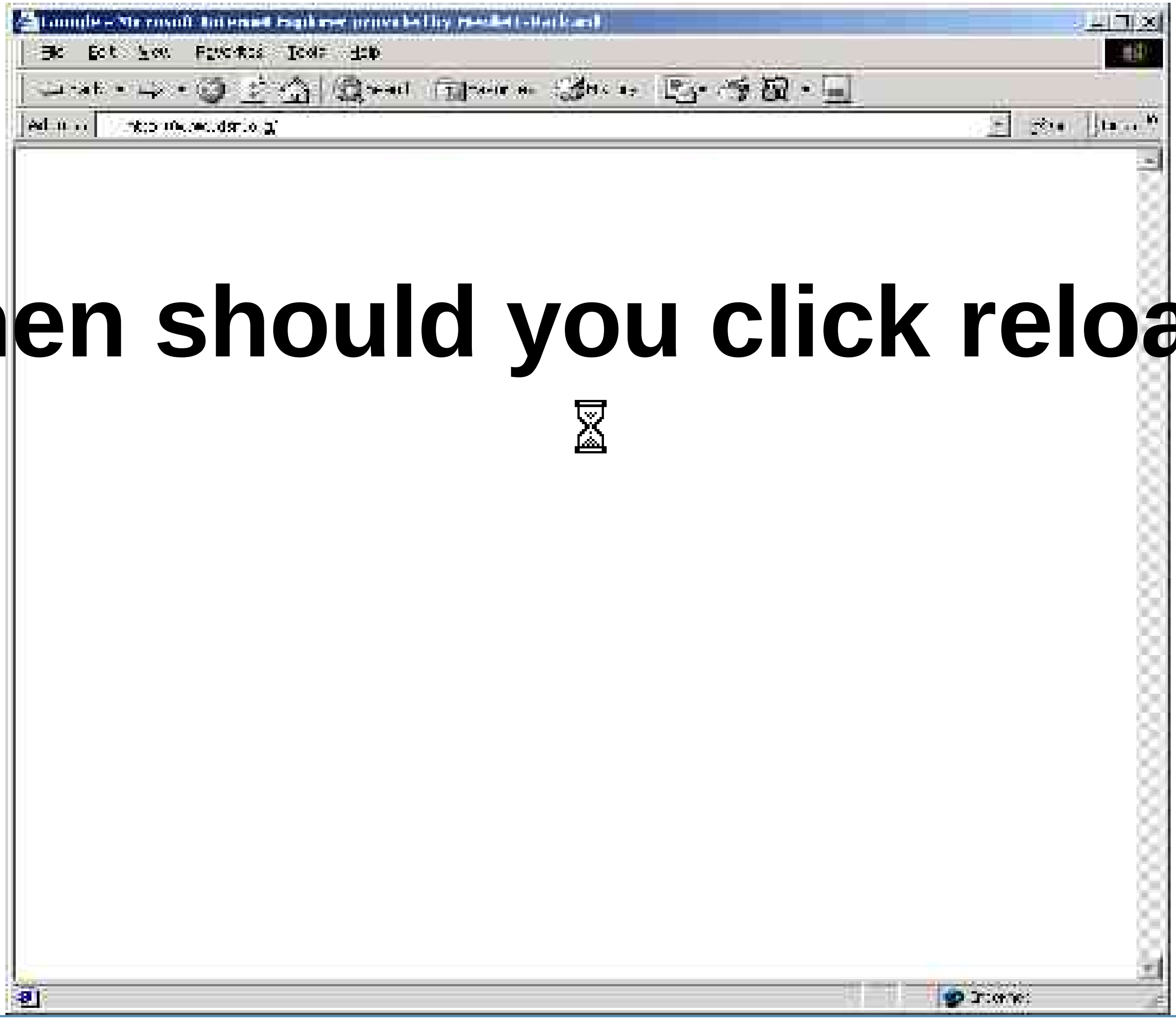
analysis and algorithms for restart

Aad van Moorsel
University of Newcastle

Katinka Wolter, and Philipp Reinecke
Humboldt University, Berlin

example application



when should you click reload?

credit

much work has been done in

- checkpointing
- rejuvenation

- restart is much simpler
- appropriate

- gives us some insights

contents

- mathematical formulation

- restart to improve first moment of completion time
- restart to improve higher moments

- geometric approximation
- optimising different moments simultaneously
- 'engineering' rule for size of restart interval

problem definition

- any task
- when it doesn’ t succeed quick enough, it is retried

model assumptions:

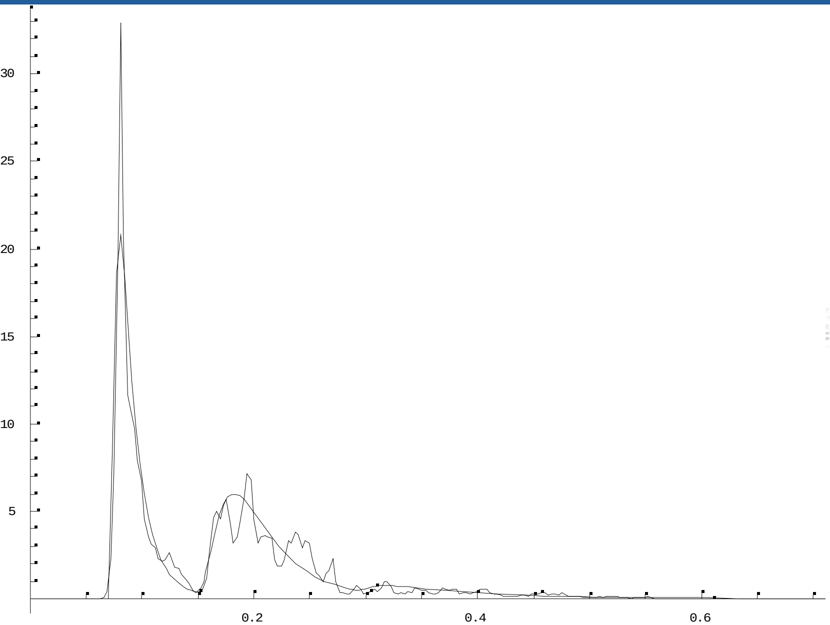
- we know the distribution (e.g. of download time)
- retries are independent, identically distributed (iid)

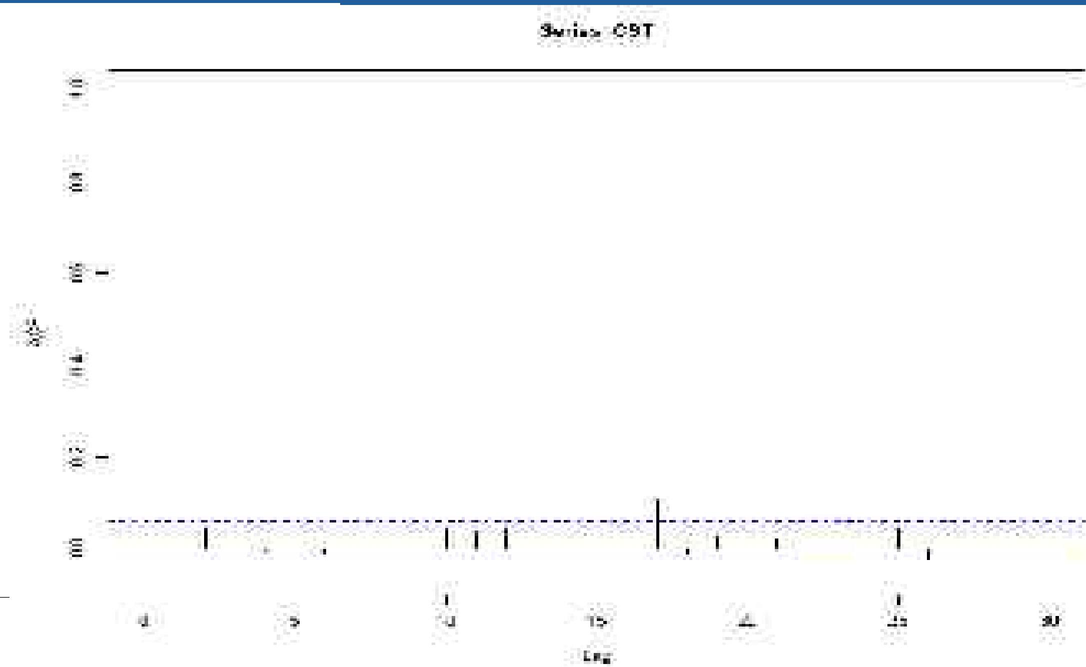
turns out to be realistic for Internet applications (SAForum 2004)

formal: $$E[T] < E[T-\tau \mid T>\tau]$$

- always true for hyper-exponential distribution
- never true for hypo-exponential distribution (Erlang)
- does not hurt nor help for exponential distribution (=)

independence





data overlaid with 3 lognormal distributions

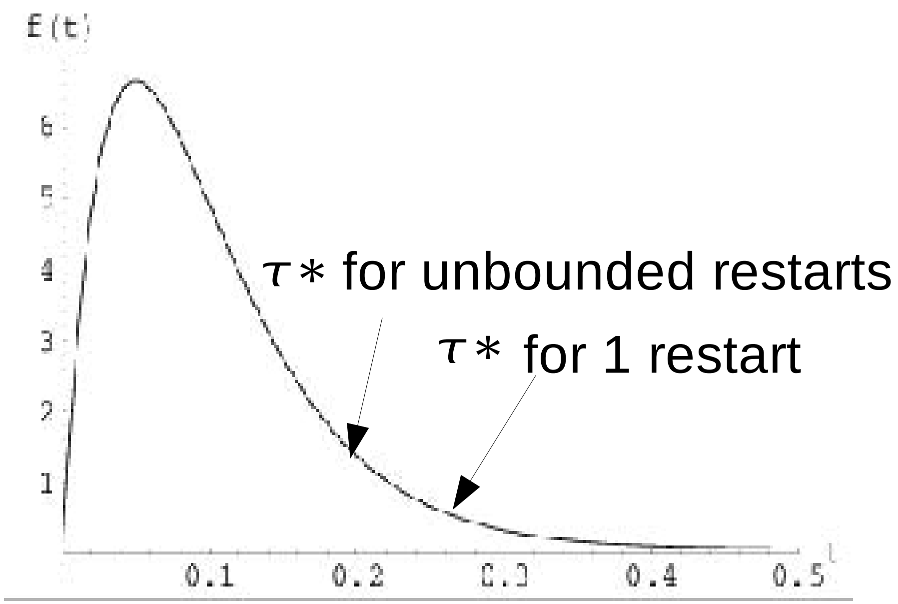
correlation for same URL

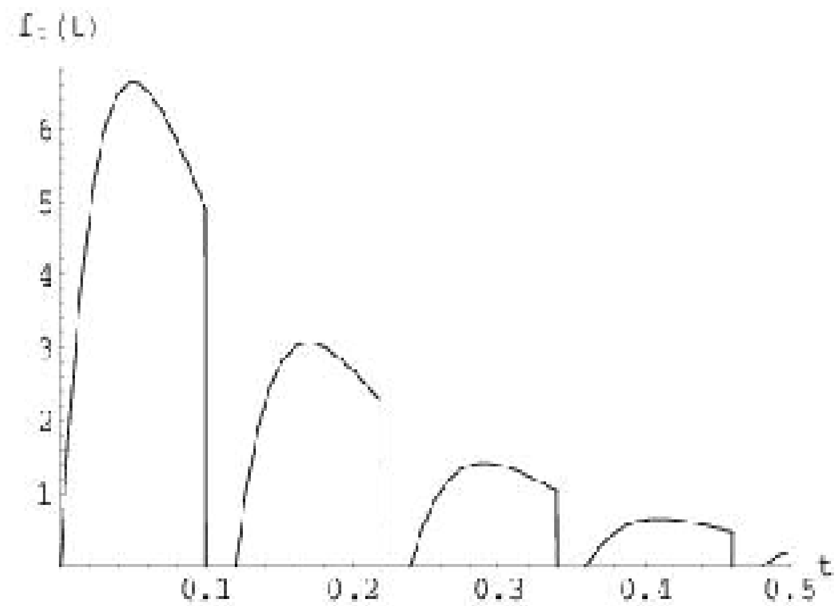
completion time distribution

T ~ hyper/hypo-exponential

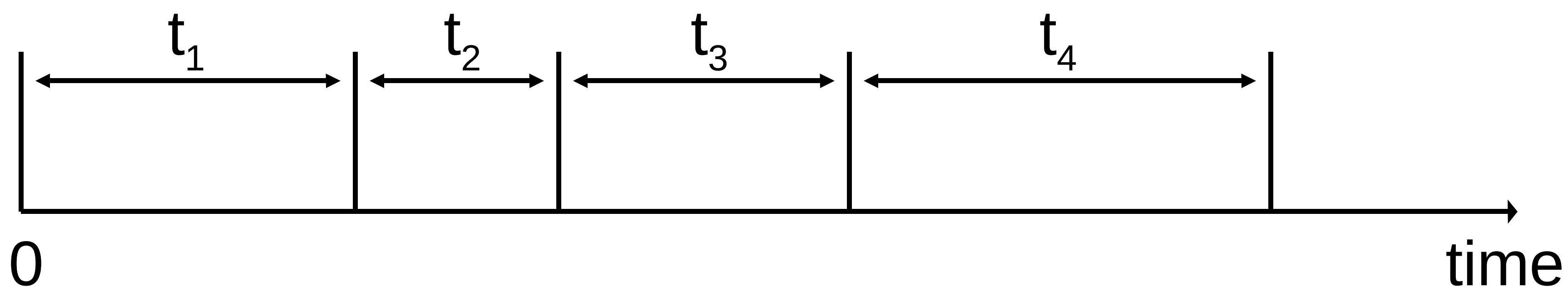
$T = 0.9\ T_1 + 0.1\ T_2$

T_1 = Erlang(2) lambda = 20, T_2 = Erlang(2) lambda = 2





problem formulation

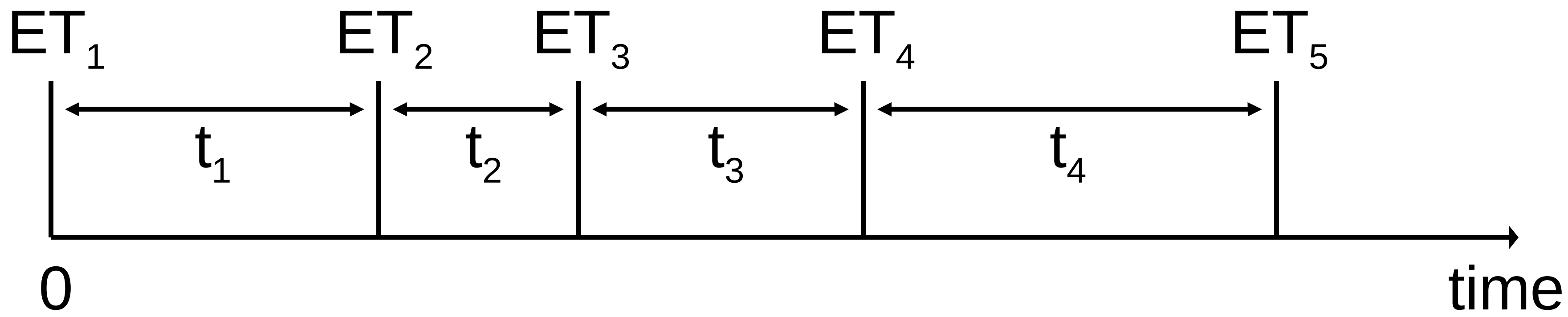


which values t_1, t_2, t_3, …minimise completion time?

	first moment	**higher moments**
finite # restarts	?	?
infinite # restarts	?	?

first moment

first moment, finite # restarts



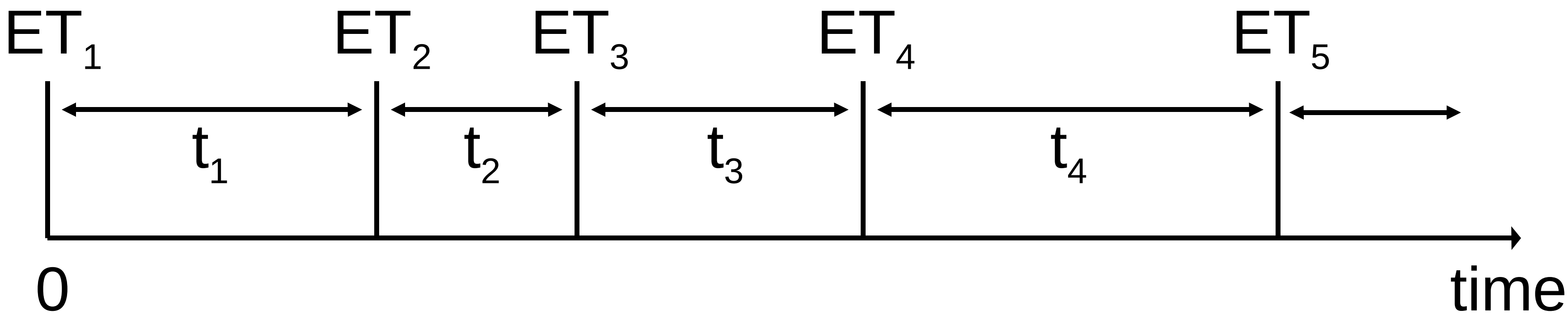
$ET_K = ET$

$ET_{K-1} = F(\, t_{K-1} \,)\, .\, M\, (\, t_{K-1} \,) + (\, 1 – F(\, t_{K-1} \,)\,)\, .\, (\, t_{K-1} \; + ET_K \,)$

backward algorithm:

for k = K, K-1, .., 1 find t_k that minimises ET_k;

first moment, infinite # restarts



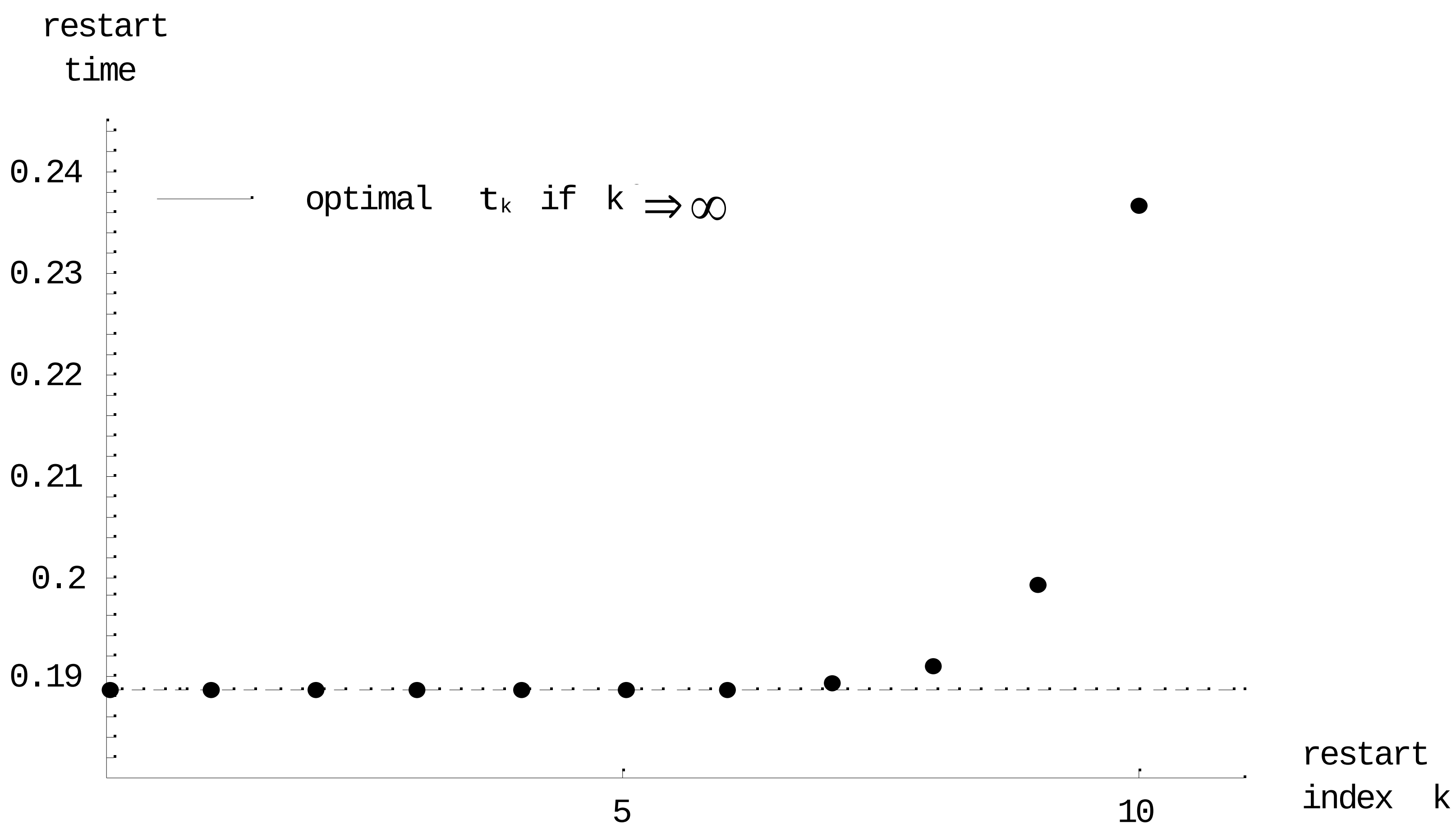
$ET_1 = ET_2 = \ldots = ET_{K-1} = ET_K = \ldots = ET_\infty$

constant optimal restart time $t_1 = t_2 = t_3 = \ldots = t_\infty$

find t_∞ that minimises

$ET_\infty = M(t_\infty) + (1 - F(t_\infty))\,.\,t_\infty\,)\,/\,F(t_\infty)$

first moment: optimal restart times

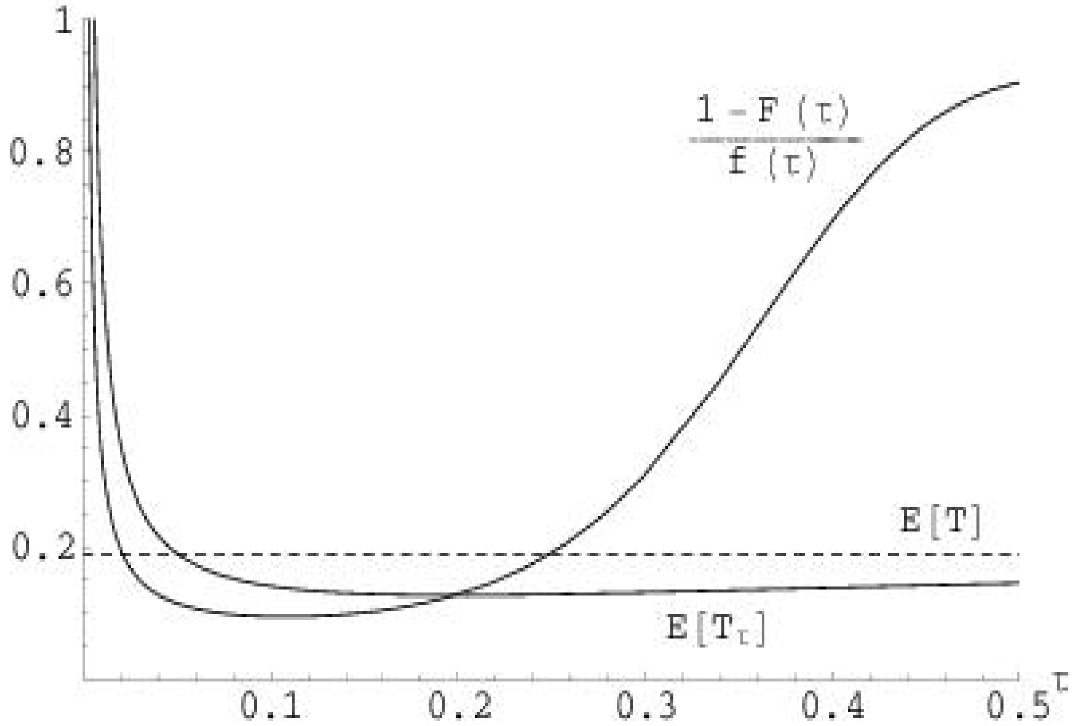


Lognormal distribution, mu = -2.31, sigma = 0.97

optimal restart time for unbounded restarts

inverse hazard rate

$$\frac{1 - F(\tau*)}{f(\tau*)} = E[\tau*] + c$$



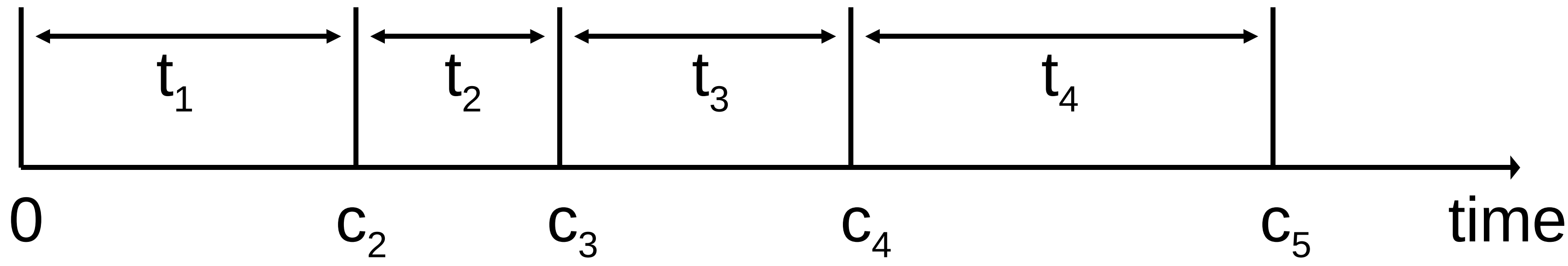
c = 0

results for first moment

	first moment	higher moments
finite # restarts	backward algorithm K optimisations	?
infinite # restarts	constant 1 optimisation	?

higher moments

sensitivity to shift



$$E(c + T) = c + ET$$

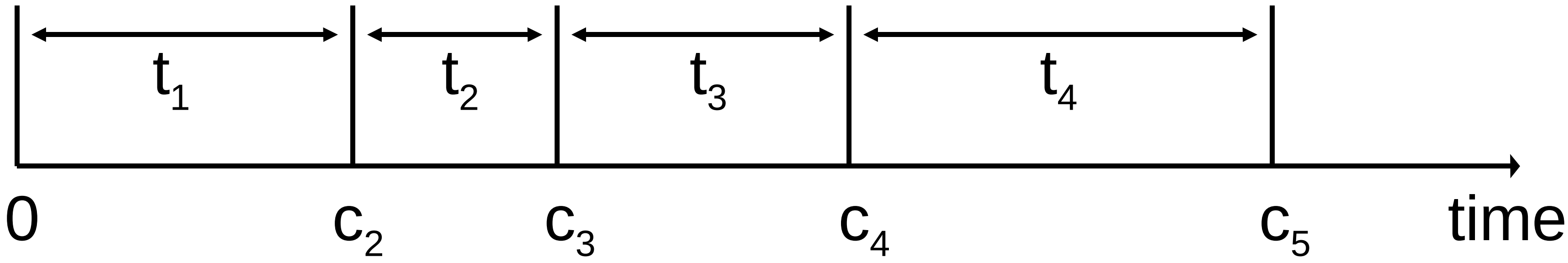
i.e., minimising $E(c + T)$ = minimising ET

but:

$$E(c + T)^2 = c^2 + 2.c.ET + ET^2$$

i.e., minimising $E(c + T)^2$!= minimising ET^2

sensitivity to shift



first moment, optimal t_k independent of earlier restarts

- finite: backward algorithm works
- infinite: constant

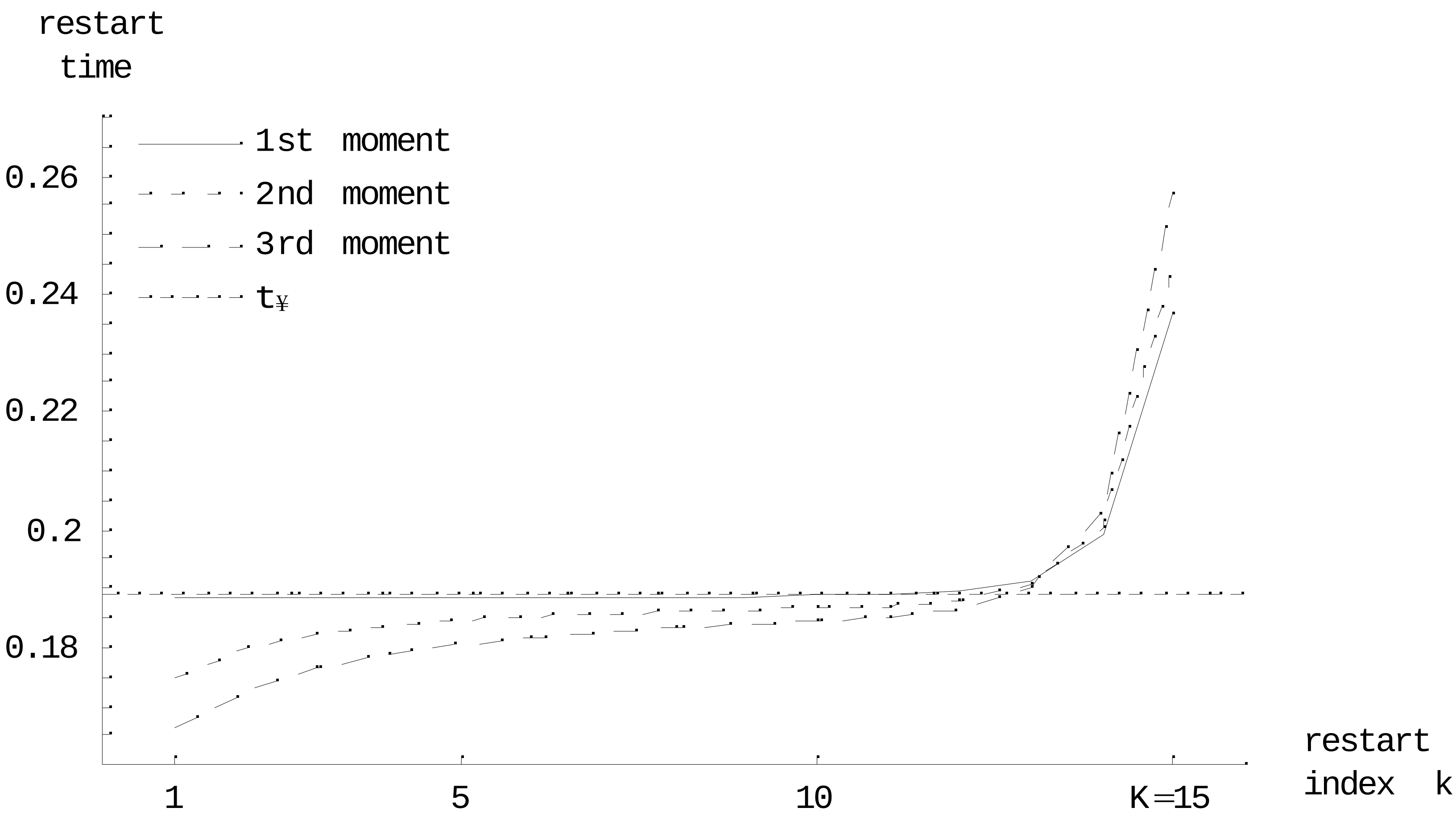
higher moments:

- finite: backward algorithm fails
- infinite: not constant (btw, papers have claimed the opposite)

luckily, optimal restart times for higher moments only depend on the *sum* of preceding restarts

resolution: backward/forward algorithm

first three moments: optimal restart times



Lognormal distribution, mu = -2.31, sigma = 0.97

results for all moments

	first moment	**higher moments**
finite # restarts	backward algorithm K optimisations	backward/forward iteration N.K optimisations
infinite # restarts	constant 1 optimisation	first moment approximation & backward/forward at start and end $N_1.K_1 + N_2.K_1$

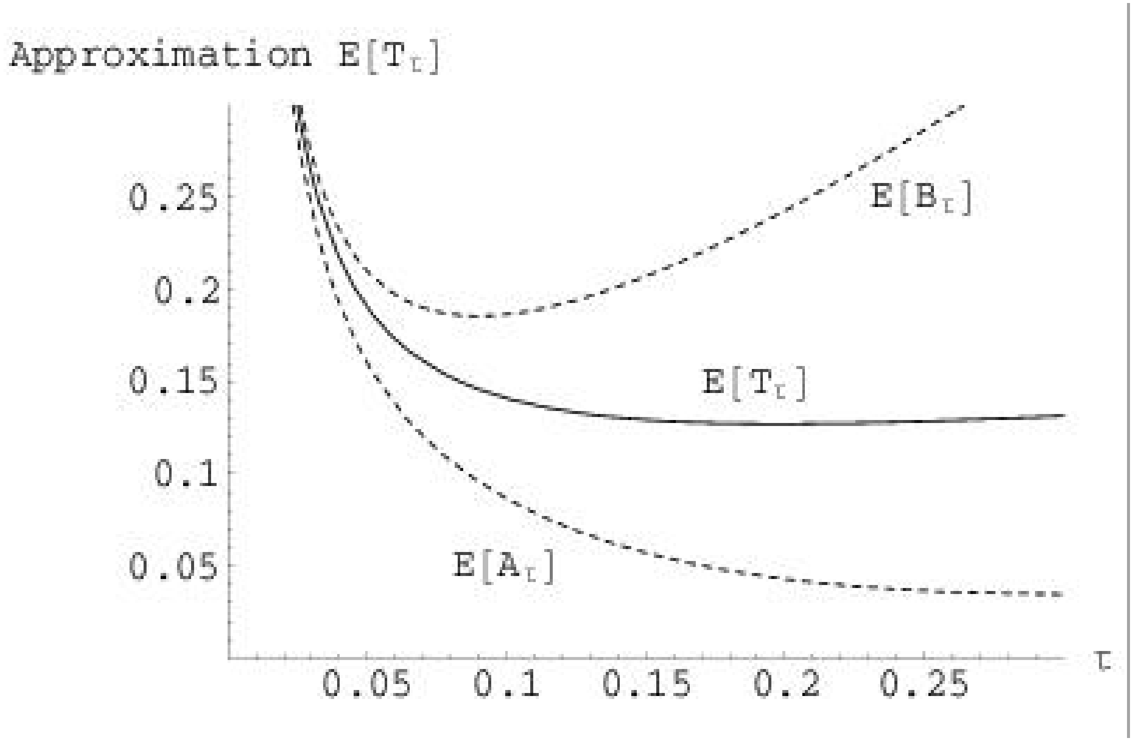
results for higher moments in UKPEW 2004

properties and related issues

geometric approximation for all moments

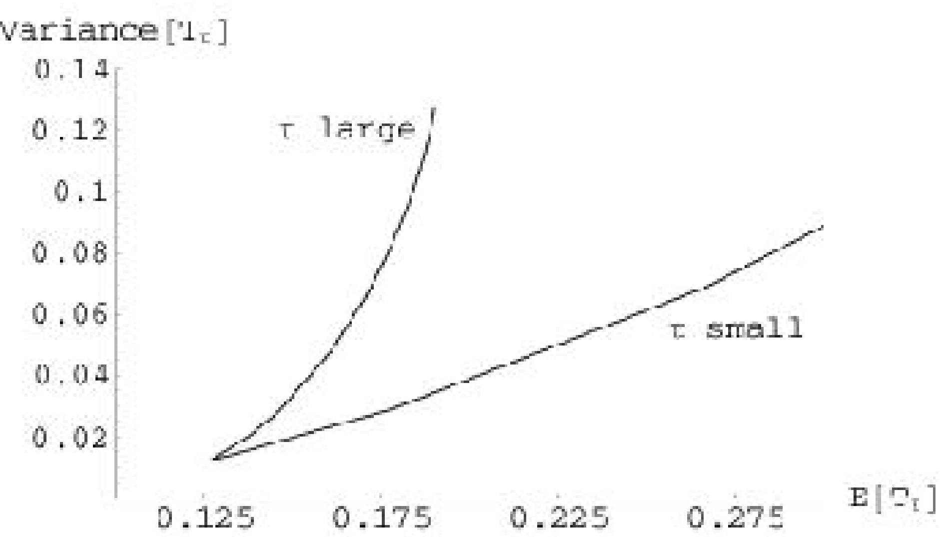
restart is bernoulli trial, approximate number of trials before completion

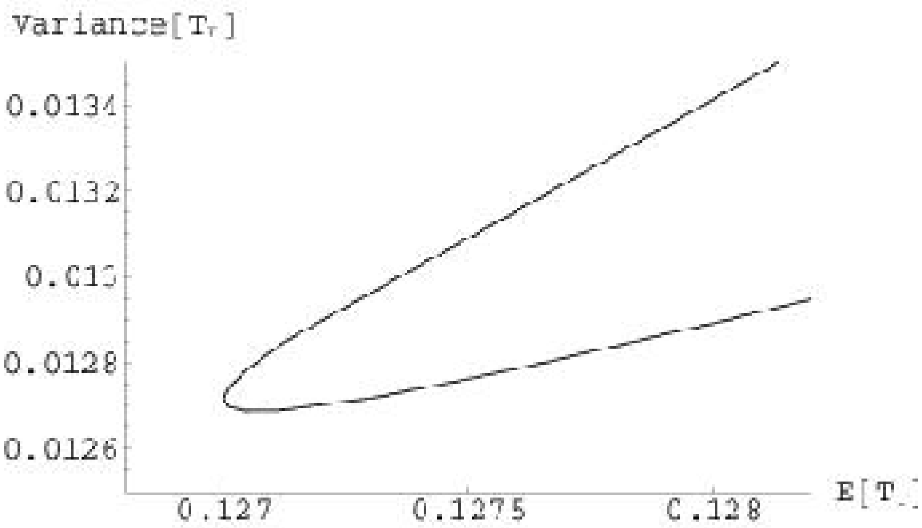
$$A_\tau = k(\tau + c) + 0 \quad \text{with prob.} \quad (1 - F(\tau))^k F(\tau)$$
$$B_\tau = k(\tau + c) + \tau \quad \text{with prob.} \quad (1 - F(\tau))^k F(\tau)$$



cusp point

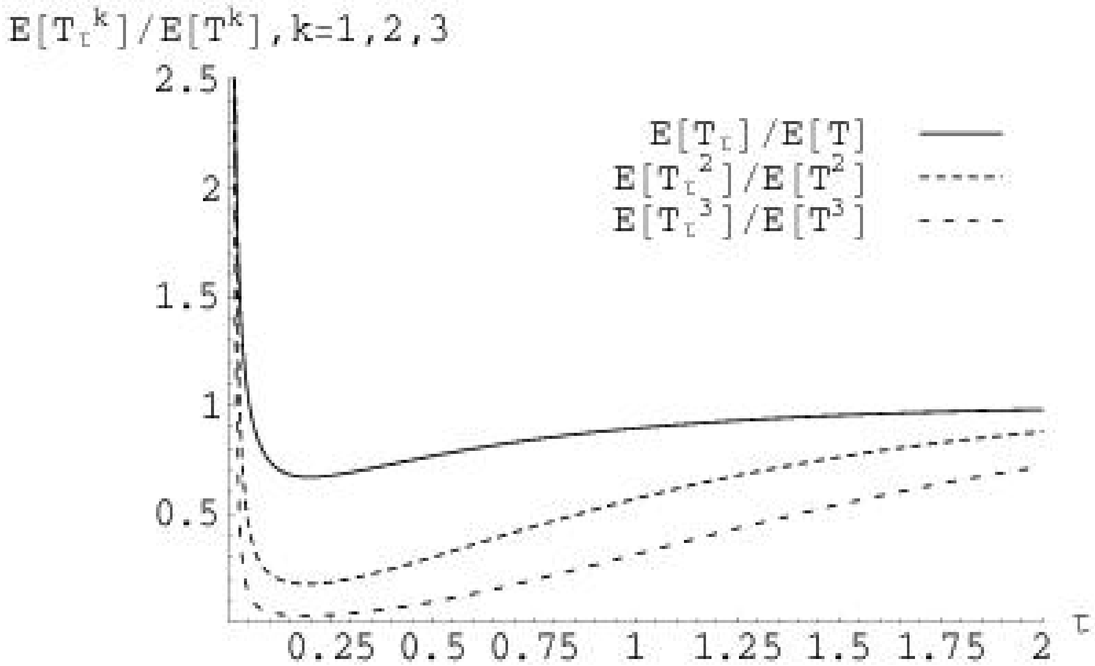
unbounded number of restarts





in general optimum for mean and variance is not the same

length of restart interval

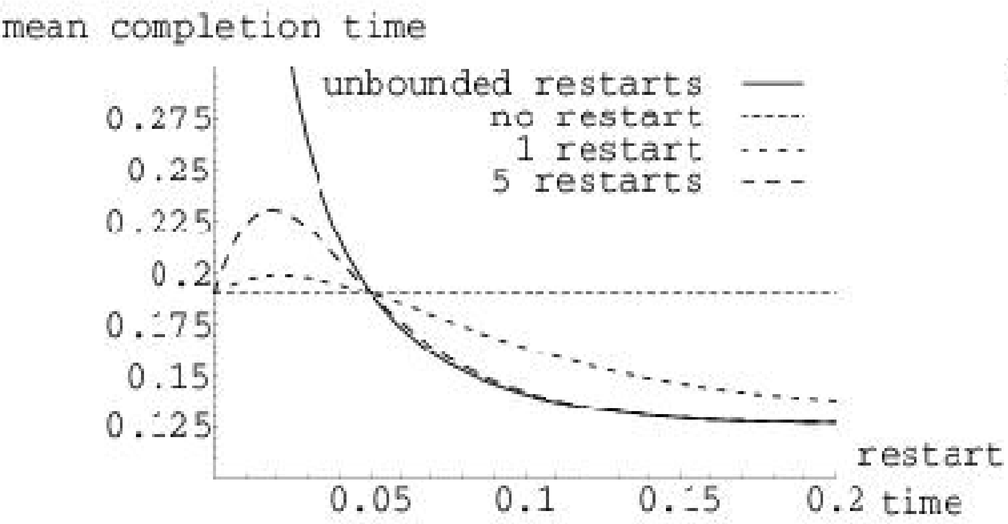


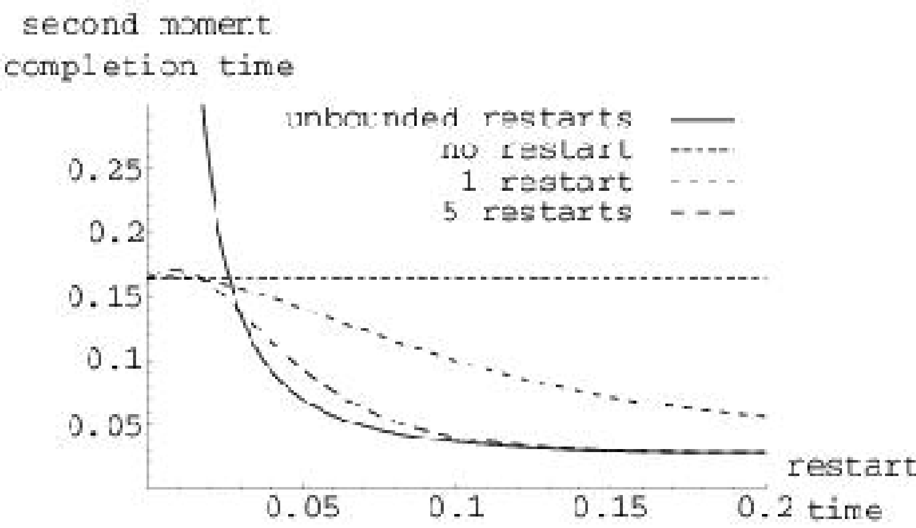
- a rough estimate should rather be a little bigger
- more gain for higher moments
- the more restarts the better:

$$E[T_\tau] < E[T] \;\Leftrightarrow\; E[T_\tau] < E[T_{\tau^{K+1}}] < E[T_{\tau^K}] < \ldots < E[T], \quad K \geq 1$$

- infinite moments become finite under restart

number of restarts





For the first moment there is one point, in which the number of restarts makes no difference. For higher moments there is no such point.

conclusion

	first moment	higher moments
finite # restarts	backward	forward/backward iteration
infinite # restarts	constant	first moment approximation & forward/backward

- higher moments benefit more than first moment
- infinite moments become finite under restart
- choose restart times not too small
- resolved problem of sensitivity to shifts with forward/backward