10^{-4}$
Conditioning, Fishman 86	2.6 e+00	1.3 e+00	4.3 e−01	1.4 e−02
GS Botev et al. 13	4.0 e+00	6.2 e+00	7.7 e+00	8.9 e+00
Splitting, Murray et al. 13	4.6 e+00	7.1 e+00	8.6 e+00	8.8 e+00
Permutation MC Gerbatsh	3.0 e+00	4.2 e+00	4.3 e+00	4.4 e+00
IS: ZVA 2010	1.2 e+00	1.7 e−01	5.7 e−02	1.7 e−02
RVR Cancela, Khadiri 1995	8.4 e−01	7.1 e−01	7.1 e−01	7.1 e−01
IS+ RVR: BRD 14	9.5 e−01	7.0 *e*-01	7.1 e−01	7.1 e−01
IS+RVR: AZVRD 14	2.8 e−01	5.1 e−02	1.6 e−02	5.0 e−03

Work in progress

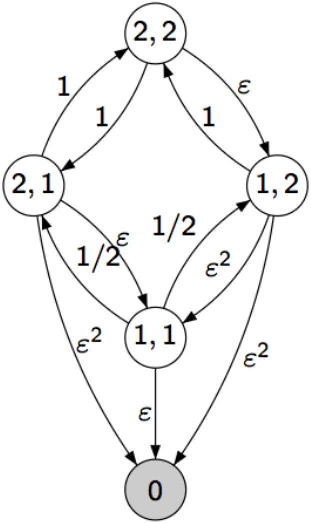
- Railway Data Communication System (DCS), with failing *nodes*



- Dependability including logistics: return to a dynamic model. Two challenges
  - Non-Markovian model
  - more complicated assumptions with logistics on repair teams, spares.

Example: Highly Reliable Markovian Systems (HRMS)

- System with c types of components. $Y = (Y_1, \ldots, Y_c)$ with Y_i number of up components.
- $\mathbf{1}$: state with all components up.
- Markov chain. Failure rates are $O(\varepsilon)$, but not repair rates. Failure propagations possible.
- System down when in grey state(s) (in Δ).
- Goal: compute $\mu(y)$ probability to hit Δ before $\mathbf{1}$.
- $\mu(\mathbf{1})$ important in dependability analysis,
- Small if ε small.



Highly Reliable Markovian Systems (HRMS)

- Failure rates are $O(\varepsilon)$, but not repair rates. Failure propagations possible.
- Simulation using the embedded DTMC. Failure probabilities are $O(\varepsilon)$ (except from $\mathbf{1}$). How to improve (accelerate) this?
- Existing method: $\forall y \neq \mathbf{1}$, increase the probability of the set of failures to constant $0.5 < q < 0.9$ and use individual probabilities proportional to the original ones (SFB), or uniformly (BFB).
- Failures not rare anymore. BRE property verified for BFB.

HRMS Example, and IS

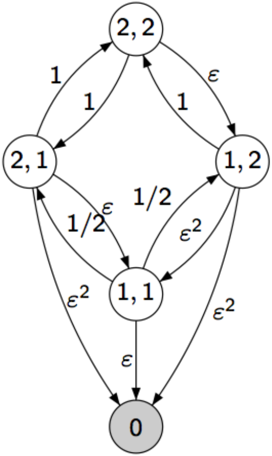


Figure: Original probabilities

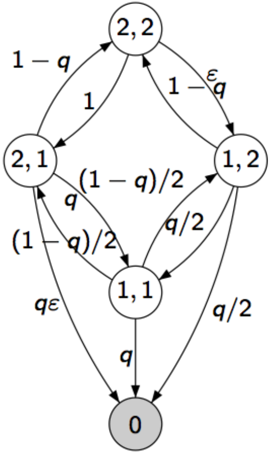


Figure: Probabilities under IS/BFB

HRMS, Zero-variance IS

L'Ecuyer & T., ANOR, 2011

- Complicates the previous model due to the multidimensional description of a state.
- The idea is to approach $\mu(y)$ by the probability of the path from y to Δ with the largest probability
- Intuition: as $\epsilon \to 0$, we get a good idea of the probability.

Proposition

Bounded Relative Error proved (as $\epsilon \to 0$) in general.
Even Vanishing Relative Error if $\hat{\mu}(y)$ contains all the paths with the smallest degree in ϵ.

- Other simple version: approach $\mu(y)$ by the (sum of) probability of paths from y with only failure components of a given type.
- Gain of several orders of magnitudes + stability of the results with respect to the literature.

HRMS: numerical illustrations

- Comparison of BFB and Zero-Variance Approximation (ZVA).
- $c = 3$ types of components, n_i of type i
- $\lambda_1 = \varepsilon$, $\lambda_2 = 1.5\varepsilon$, and $\lambda_3 = 2\varepsilon^2$, $\mu = 1$
- System is down whenever fewer than two components of any one type are operational.

n_i	ε	μ_0	BFB est	ZVA est	BFB σ^2	ZVA σ^2
3	0.001	2.6×10^{-3}	2.7×10^{-3}	2.6×10^{-3}	6.2×10^{-5}	2.2×10^{-8}
6	0.01	1.8×10^{-7}	1.9×10^{-7}	1.8×10^{-7}	6.3×10^{-11}	2.0×10^{-14}
6	0.001	1.7×10^{-11}	1.8×10^{-11}	1.7×10^{-11}	8.8×10^{-19}	1.2×10^{-23}
12	0.1	6.0×10^{-8}	4.8×10^{-8}	6.0×10^{-8}	8.1×10^{-10}	1.6×10^{-10}
12	0.001	3.9×10^{-28}	(1.8×10^{-40})	3.9×10^{-28}	(3.2×10^{-74})	1.4×10^{-55}