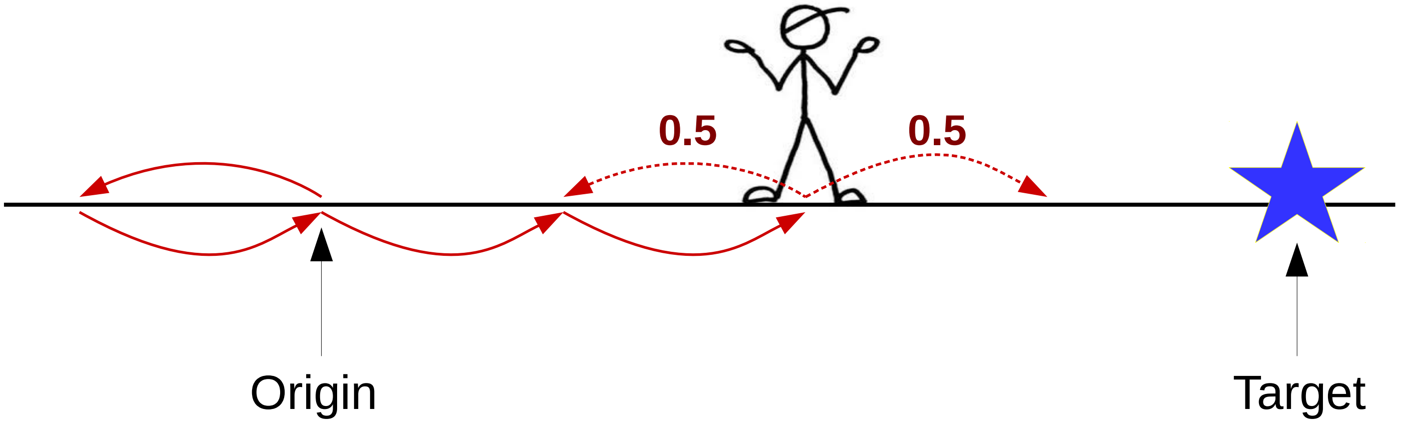
Efficiency of Stochastic Searches with Costly and Punctual Restarts

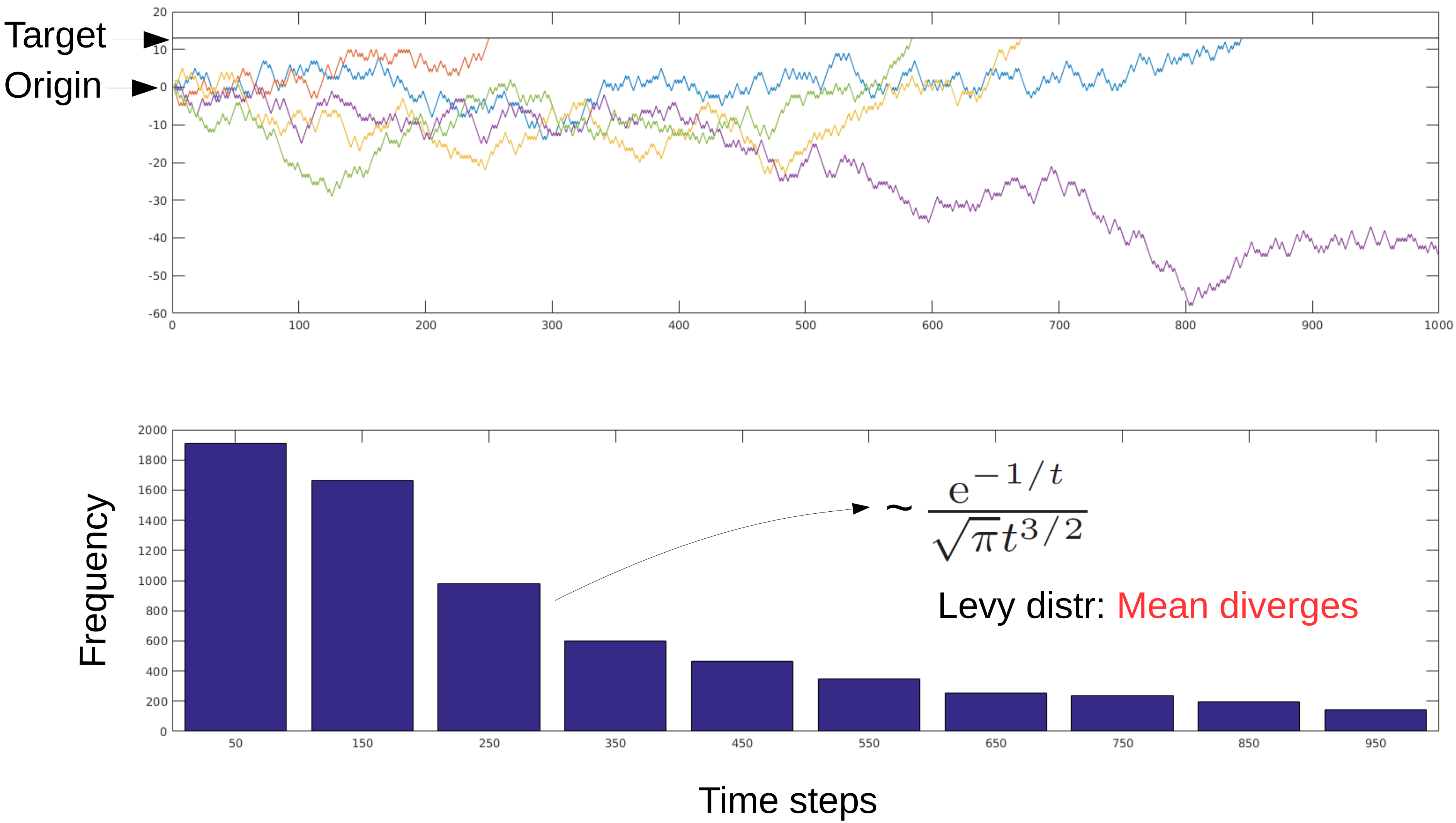
Kabir Husain, Sandeep Krishna
National Centre for Biological Sciences, Bangalore

More details:
Kabir's poster
arXiv:1609.03754

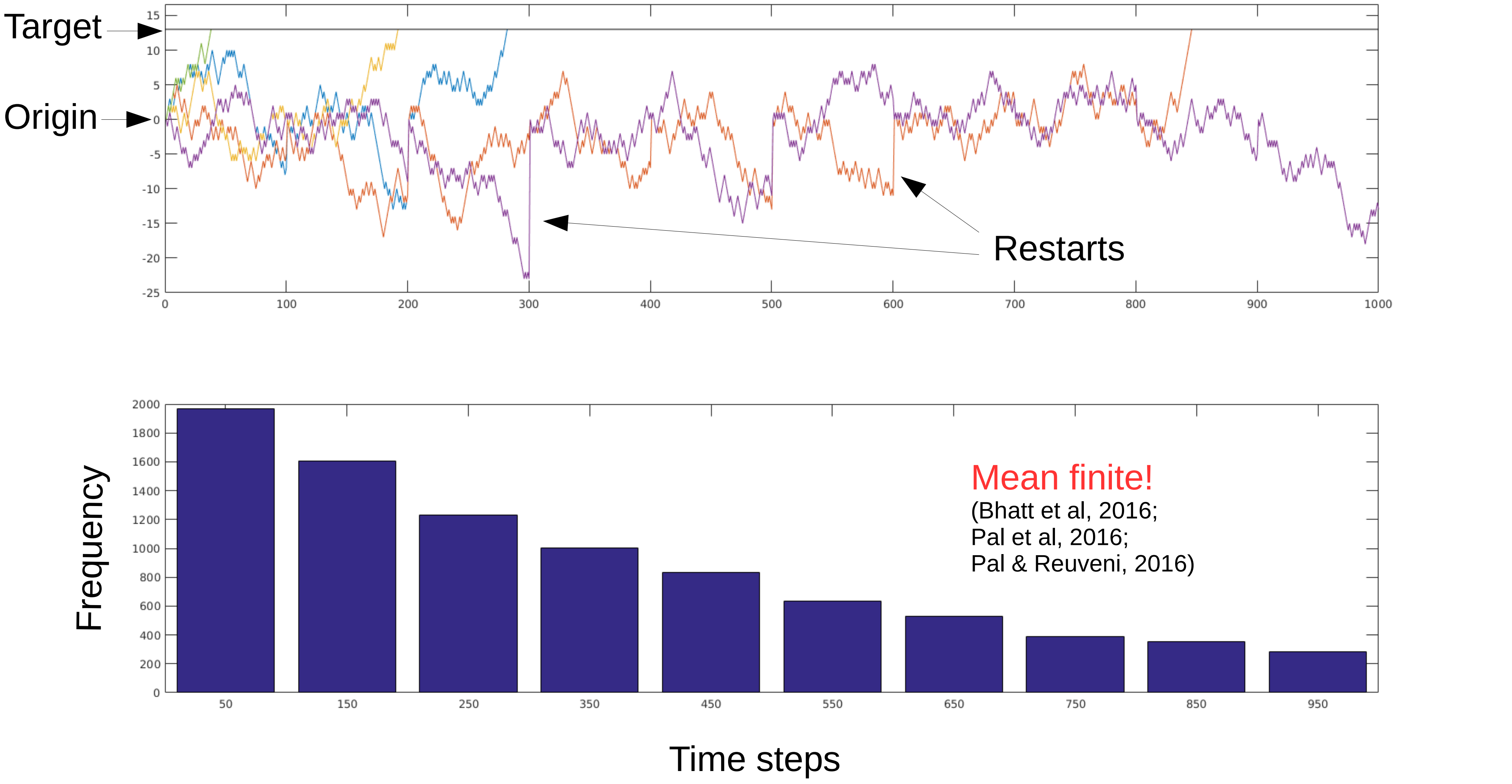
1D Diffusive Search for a Fixed Target



1D Diffusive Search for a Fixed Target



1D Diffusive Search for a Fixed Target with Restarts



Stochastic Searches with Restarts in ...

Computer science:

Satisfiability (SAT) problems
GSAT Algorithm [Selman, Mitchell, Levesque, 1992]
WalkSAT algorithm family [McAllester et al., 1997]
Hoos 1999

Travelling Salesman problem
Johnson & McGeoch 1997
Stutzle & Hoos 1997–2000

Graph partitioning
Martin, Houdayer, Schreiber, 1999

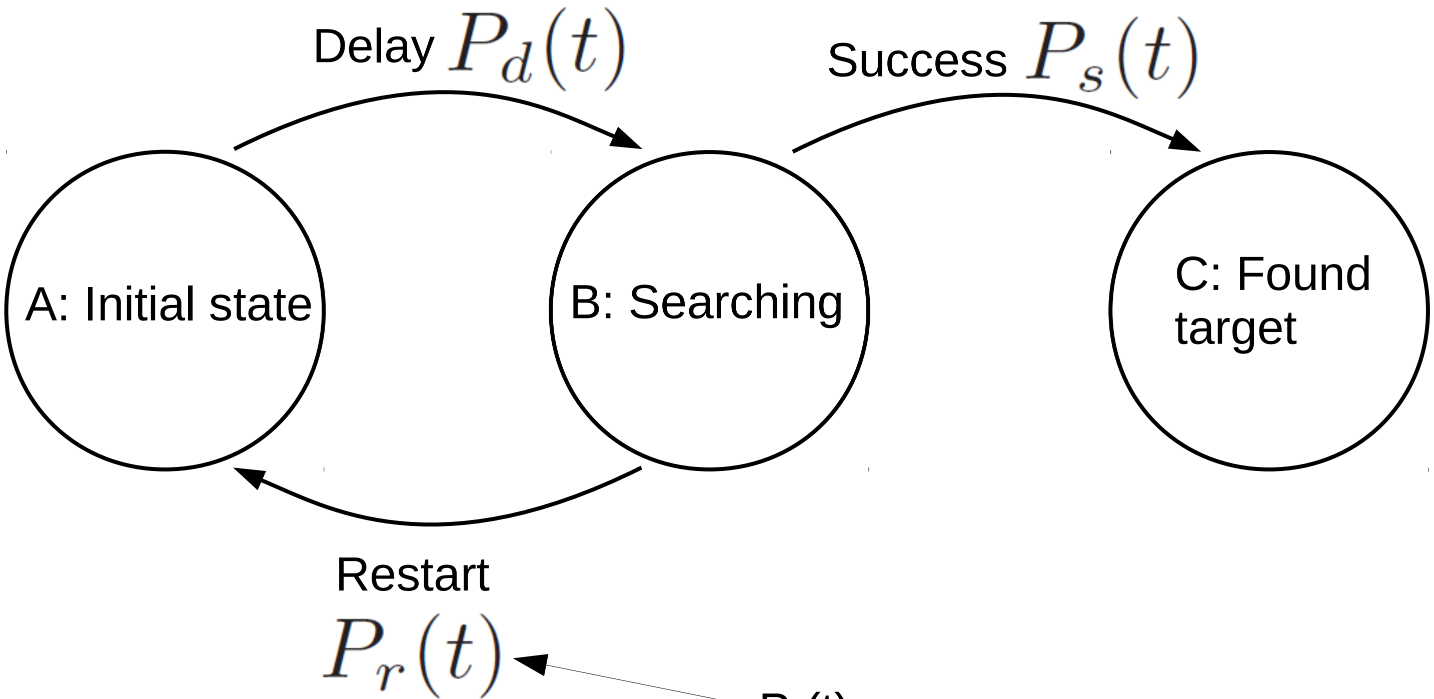
Quadratic Assignment Problem
Stutzle 1997-1999; Taillard 1995

Combinatorial Auctions
Hoos & Boutilier 2000

Biology (the role of restarts is speculation):

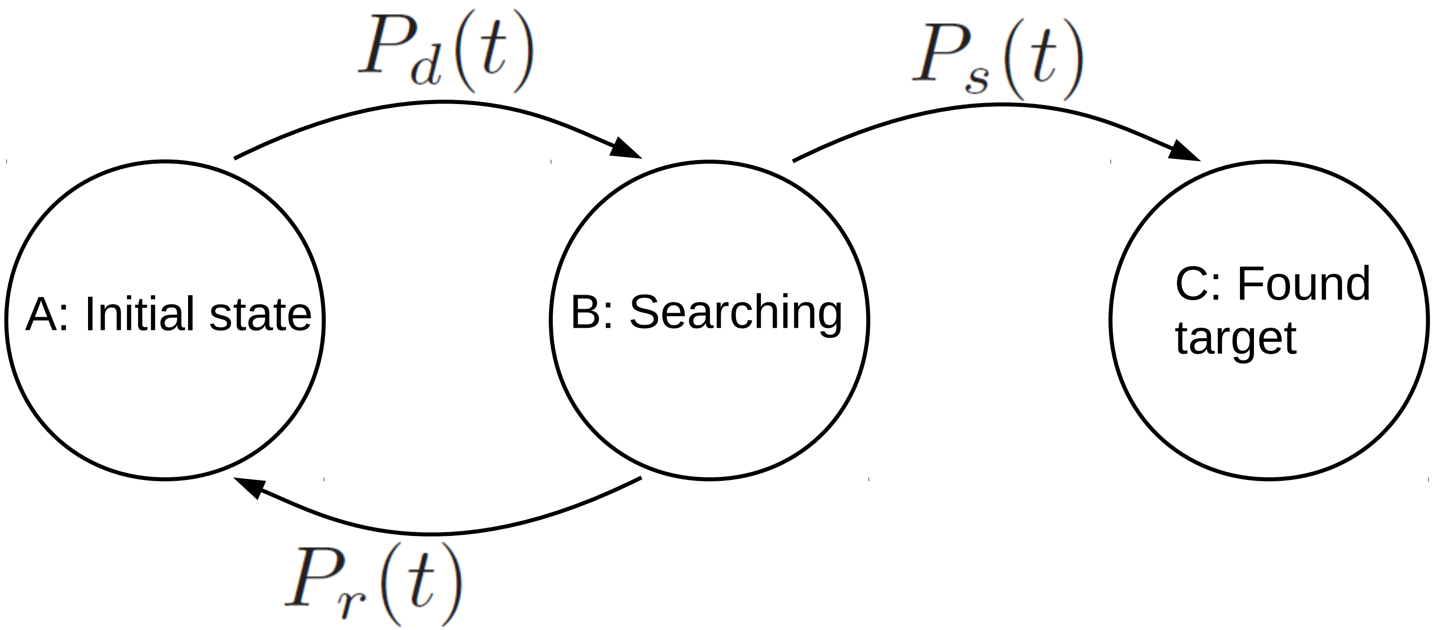
Animal foraging
Immune system recognition of self vs. non-self
Protein folding (do some chaperones implement restarts?)
Binding of regulatory proteins to DNA

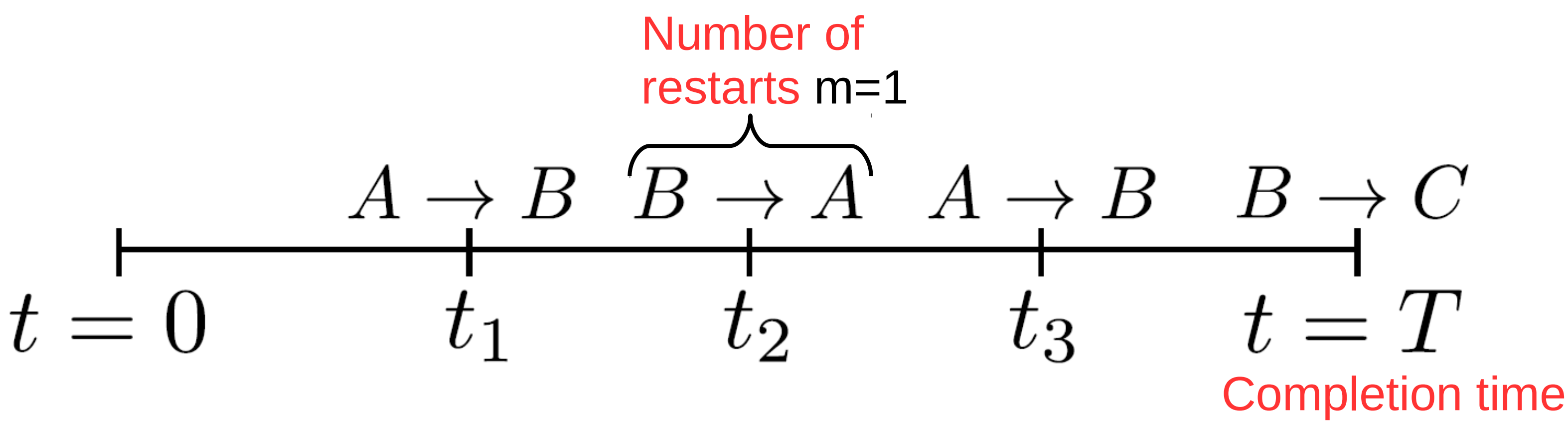
General Framework



$P_x(t)$ are "waiting time" distributions for each state x

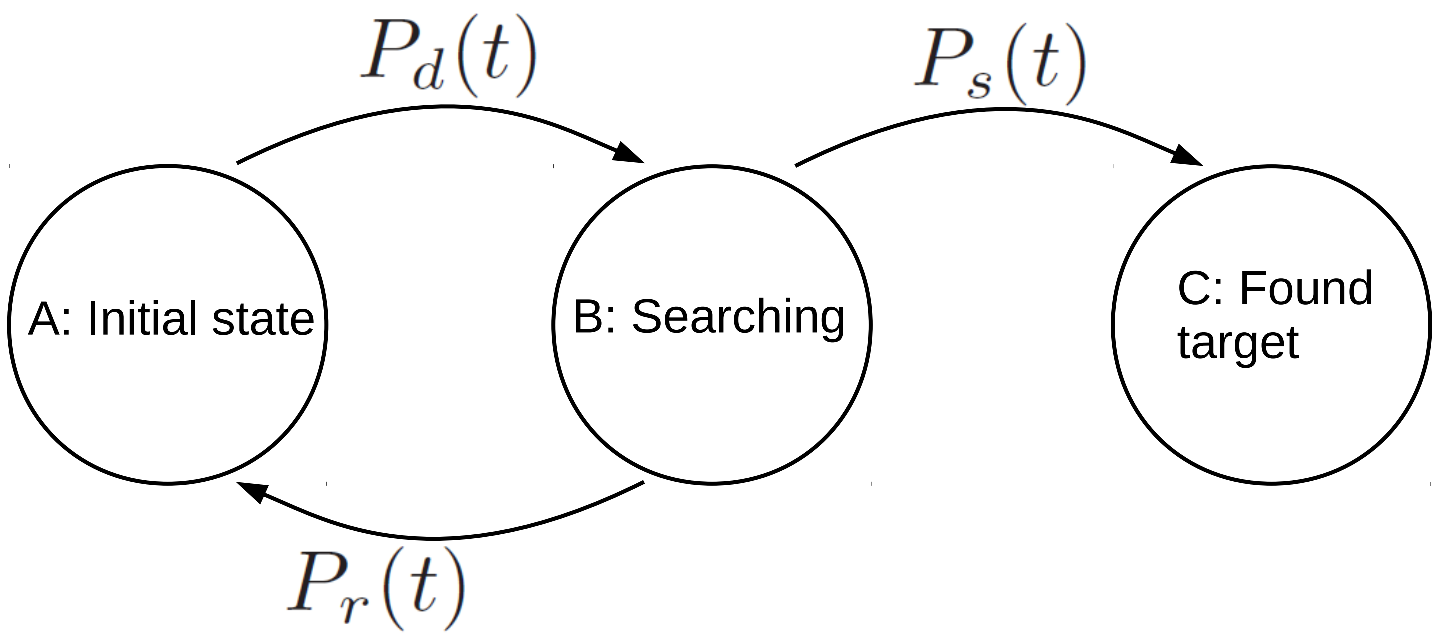
General Framework





Joint probability distribution: $P(m, T)$

General Framework



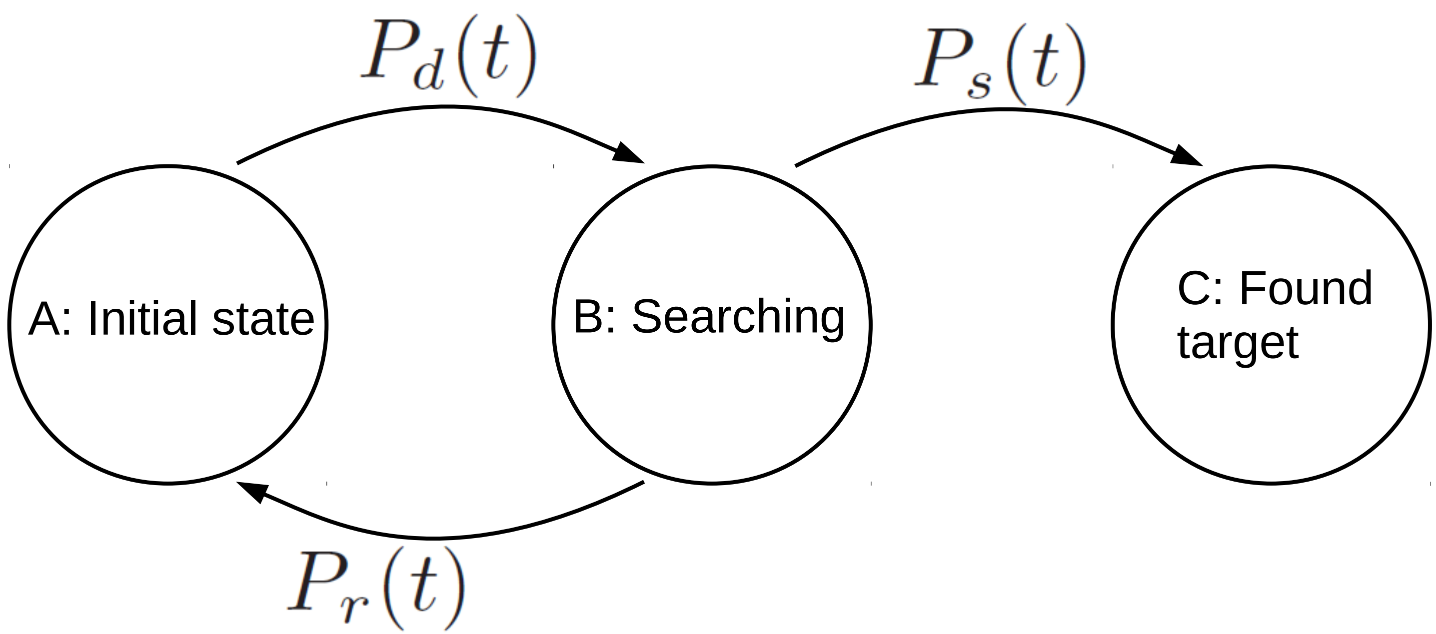
$$\mathcal{L}\left(P(m,T)\right) \equiv \int_0^\infty dT\, P(m,T)\, \mathrm{e}^{-sT} = \hat{G}_f(s)\left(\hat{G}_i(s)\right)^m$$

where $\hat{G}_i(s) = \mathcal{L}\left[P_d(t)\right] \times \mathcal{L}\left[P_r(t) S_s(t)\right]$,

$\hat{G}_f(s) = \mathcal{L}\left[P_d(t)\right] \times \mathcal{L}\left[P_s(t) S_r(t)\right]$, $S_x(t) = 1 - \int_0^t dt' \; P_x(t')$

$$\mathcal{L}\left(P(T)\right) = \hat{G}_f \sum_{m=0}^{\infty} \hat{G}_i^m = \frac{\hat{G}_f(s)}{1 - \hat{G}_i(s)}$$

General Framework



The case of zero delay: $P_d(t) = \delta(t)$

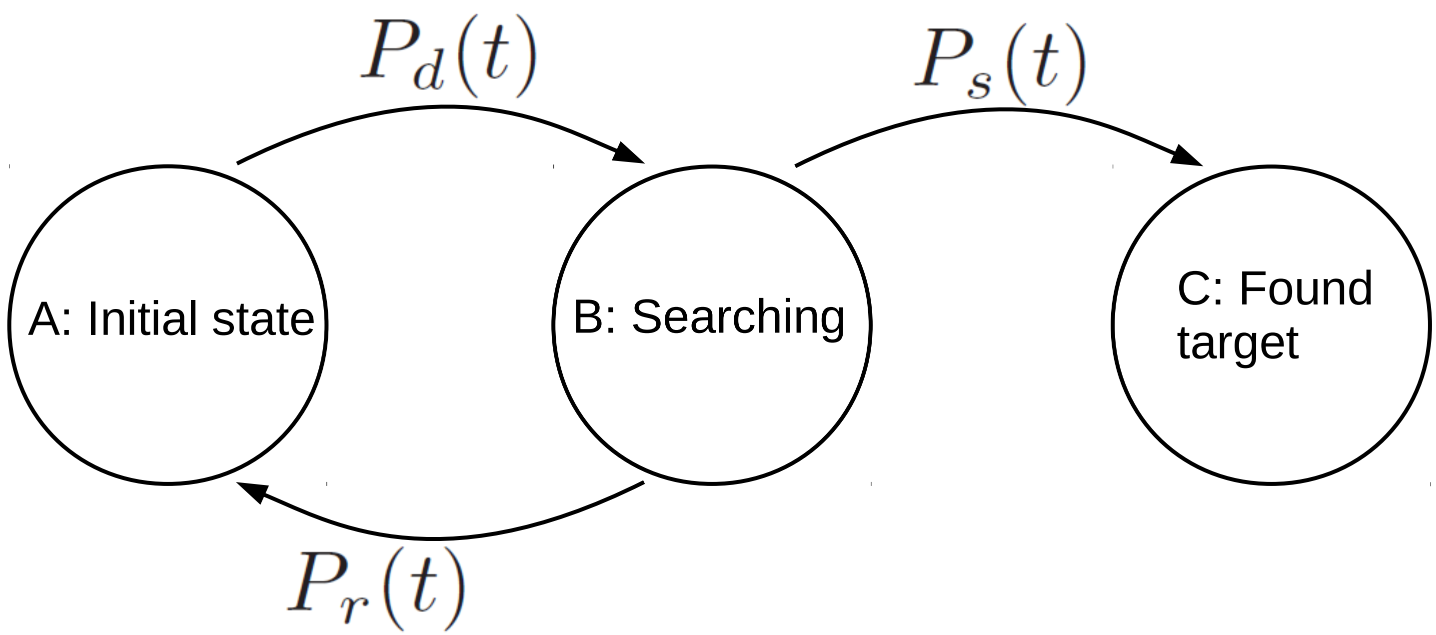
$$\langle T \rangle = \frac{G_f^{(1)} + G_i^{(1)}}{G_f^{(0)}}, \qquad \begin{aligned} G_i(t) &= P_r(t)\ S_s(t) \\ G_f(t) &= P_s(t)\ S_r(t) \end{aligned} \qquad G_x^{(n)} = \int dt\, t^n\, G_x(t)$$

$$\langle T^2 \rangle = \frac{2G_i^{(1)}\left(G_f^{(1)} + G_i^{(1)}\right) + G_f^{(0)}\left(G_f^{(2)} + G_i^{(2)}\right)}{(G_f^{(0)})^2}$$

Computation of all moments reduced to quadratures

$$\mathcal{L}\left(P(T)\right) = \hat{G}_f \sum_{m=0}^{\infty} \hat{G}_i^m = \frac{\hat{G}_f(s)}{1 - \hat{G}_i(s)}$$

General Framework



Re-deriving previously known results:

(1) No delay, deterministic restart: $P_r(t) = \delta(t-\tau); P_d(t) = \delta(t)$

$$\langle T \rangle_\delta = \frac{\int_0^\tau dt\, S_s(t)}{1 - S_s(\tau)}$$

(Bhatt et al, 2016; Pal et al, 2016; Pal & Reuveni, 2016)
Consequence: *any* deterministic restart renders <T> finite, as long as the denominator is non-zero

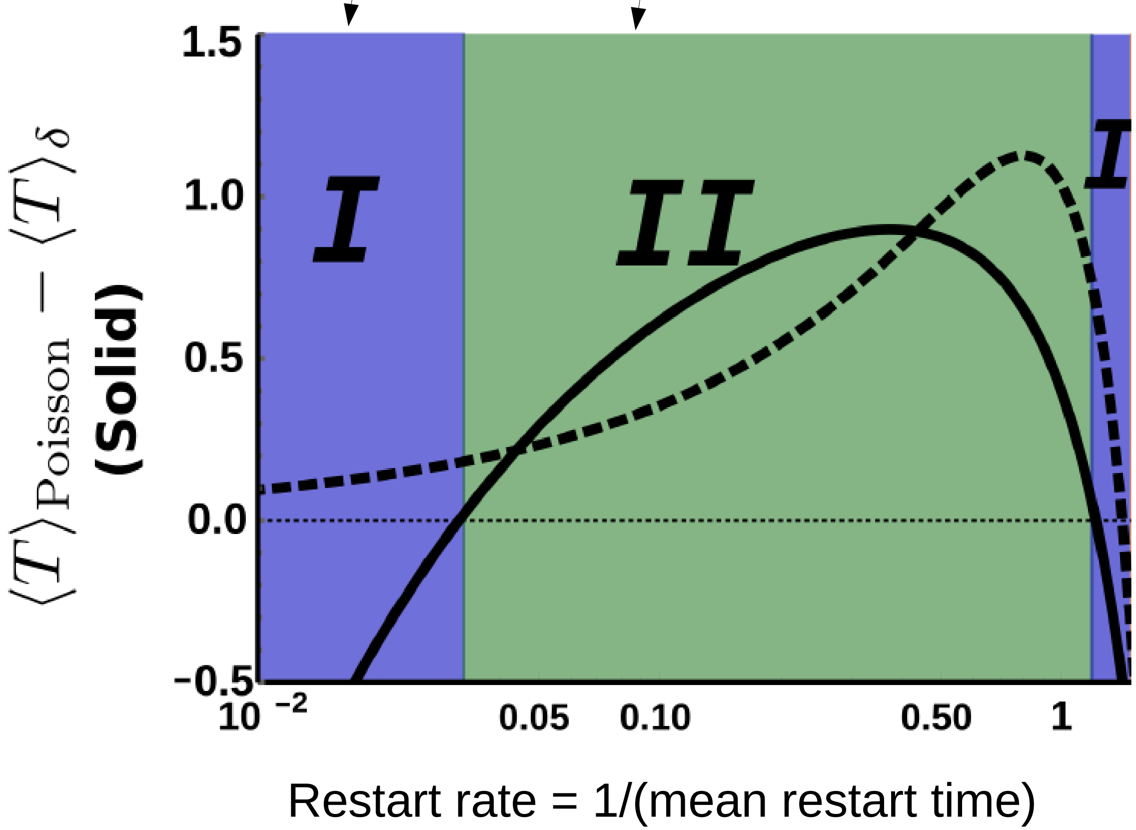
(2) No delay, Poisson restarts, 1D diffusive search: $P_s(t) = \frac{e^{-1/t}}{\sqrt{\pi} t^{3/2}}; P_r(t) = k \exp(-kt)$

$$\langle T \rangle = \left(\exp(2\sqrt{k}) - 1 \right) / k$$ (Evans & Majumdar 2011)

(3) No delay, power law restarts, 1D diffusive search (Nagar & Gupta 2016): e.g. $P_r(t) = \frac{\eta}{(\eta + t)^2}$
<T> can be finite even if search process *and* restarts have diverging mean waiting times

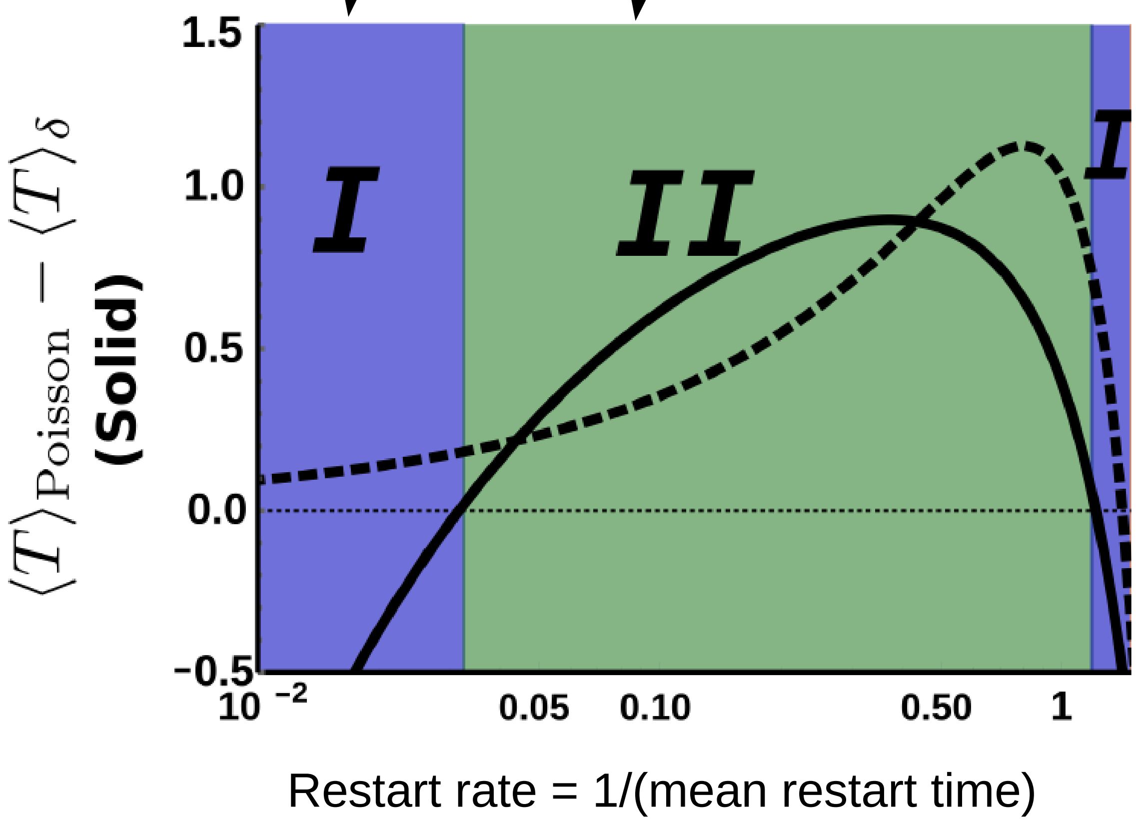
New Results: When does stochasticity in restart distribution help in reducing mean completion time?

For a fixed restart time, adding stochasticity:
- decreases the completion time when restart rate is too small or too large
- increases it when restart rate is in between



New Results: When does stochasticity in restart distribution help in reducing mean completion time?

For a fixed restart time, adding stochasticity:
- decreases the completion time when restart rate is too small or too large
- increases it when restart rate is in between



However, the *optimal* restart distribution is the deterministic one (Evans & Majumdar 2011; Pal & Reuveni 2016)

What if there are un-avoidable sources of noise in the restarts?

New Results: What if restarts are ‘punctual’, i.e., almost-deterministic?

$$P_r(t) = \delta\left(t-\tau\right) + \sum_{n=2}^{\infty} \frac{(-1)^n}{n!} \mu_n(\tau)\, \delta^{(n)}\left(t-\tau\right)$$

$$\langle T \rangle \;=\; \langle T \rangle_\delta(\tau) \,+\, R_{\sigma^2}(\tau)\sigma_r^2 \,+\, \mathcal{O}(\mu_3(\tau))$$

$$\langle T \rangle_\delta = \frac{\int_0^\tau dt\, S_s(t)}{1 - S_s(\tau)}$$

Deterministic restarts

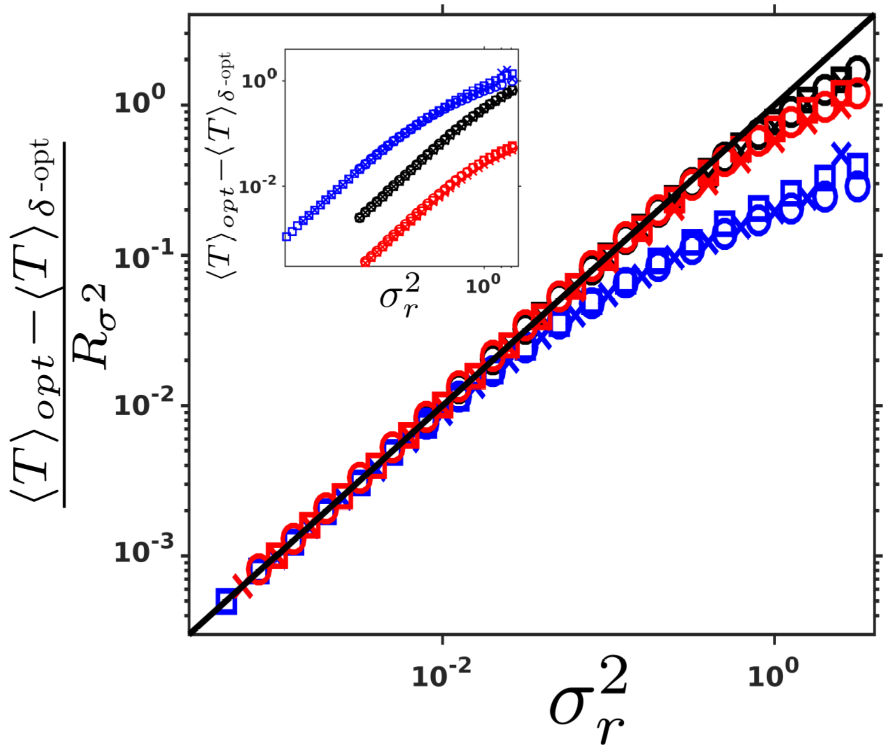
Variance of the
restart distribution

New Results: What if restarts are 'punctual', i.e., almost-deterministic?

$$P_r(t) = \delta\left(t-\tau\right) + \sum_{n=2}^{\infty} \frac{(-1)^n}{n!} \mu_n(\tau)\, \delta^{(n)}\left(t-\tau\right)$$

$$\langle T \rangle \; = \; \langle T \rangle_\delta(\tau) + R_{\sigma^2}(\tau)\sigma_r^2 + \mathcal{O}(\mu_3(\tau))$$

Variance of the restart distribution



Simple scaling of optimal completion time with variance of restart time distribution

Search distr:
Levy, log-normal, Frechet

Restart distr:
Gamma (X), Weibull (O), truncated normal (□)

New Results: What if restarts are 'punctual', i.e., almost-deterministic?

$$P_r(t) = \delta\left(t-\tau\right) + \sum_{n=2}^{\infty} \frac{(-1)^n}{n!} \mu_n(\tau)\, \delta^{(n)}\left(t-\tau\right)$$

$$\langle T \rangle \;=\; \langle T \rangle_\delta(\tau) + R_{\sigma^2}(\tau)\sigma_r^2 + \mathcal{O}(\mu_3(\tau))$$

$$R_{\sigma^2} = -\frac{\langle T \rangle_\delta}{2\left(1 - S_s(\tau)\right)} \left[\frac{P_s(\tau)}{\langle T \rangle_\delta} + \left.\frac{\partial P_s}{\partial t}\right|_{t=\tau}\right]$$

Simple bounds on optimal restart time: $\tau_{\text{opt}} \;>\; t_{\text{mode}}$

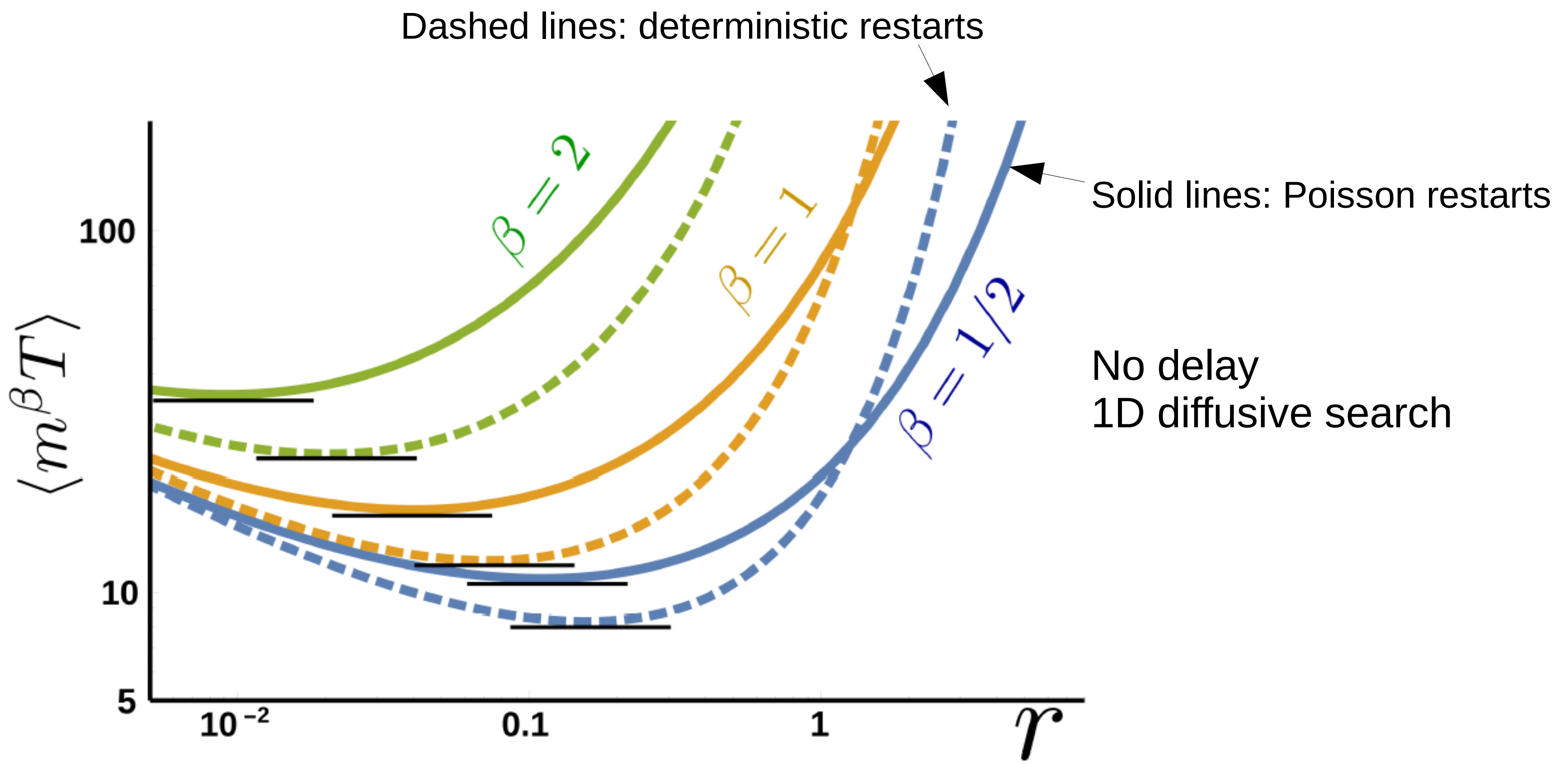
Optimal deterministic restart time — Mode of the search distribution time

New Results: What if restarts are 'costly', i.e., one wants to minimise some combination of completion time and number of restarts?

Moments of m and cross-moments are also easy to compute

$$\langle m^\beta T\rangle = \frac{1}{\langle m\rangle} G_f^{(0)} G_i^{(1)} \Phi\left(G_i^{(0)}, -\beta-1, 0\right) + G_f^{(1)} \Phi\left(G_i^{(0)}, -\beta, 0\right)$$

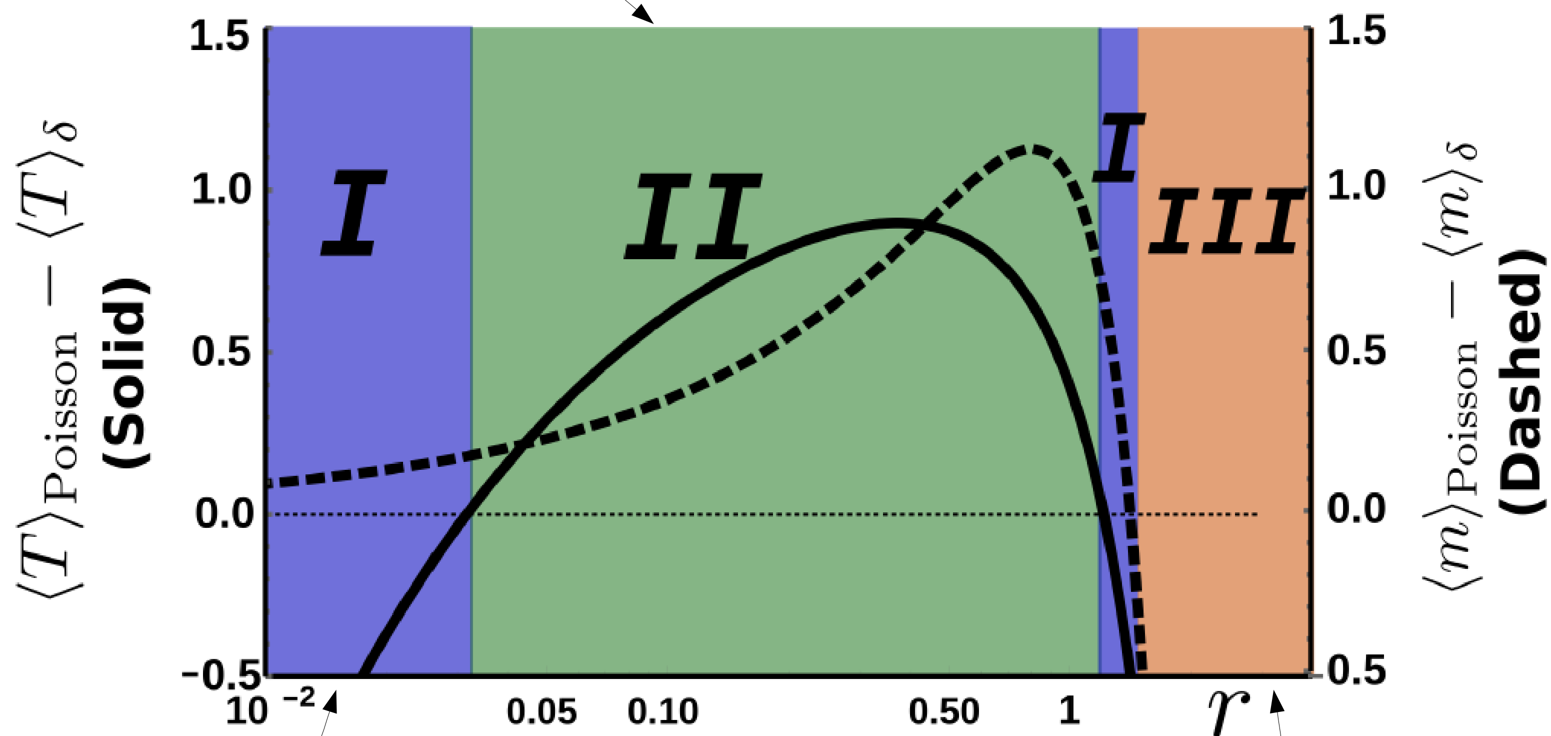
Lerch transcendent



No delay
1D diffusive search

New Results: What if restarts are 'costly', i.e., one wants to minimise some combination of completion time and number of restarts?

Deterministic restarts more Efficient and Faster



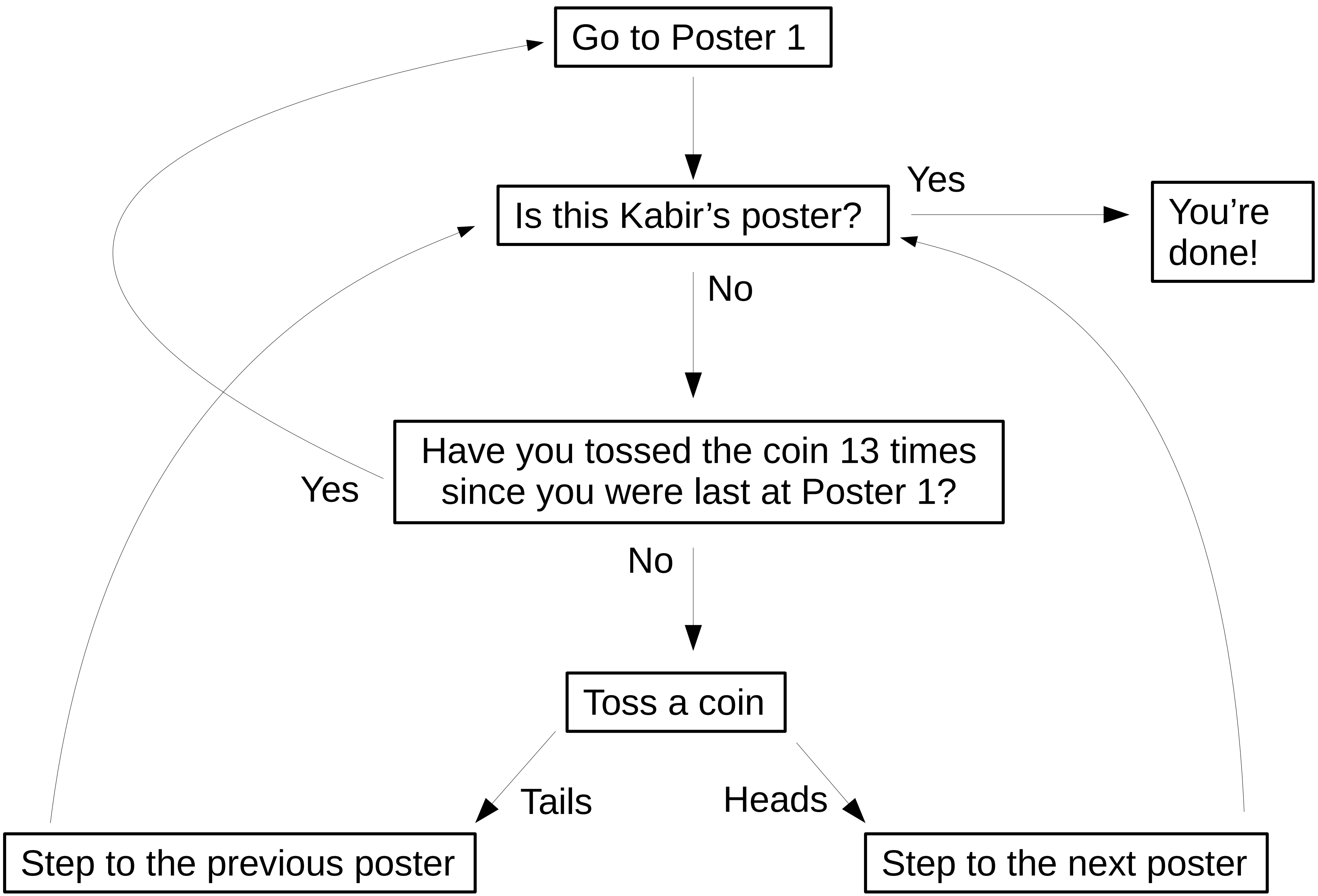
Restart rate = 1/(mean restart time)

Deterministic restarts more Efficient but Stochastic restarts Faster

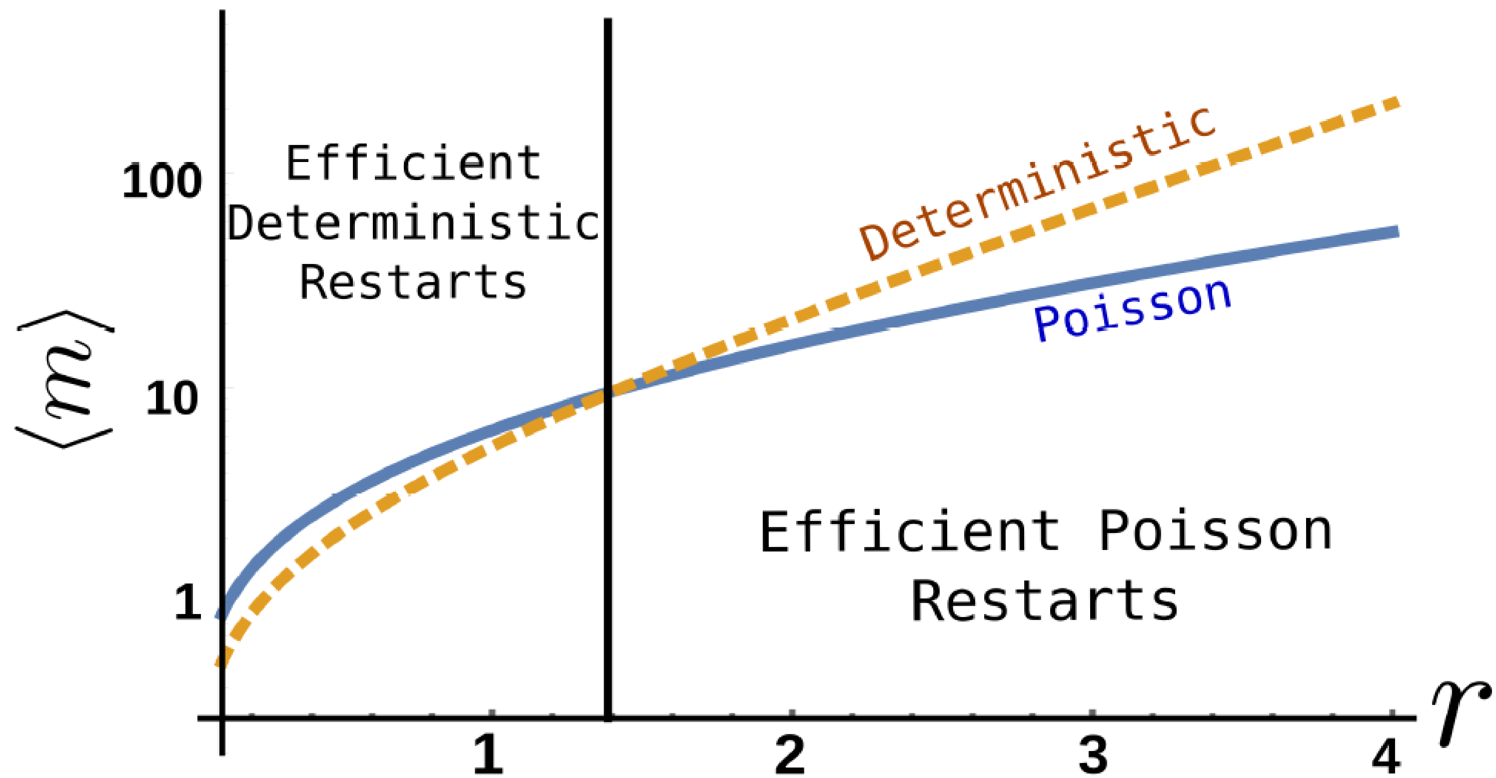
Stochastic restarts more Efficient and Faster

('more Efficient' => has smaller <m>; 'Faster' => has smaller <T>)

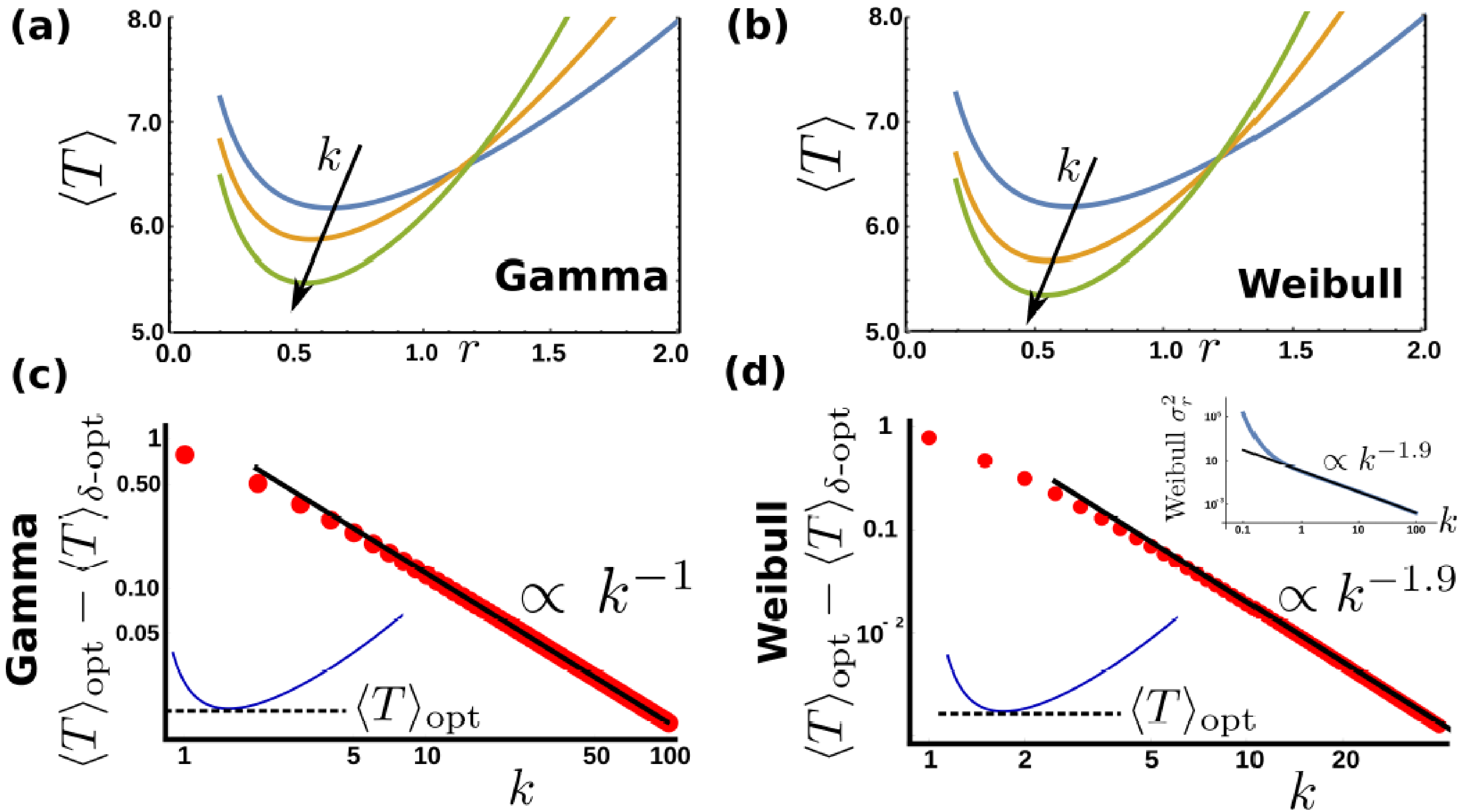
How to find Kabir's poster in a finite time



More details: Kabir Husain, Sandeep Krishna, arXiv:1609.03754

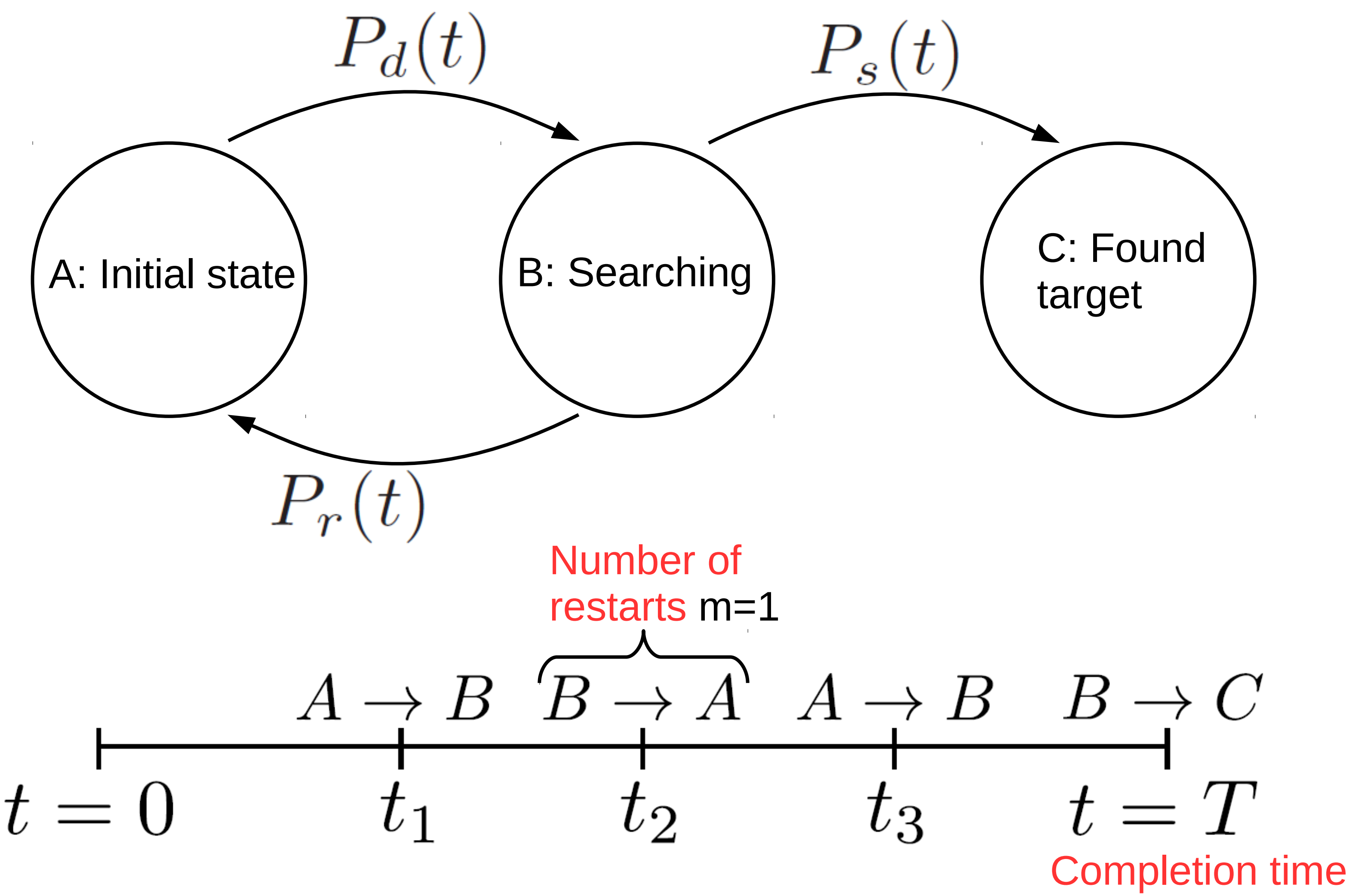


$$\langle m\rangle \approx \langle m\rangle_\delta - \frac{\sigma_r^2}{2\left[1-S_s(\tau)\right]^2}\left.\frac{\partial P_s}{\partial t}\right|_{t=\tau}$$



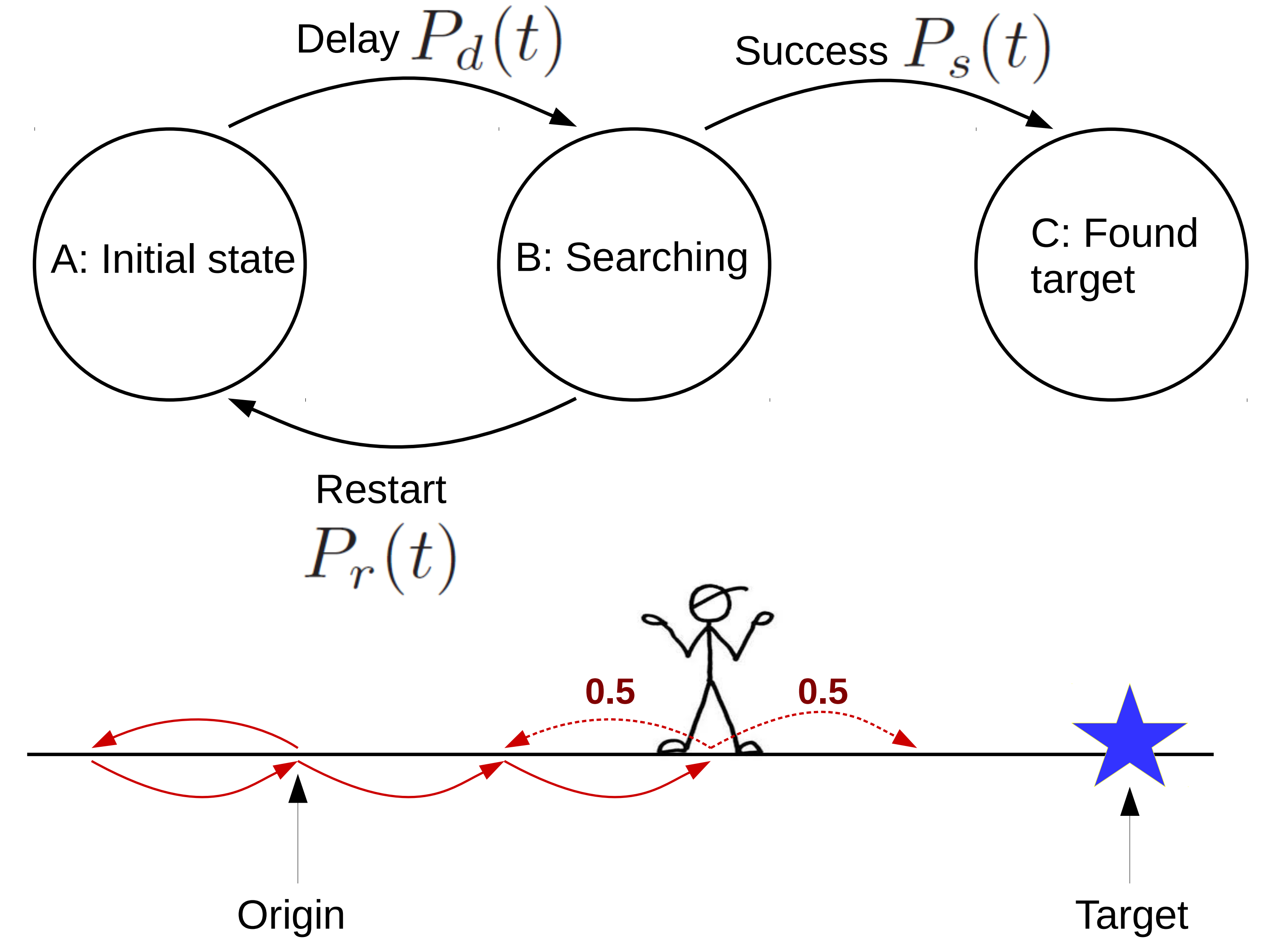
(a)
(b)
(c)
(d)
Gamma
Weibull
k
r
⟨T⟩
8.0
7.0
6.0
5.0
0.0
0.5
1.0
1.5
2.0
⟨T⟩opt − ⟨T⟩δ-opt
∝ k^−1
∝ k^−1.9
⟨T⟩opt
1
0.50
0.10
0.05
5
10
50
100
0.1
10^-2
2
20
Weibull σ_r^2

General Framework



Joint probability distribution: $P(m,T) = \int_0^T dt_{N-1} \ldots \int_0^{t_2} dt_1 \; P(m,T,\{t_i\})$

where $P(m,T,\{t_i\}) = P_d(t_1)\,[P_r(t_2 - t_1)S_s(t_2 - t_1)] \ldots \times [P_s(T - t_{N-1})S_r(T - t_{N-1})]$



$$\mathcal{L}\left(P(m,T)\right) \equiv \int_0^\infty dT\, P(m,T)\, \mathrm{e}^{-st} = \hat{G}_f(s)\left(\hat{G}_i(s)\right)^m$$

Stochastic Searches with Restarts in ...

Computer science:

Satisfiability (SAT) problems
GSAT Algorithm [Selman, Mitchell, Levesque, 1992]
WalkSAT algorithm family [McAllester et al., 1997]
Hoos 1999
Travelling Salesman problem
Johnson & McGeoch 1997
Stutzle & Hoos 1997–2000
Graph partitioning [Martin, Houdayer, Schreiber, 1999]
SAT [e.g., Selman et al. 1992–1997, Hoos et. al. 1994–2000]
TSP [e.g., Johnson & McGeoch 1997, Stutzle & Hoos ¨
1997–2000]
Quadratic Assignment Problem [e.g., Stutzle 1997-1999; ¨
Taillard 1995]
Scheduling [e.g., den Besten, Stutzle, Dorigo 2000] ¨ Planning [e.g., Kautz et al.
1996–1999, Brafman & Hoos
1998–2000]
Combinatorial Auctions [Hoos & Boutilier 2000

Biology (the role of restarts is speculation):

Animal foraging
Immune system recognition of self vs. non-self
Protein folding (do some chaperones implement restarts?)
Binding of regulatory proteins to DNA