Universal performance bounds of restart

Dmitry Starkov and Sergey Belan

Landau Institute for Theoretical Physics

Landau Week 2023

Outline

- Introduction
  - restart-induced speed up of process completion
  - applications: computer science and chemical kinetics
  - mathematical model, analytical methods and key rigorous results
- Main part
  - performance bounds of restart
  - how good is the best restart protocol?
  - novel sufficient condition for restart to be effective
- Conclusion and outlook

Random processes under restart

Restart - interruption of a process followed by starting a new statistically independent realization.

Motivation to restart: a proper restart policy could decrease the characteristic time scale of process completion (mean, median or mode) and/or increase the deadline meeting probability.

Key application: restart-induced speed up of randomized computer algorithms

[1]***Alt et.al. A method for obtaining randomized algorithms with small tail probabilities, Int. Comp. Sci. Inst. (1991)***
[2]***Luby et.al. Optimal speedup of Las Vegas algorithms, Inf. Proc. Lett. (1993)***

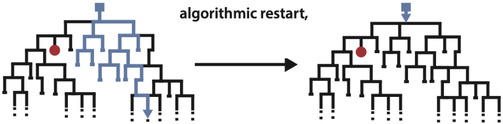
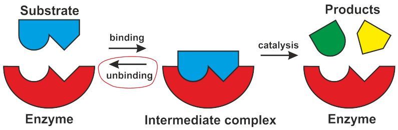


Рис. 1: (from Pal & Reuveni PRL (2017)) A computer algorithm randomly scans a tree of possibilities in search of a solution. When a given number of steps elapses without being able to solve the problem, the algorithm is restarted to avoid prolonged wandering in the region of the search space far from the actual solution

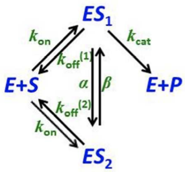
Another possible application: enzymatic reactions

Simple Michaelis–Menten kinetics assumes that the enzyme–substrate complex has only one conformation that decays as a single exponential. As a consequence, the enzymatic velocity decreases as the dissociation rate constant of the complex increases.



Recent theoretical works showed that it is possible for the enzymatic velocity to increase when the unbinding rate k_{off} becomes higher, if the enzyme–substrate complex has many conformations characterised by different catalysis rates.

[3] Reuveni et al. Proc. Natl. Acad. Sci. (2014)
[4] Berezhkovskii et al. JCP (2017)



Barrier to practical use: conceivable methods to control unbinding rate k_{off} (temperature change, pH change, structure modifications of enzyme) affects also other kinetic constant (binding rate k_{on}, catalysis rate k_{cat}).

Model formulation

Assumptions:
– the process halts with unit probability
– restart mechanism is uncoupled from the process internal dynamics
– each restart event has negligible duration

Key ingredients of model:
T - random completion time of stochastic process;
$P(T)$ - probability density of random variable T;
$\mathcal{R} = \{\tau_1, \tau_2, \dots\}$ - restart protocol characterised by a (possibly infinite) sequence of inter-restarts time intervals.

If the process is completed prior to the first restart event, the story ends there. Otherwise, the process will start from scratch and begin anew. Next, the process may either complete prior to the second restart or not, with the same rules. This procedure repeats until the process finally reaches completion.

Metric of interest: mean completion time $\langle T_{\mathcal{R}} \rangle$ in the presence of protocol $\mathcal{R}$.

Simple case studies

- Lévy-Smirnov distribution (first-passage time density for 1d diffusion)
  Completion time probability density: $P(T) = \frac{L}{2\sqrt{\pi D T^3}} \exp(-\frac{L^2}{4DT})$
  (L - initial distance to a target; D - diffusion constant)
  In the absence of restart mean search time diverges: $\langle T \rangle \to \infty$.
  Poisson restart at rate r renders expected search time finite:
  $\langle T_r \rangle = \frac{1}{r}(\exp\sqrt{\frac{rL^2}{D}} - 1)$ (***Evans & Majumdar, PRL 2011***). What is more, optimal restart rate $r^* \approx 2.54D/L^2$ brings $\langle T_r \rangle$ to minimum.

- Double-exponential distribution
  Completion time probability density: $P(T) = p\alpha_1 e^{-\alpha_1 T} + (1-p)\alpha_2 e^{-\alpha_2 T}$
  The mean completion time in the absence of restart: $\langle T \rangle = \frac{p}{\alpha_1} + \frac{1-p}{\alpha_2}$.
  Poisson restart at rate r reduces the mean completion time:
  $\langle T_r \rangle = \frac{(1-p)\alpha_1 + p\alpha_2 + r}{\alpha_1\alpha_2 + (p\alpha_1 + (1-p)\alpha_2)r} < \langle T \rangle$. Optimal restart rate $r^* \to +\infty$ brings $\langle T_r \rangle$ to minimum.

Analytical method: renewal approach

The random completion time $T_{\mathcal{R}}$ of the process in the presence of restart protocol $\mathcal{R} = \{\tau_1, \tau_2, \dots\}$ obeys the infinite chain of equations

$$T_{\mathcal{R}} = T_1 \cdot I[T_1 < \tau_1] + (\tau_1 + T_1^{\text{res}}) \cdot I[T_1 \geq \tau_1], \tag{1}$$

$$T_1^{\text{res}} = T_2 \cdot I[T_2 < \tau_2] + (\tau_2 + T_2^{\text{res}}) \cdot I[T_2 \geq \tau_2], \tag{2}$$

$$\dots$$

$$T_n^{\text{res}} = T_n \cdot I[T_n < \tau_n] + (\tau_{n+1} + T_{n+1}^{\text{res}}) \cdot I[T_{n+1} \geq \tau_{n+1}], \tag{3}$$

$$\dots$$

where
$T_1, T_2, \dots$ - independent random variables sampled from the probability distribution $P(T)$;
$I[\dots]$ - indicator random variable which is equal to unity the inequality in its argument is justified and is zero otherwise;

T_i^{res} - the i-th residual time, i.e. the time remaining to the process completion just after the i-th restart.

Performing averaging over the statistics of random variables $T_1, T_2, \dots$ one obtains

$$\langle T_{\mathcal{R}} \rangle = \sum_{k=1}^{\infty} \left(\frac{\int_0^{\tau_k} dT P(T) T}{\int_{\tau_k}^{\infty} dT P(T)} + \tau_k \right) \prod_{i=1}^{k} \int_{\tau_i}^{\infty} dT P(T) \tag{4}$$

periodic restart with period τ: $\langle T_\tau \rangle = \frac{\int_0^\tau P(T) T dT + \tau \int_\tau^\infty P(T) dT}{\int_0^\tau P(T) dT}$

Poisson restart at rate r: $\langle T_r \rangle = \frac{1 - \tilde{P}(r)}{r \tilde{P}(r)}$ (where $\tilde{P}(r)$ - Laplace transform of $P(T)$)

Some general results

- *When is the restart method effective?*

  Simple sufficient conditions for existence of effective restart strategy:
  – power-law tail of completion time density $P(T)$;
  – $\frac{\sigma(T)}{\langle T \rangle} > 1$, where $\sigma(T)$ - is the standard deviation of the random completion time in the absence of restart (***Reuveni PRL 2016***).

- *Which protocol is the most effective?*

  Optimally tuned periodic restart beats any other restart strategy: if you found a value $\tau_* \geq 0$ (probably $\tau_* = +\infty$) such that $\langle T_{\tau_*} \rangle \leq \langle T_\tau \rangle$ for any $\tau \geq 0$, then $\langle T_{\tau_*} \rangle \leq \langle T_{\mathcal{R}} \rangle$ for all $\mathcal{R}$ (***Luby et al. 1993, Lorenz 2021***).

Research question 1: performance limit of restart

What, if any, are the fundamental limitations of the optimization via restart?

We are wondering if the mean performance of restart is bounded from below by some simple statistical characteristics of the original process. More specifically, we seek to derive inequality of the form

$$\boxed{\langle T_{\mathcal{R}} \rangle \geq C\mathcal{T}} \quad (5)$$

where
$T_{\mathcal{R}}$ - random completion time of the generic stochastic process under arbitrary restart protocol $\mathcal{R}$;
$\mathcal{T}$ - some time scale expressed through the statistical moments, quantiles or mode of the probability density $P(T)$;
C - universal positive constant which depends neither on specific form of $P(T)$ nor on the particular restart schedule $\mathcal{R}$.

No-go results

Previous works have shown the importance of relative fluctuation σ/μ, where $\mu = \langle T \rangle$ and $\sigma = \sqrt{\langle T^2 \rangle - \langle T \rangle^2}$, for the analysis of the potential response of stochastic process to restart. Namely, the inequality $\sigma/\mu > 1$ represents a sufficient condition for the existence of a restart protocol that reduces the expected completion time.

Given this result, let us first find out if knowledge of the mean value μ and the standard deviation σ allows one to write a lower bound on the average performance of restart. Consider, probability density $P(T) = p\delta(T - t_1) + (1 - p)\delta(T - t_2)$, where $0 \leq t_1 \leq t_2$ and $0 \leq p \leq 1$. Putting $t_2 = \frac{\mu^2+\sigma^2}{\mu}$, $p = \frac{\sigma^2}{\mu^2+\sigma^2}$, $\tau = t_1 + 0$ and $t_1 \to 0$, one immediately obtains from the relation

$$\langle T_\tau \rangle = \frac{\int_0^\tau P(T)TdT + \tau \int_\tau^\infty P(T)dT}{\int_0^\tau P(T)dT}, \tag{6}$$

that $\langle T_\tau \rangle = t_1/p \to 0$. We see that for the fixed values of μ and σ, the completion time $\langle T_\tau \rangle$ can be arbitrarily small. Therefore, the pair (μ,σ) does not produce any non-trivial lower bound.

Universal performance bound

Restart performance is limited to a quarter of the harmonic mean completion time:

$$\boxed{\langle T_{\mathcal{R}} \rangle \geq \frac{1}{4} h} \tag{7}$$

where $h = \langle T^{-1} \rangle^{-1}$.

<u>Proof</u>:
Let τ_* be the best period of regular restart protocol for a given stochastic process, i.e.

$$\langle T_{\tau_*} \rangle \leq \langle T_\tau \rangle \tag{8}$$

for any $\tau \geq 0$. Luby et al. 1993 and Lorenz 2021 proved that $\langle T_{\tau_*} \rangle \leq \langle T_{\mathcal{R}} \rangle$ for any $\mathcal{R}$. Also, as shown by the same authors the mean performance of an optimal periodic restart obeys the condition

$$\langle T_{\tau_*} \rangle \geq \frac{1}{4} \min_{\tau} \frac{\tau}{Pr[T \leq \tau]}. \tag{9}$$

Applying Markov's inequality to the variable $\omega = 1/T$ we find
Pr$[T \leq \tau] =$ Pr$[\omega \geq \frac{1}{\tau}] \leq \tau \langle \omega \rangle = \tau \langle \frac{1}{T} \rangle$. Next, taking into account Eq. (9) one obtains $\langle T_{\tau_*} \rangle \geq \frac{1}{4} h$, where $h = \langle T^{-1} \rangle^{-1}$ is the harmonic mean completion time of the original process. And finally, since $\langle T_{\mathcal{R}} \rangle \geq \langle T_{\tau_*} \rangle$ for any $\mathcal{R}$, this yields (7). No constraints have been imposed on the form of $P(T)$, and, therefore, Eq. (7) is universally valid for any setting.

Particular case of smooth unimodal distribution

Somewhat less general, but still informative, result can be obtained if we assume that the completion time distribution $P(T)$ is smooth and exhibits single local maximum at some non-zero value of T. The efficiency of any restart protocol in this case satisfies the inequality

$$\boxed{\langle T_{\mathcal{R}} \rangle \geq \frac{1}{4} M} \qquad (10)$$

where $M = \text{argmax}_T P(T) > 0$ is the mode of the probability distribution $P(T)$, i.e. the value of the random completion time T that occurs most frequently.

Proof:
To prove Eq. (10) let us introduce $\tau_0 \equiv \text{argmin}_\tau \frac{\tau}{\text{Pr}[T \leq \tau]}$. Clearly, assumption $M > 0$ implies that $\tau_0 > 0$. Since the smooth function $f(\tau) = \frac{\tau}{\text{Pr}[T \leq \tau]}$ attains its minimal value at $\tau = \tau_0$, one obtains $df(\tau_0)/d\tau = 0$ or, equivalently, $P(\tau_0)\tau_0 = \int_0^{\tau_0} P(T)dT$. Next, as the unimodal function $P(T)$ is non-decreasing on the interval form 0 to M, this extrema condition implies the inequality $\tau_0 \geq M$ and, therefore, $\frac{\tau_0}{\text{Pr}[T \leq \tau_0]} \geq \frac{M}{\text{Pr}[T \leq \tau_0]} \geq M$. Together with Eq. (9) this yields inequality $\langle T_{\tau_*} \rangle \geq \frac{1}{4} M$. Recalling that $\langle T_{\mathcal{R}} \rangle \geq \langle T_{\tau_*} \rangle$ for all $\mathcal{R}$, we then obtain Eq. (10). Note also, that if the probability distribution $P(T)$ has multiple local maxima, then $\langle T_{\mathcal{R}} \rangle \geq \frac{1}{4} M^{min}$, where M^{min} is the leftmost mode.

Research question 2: How good is the best protocol?

Next let us construct an inequality of the form

$$\boxed{\langle T_{\tau_*} \rangle \leq C\mathcal{T}} \tag{11}$$

where
$\langle T_{\tau_*} \rangle$ - expected completion time at optimal restart period τ_*;
$\mathcal{T}$ - some time scale determined by the original stochastic process;
C - universal positive constant which depends neither on specific form of $P(T)$ nor on the optimal restart period τ_*.

No-go result:
It is easy to understand that the upper bound limit on optimal performance cannot be expressed via the harmonic mean h or the mode M. Indeed, for the half-normal distribution $P(T) = \sqrt{\frac{2}{\pi\sigma^2}} e^{-\frac{T^2}{2\sigma^2}}$ one has $\tau_* = +\infty$, so that $\langle T_{\tau_*} \rangle = \langle T \rangle > 0$, whereas $h = M = 0$. Therefore, inequalities of the form $\langle T_{\tau_*} \rangle \leq C_1 h$ and $\langle T_{\tau_*} \rangle \leq C_2 M$, where C_1 and C_2 are positive constants, cannot be universally valid.

Performance bound for optimal restart

The desired universal upper bound can be expressed in terms of the median completion time m of the original process obeying by definition the equation $Pr[T \leq m] = 1/2$. Namely

$$\boxed{\langle T_{\tau_*} \rangle \leq 2m} \tag{12}$$

Thus, no matter how heavy the tails of $P(T)$ are, in the presence of an optimally tuned periodic restart, the average completion time does not exceed twice the median of the unperturbed process.

Proof:
Taking into account that $\langle T_{\tau_*} \rangle \leq \langle T_\tau \rangle$ for any $\tau \geq 0$ together with the inequality $\langle T_\tau \rangle = \frac{\int_0^\tau P(T) T dT + \tau \int_\tau^\infty P(T) dT}{\int_0^\tau P(T) dT} \leq \frac{\tau}{Pr[T \leq \tau]}$, we find $\langle T_{\tau_*} \rangle \leq \frac{\tau}{Pr[T \leq \tau]}$.
Substituting m for τ in the last inequality one obtains (12).

Importantly, the bound dictated by Eq. (12) is sharp. Indeed, for $P(T) = \frac{1}{2}\delta(T - t) + \frac{1}{2}\delta(T - 3t)$ one obtains $m = t$ and $\langle T_{\tau_*} \rangle = 2t$, where $\tau_* = t$.

Yet another no-go result

Given this result, it is natural to ask if the median value can be used to construct the bottom bound of restart performance in the spirit of Eqs. (7) and (10). The answer is no. A simple counterexample demonstrating that the inequality $\langle T_{\mathcal{R}} \rangle \geq Cm$, where C is universal non-zero constant, cannot be valid is given by the Weibull distribution $P(T) = \frac{k}{\lambda^k} T^{k-1} e^{-(\frac{T}{\lambda})^k}$ with $0 < k < 1$ for which $\langle T_{\tau_*} \rangle = 0$, where $\tau_* \to 0$, and $m > 0$.

Checking consistency

Inequalities $\langle T_{\mathcal{R}} \rangle \geq \frac{1}{4}h$ and $\langle T_{\tau_*} \rangle \leq 2m$ do not contradict each other since

$$\boxed{h \leq 2m} \tag{13}$$

for any probability density $P(T)$. Indeed, applying Markov's inequality we find $1/2 = \Pr[T \leq m] = \Pr[\omega \geq \frac{1}{m}] \leq \langle \omega \rangle m = T_{1/2}/h$, where $\omega = 1/T$.

Also, $\langle T_{\tau_*} \rangle \leq 2m$ does not contradict $\langle T_{\mathcal{R}} \rangle \geq \frac{1}{4}M$ since

$$\boxed{M \leq 2m} \tag{14}$$

for any continuous unimodal probability density $P(T)$. To prove this, assume the contrary, that is, let $M > 2m$ be true for some non-negative random variable T. By the virtue of definition we have $\int_0^m P(T)dT = \int_m^{+\infty} P(T)dT = 1/2$. Further, it follows from the definition of the mode that the function $P(T)$ is non-decreasing on the interval $[0, M]$. Then, on the one hand $\int_m^M P(T)dT > \int_m^{2m} P(T)dT \geq \int_0^m P(T)dT = 1/2$, and on the other $\int_m^M P(T)dT \leq \int_m^{+\infty} P(T)dT = 1/2$. We got a contradiction.

Numerical results

For the sake of illustration we explored several probability distributions $P(T)$, whose response to restart has been extensively discussed in the physical and computer science literature.

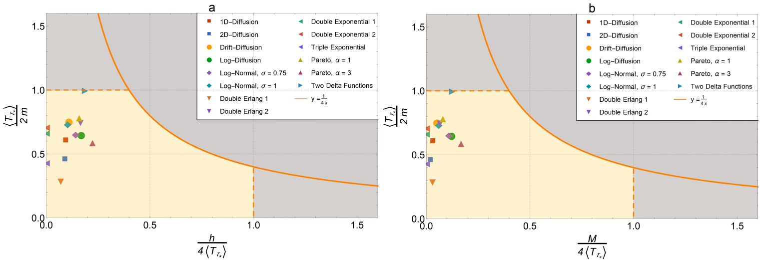


Рис. 2: **In accordance with our analytical predictions all points belong to the light orange region determined by the conditions** $0 \leq x \leq 1$, $0 \leq y \leq 1$, $y \leq 1/(4x)$.

Beyond the mean performance

Inequality constraints derived above can be generalized to higher order statistical moments of random completion time. First of all, since $\sqrt[k]{\langle T_{\mathcal{R}}^k \rangle} \geq \langle T_{\mathcal{R}} \rangle$ for any natural k due to Jensen's inequality, we immediately find from relation $\langle T_{\mathcal{R}} \rangle \geq \frac{1}{4} h$ that $\sqrt[k]{\langle T_{\mathcal{R}}^k \rangle} \geq \frac{1}{4} h$ for a generic stochastic process under an arbitrary restart protocol.

A similar extension of relation $\langle T_{\tau_*} \rangle \leq 2m$ is more tricky. It turns out that the statistical moments of the optimal completion time T_{τ_*} satisfy the inequality

$$\boxed{\sqrt[k]{\langle T_{\tau_*}^k \rangle} \leq 2 \sqrt[k]{k!} m.} \tag{15}$$

Proof:
To prove Eq. (15) let us assume that the process, which is being restarted periodically in an optimal way, becomes subject to an additional restart protocol $\mathcal{R}_\Gamma$ characterized by random restart-intervals $\tau_1, \tau_2, \dots$ independently sampled from Gamma distribution $\rho(\tau) = \frac{\beta^k}{\Gamma(k)} \tau^{k-1} e^{-\beta\tau}$ with shape parameter k and infinitesimally small rate parameter β. This produces a deferential correction $\langle T_{\tau_* + \mathcal{R}_\Gamma} \rangle - \langle T_{\tau_*} \rangle \approx \frac{1}{k!} \left(\langle T_{\tau_*} \rangle \langle T_{\tau_*}^k \rangle - \frac{1}{k+1} \langle T_{\tau_*}^{k+1} \rangle \right) \beta^k$ to the mean completion time attained by the optimal periodic restart. Because of the dominance of a periodic restart over other restart strategies, one can be sure that this difference is positive, and therefore $\langle T_{\tau_*}^k \rangle \leq k! \langle T_{\tau_*} \rangle^k$. Together with relation $\langle T_{\tau_*} \rangle \leq 2m$ this yields Eq. (15).

Potentially non-stopping processes

So far we have assumed that the process terminates in finite time with probability 1. however, our results remain unchanged or require a trivial modification when one introduces the non-zero probability of never stopping.

Namely, once the probability $q = Pr[T < +\infty]$ of halting for a given stochastic process is known, one can improve the estimate $\langle T_{\mathcal{R}} \rangle \geq \frac{1}{4} h$ as follows:

$$\boxed{\langle T_{\mathcal{R}} \rangle \geq \frac{1}{4} \frac{h}{q}} \quad (16)$$

Also, inequality $\langle T_{\mathcal{R}} \rangle \geq \frac{1}{4} M$ holds even for potentially non-stopping processes, with the obvious caveat that M should now be considered as the most frequent completion time of halting trials.

Finally, if the process has non-zero probability of never halting, the mean performance of optimal periodic restart obeys

$$\boxed{\langle T_{\tau_*} \rangle \leq 2 \frac{m_s}{q}} \quad (17)$$

where m_s denotes the median completion times of the halting trials.

Conclusion

We derived a range of statistical inequalities that offer constraints on the effect that restart could impose on the completion time of a generic stochastic process.

Publication: Starkov & Belan PRE 2023

Key results:

- Restart performance is limited to a quarter of the harmonic mean completion time: $\langle T_{\mathcal{R}} \rangle \geq \frac{1}{4} h$
- For the tasks with smooth unimodal completion time distribution one obtains: $\langle T_{\mathcal{R}} \rangle \geq \frac{1}{4} M$
- The twice median sets upper bound on the optimized mean completion time: $\langle T_{\tau_*} \rangle \leq 2m$
- The statistical moments of completion time at optimal restart conditions obey: $\sqrt[k]{\langle T_{\tau_*}^k \rangle} \leq 2 \sqrt[k]{k!} m$
- Novel sufficient condition for restart to be efficient: $\langle T \rangle > 2m$

Some open questions:

- Our analyses does not answer the question of whether the bounds $\langle T_{\mathcal{R}} \rangle \geq \frac{1}{4} h$ and $\langle T_{\mathcal{R}} \rangle \geq \frac{1}{4} M$ are sharp.
- How does accounting for non-zero time penalty T_p for restart modifies the bounds constructed here? $\langle T_{\tau_*} \rangle \leq 2(m + T_p)$.
- Performance bounds for restart-induced optimization of other metrics: median completion time, deadline meeting probability, etc.

Acknowledgments: The work was supported by the Russian Science Foundation, project no. 22-72-10052.

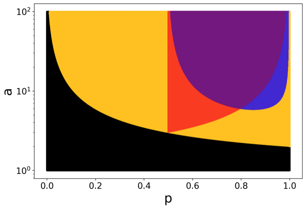
Useful corollary: novel criterion of restart efficiency

As follows from relation $\langle T_{\tau_*} \rangle \leq 2m$ the inequality

$$\boxed{\langle T \rangle > 2m} \tag{18}$$

guarantees that there exists finite restart period decreasing the expected completion time.

What is particularly interesting is that this simple inequality makes it possible to capture the benefit of restarting in those cases when analysis of the relative fluctuation cannot. In figure below we compare the applicability of two criteria, $\langle T \rangle / m > 2$ and $\sigma(T)/\langle T \rangle > 1$, using the mix of two delta-functions $P(T) = p \cdot \delta(T - t_1) + (1-p) \cdot \delta(T - t_2)$ as a model distribution.



$a = t_2/t_1$
Black area: restart is not efficient.
Purple area: restart is efficient and both criteria, $\sigma(T) > \langle T \rangle$ and $\langle T \rangle > 2m$, are fulfilled.
Blue area: restart efficiency is captured only by the inequality $\sigma(T) > \langle T \rangle$.
Red area: only the condition $\langle T \rangle > 2m$ is satisfied.
Orange area: neither of the two sufficient conditions is met, but restart is useful nevertheless.

Optimal property of periodic restart

Let τ_* be the best period of regular restart protocol for a given stochastic process, i.e.

$$\langle T_{\tau_*} \rangle \leq \langle T_\tau \rangle \tag{19}$$

for any $\tau \geq 0$. Also, let $\mathcal{R}_* = \{\tau_1^*, \tau_2^*, \dots\}$ be an optimal restart protocol for the same process. By the definition of optimal protocol this means that

$$\langle T_{\mathcal{R}_*} \rangle \leq \langle T_{\mathcal{R}} \rangle \tag{20}$$

for any $\mathcal{R}$. Below we show that $\langle T_{\tau_*} \rangle = \langle T_{\mathcal{R}_*} \rangle$.
The random completion time in the presence of restart events scheduled accordingly to the protocol $\mathcal{R}_*$ can be represented as

$$T_{\mathcal{R}_*} = TI(T < \tau_1^*) + (\tau_1^* + T_{\mathcal{R}_*'})I(T \geq \tau_1^*) \tag{21}$$

where $\mathcal{R}_*' = \{\tau_2^*, \tau_3^*, \dots\}$. Averaging over the statistics of original process yields

$$\langle T_{\mathcal{R}_*} \rangle = \langle TI(T < \tau_1^*) \rangle + \tau_1^* \langle I(T \geq \tau_1^*) \rangle + \langle T_{\mathcal{R}_*'} \rangle \langle I(T \geq \tau_1^*) \rangle, \tag{22}$$

where we exploited statistical independence of T and $T_{\mathcal{R}_*'}$. Since $\mathcal{R}_*$ is optimal, then $\langle T_{\mathcal{R}_*'} \rangle \geq \langle T_{\mathcal{R}_*} \rangle$, and, therefore,

$$\langle T_{\mathcal{R}_*} \rangle \geq \langle TI(T < \tau_1^*) \rangle + \tau_1^* \langle I(T \geq \tau_1^*) \rangle + \langle T_{\mathcal{R}_*} \rangle \langle I(T \geq \tau_1^*) \rangle, \tag{23}$$

and

$$\langle T_{\mathcal{R}_*} \rangle \geq \frac{\langle TI(T < \tau_1^*) \rangle + \tau_1^* \langle I(T \geq \tau_1^*) \rangle}{\langle I(T < \tau_1^*) \rangle} \tag{24}$$

Comparing right-hand sides of Eqs. (24) and $\langle T_\tau \rangle = \frac{\int_0^\tau P(T) T dT + \tau \int_\tau^\infty P(T) dT}{\int_0^\tau P(T) dT}$, one obtains

$$\langle T_{\mathcal{R}_*} \rangle \geq \langle T_{\tau_1^*} \rangle. \tag{25}$$

From Eqs. (19), (20) and (25) we may conclude that $\langle T_{\mathcal{R}_*} \rangle = \langle T_{\tau_*} \rangle$, and, therefore, $\langle T_{\tau_*} \rangle \leq \langle T_{\mathcal{R}} \rangle$ for any $\mathcal{R}$.

Useful inequality from Luby et al. 1993

As shown by (***Luby et al. 1993, Lorenz 2021***) the mean performance of an optimal periodic restart obeys the condition

$$\langle T_{\tau_*} \rangle \geq \frac{1}{4} \min_{\tau} \frac{\tau}{Pr[T \leq \tau]}. \tag{26}$$

<u>Proof</u>:
If $\langle T_{\tau_*} \rangle \geq \frac{\tau_*}{2Pr[T \leq \tau_*]}$, then the inequality $\langle T_{\tau_*} \rangle \geq \frac{1}{4} \min_\tau \frac{\tau}{Pr[T \leq \tau]}$ is evident.
Next, we assume that $\langle T_{\tau_*} \rangle < \frac{\tau_*}{2Pr[T \leq \tau_*]}$. Let us introduce a variable $\tilde{T} \equiv \min(T, \tau_*)$. It follows from the relation $\langle T_\tau \rangle = \frac{\langle \min(T,\tau) \rangle}{Pr[T \leq \tau_*]}$ that $\langle \tilde{T} \rangle = \langle T_{\tau_*} \rangle Pr[T \leq \tau_*]$. Next, due to the assumption $\tau_* > 2\langle T_{\tau_*} \rangle Pr[T \leq \tau_*]$ and the Markov's inequality $Pr[\tilde{T} > 2\langle \tilde{T} \rangle] \leq \frac{1}{2}$ we find
$Pr[T > 2\langle T_{\tau_*} \rangle Pr[T \leq \tau_*]] = Pr[\tilde{T} > 2\langle T_{\tau_*} \rangle Pr[T \leq \tau_*]] = Pr[\tilde{T} > 2\langle \tilde{T} \rangle] \leq \frac{1}{2}$.
Therefore

$$Pr[T \leq 2\langle T_{\tau_*} \rangle Pr[T \leq \tau_*]] \geq \frac{1}{2}. \tag{27}$$

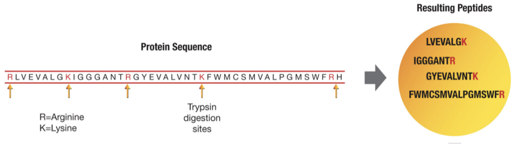
Now, let us denote $t \equiv 2\langle T_{\tau_*} \rangle Pr[T \leq \tau_*]$. It follows from the inequalities (27) and $Pr[T \leq \tau] \leq 1$ that

$$\min_\tau \frac{\tau}{Pr[T \leq \tau]} \leq \frac{t}{Pr[T \leq t]} = \frac{2\langle T_{\tau_*} \rangle Pr[T \leq \tau_*]}{Pr[T \leq 2\langle T_{\tau_*} \rangle Pr[T \leq \tau_*]]} \leq 4\langle T_{\tau_*} \rangle. \tag{28}$$

This completes the proof.

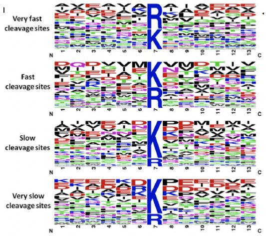
Example: trypsin-catalyzed protein digestion

Trypsin catalyzes the hydrolysis of peptide bonds immediately after lysine/arginine (K/R) residues in proteins.



(from https://at.promega.com)

However, the cleavage rate is affected by specific surrounding amino acids. In other words, the enzyme–substrate complex has many conformations so that distribution of catalysis time should be treated as multiexponential.



(from Pan et al. "Quantitative proteomics reveals the kinetics of trypsin-catalyzed protein digestion."A. B. C. (2014))