Statistics of Basketball Scoring and Lead Changes

in collaboration with Alan Gabel, Marina Kogan, Aaron Clauset

2015 SIAM Conference on Applications of Dynamical Systems

Statistics of Basketball Scoring and Lead Changes

in collaboration with Alan Gabel, Marina Kogan, Aaron Clauset

2015 SIAM Conference on Applications of Dynamical Systems

dunk courtesy of Dr. J

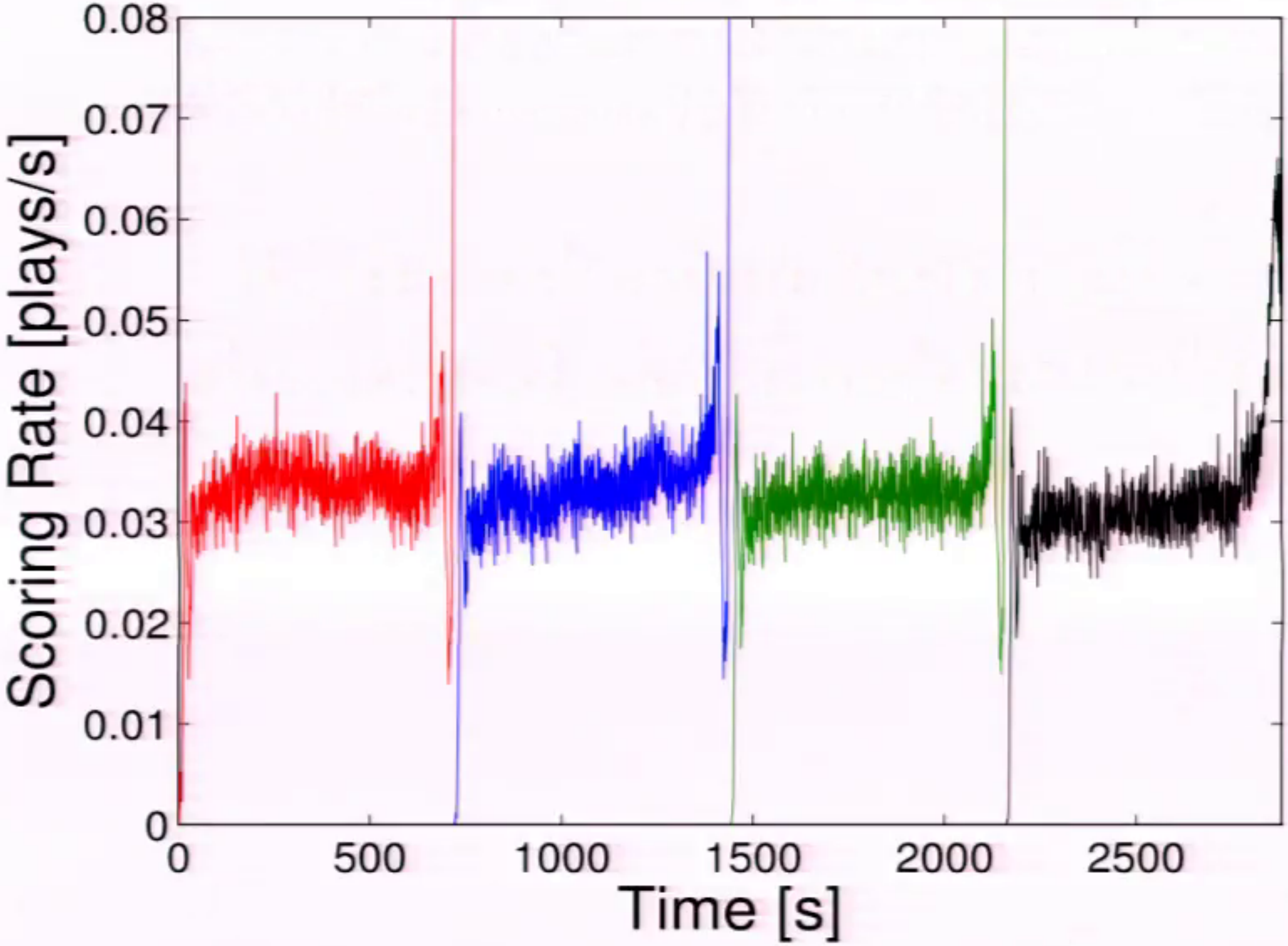
Statistics of Basketball Scoring and Lead Changes

in collaboration with Alan Gabel, Marina Kogan, Aaron Clauset

2015 SIAM Conference on Applications of Dynamical Systems

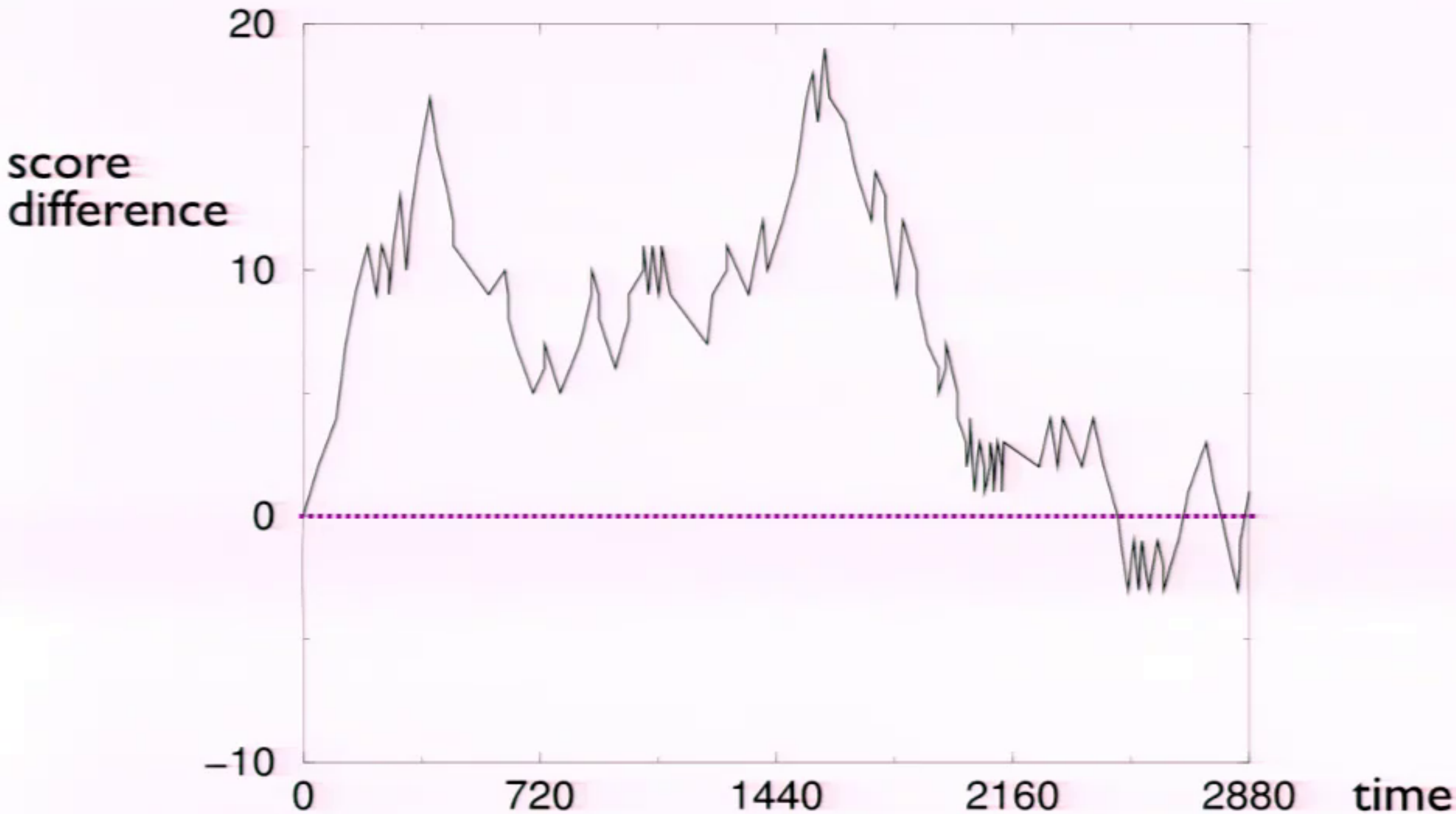
Can basketball scoring be described as a random walk?

(Almost) Constant Scoring Rate

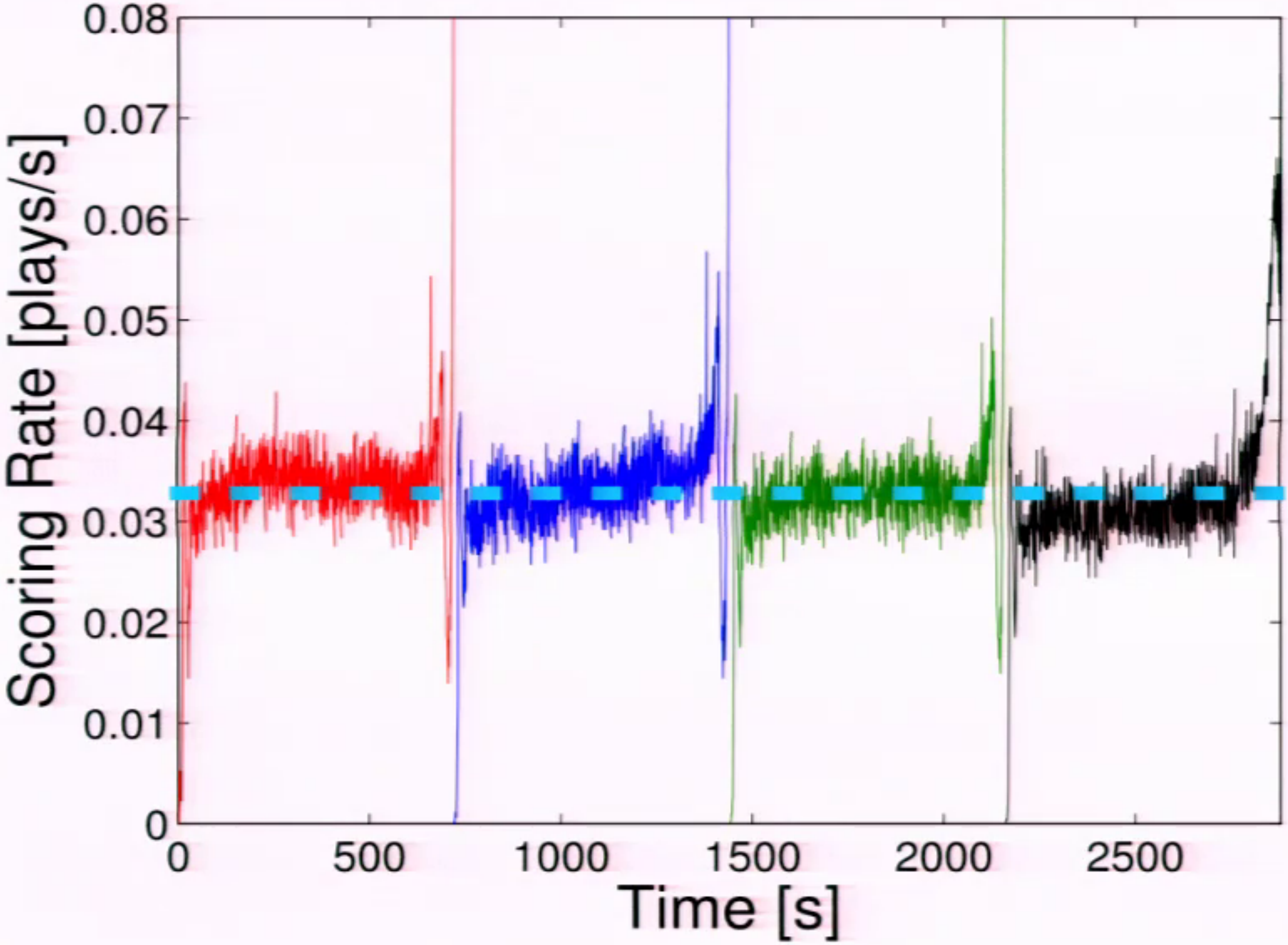


Score Evolution in a Typical NBA Basketball Game

Chicago-Denver 11/26/2010

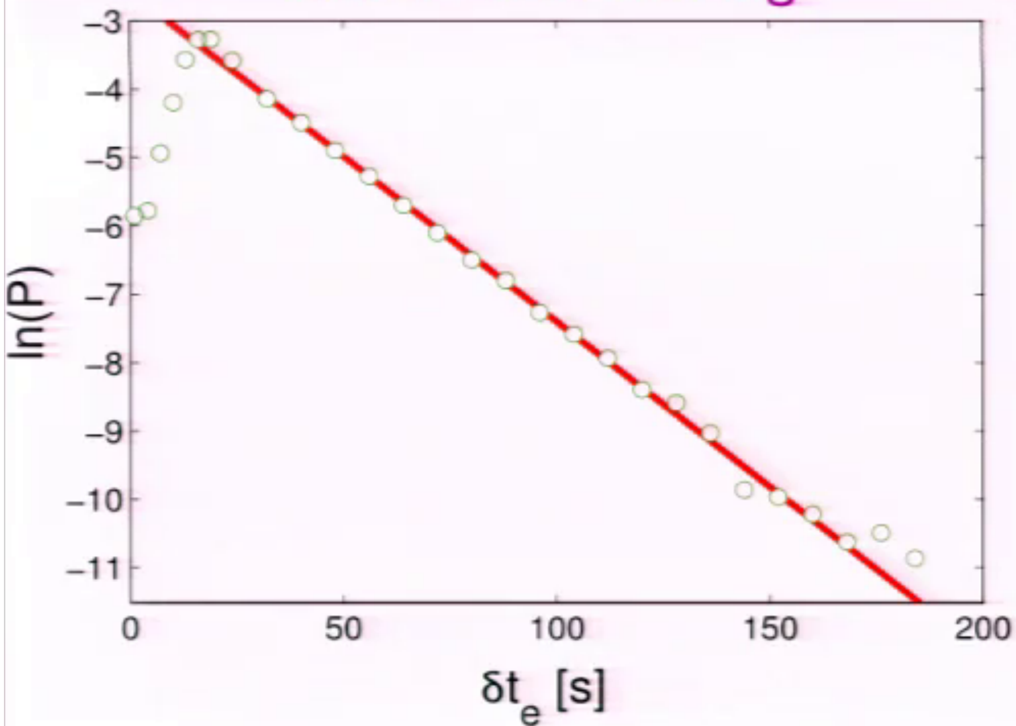


(Almost) Constant Scoring Rate



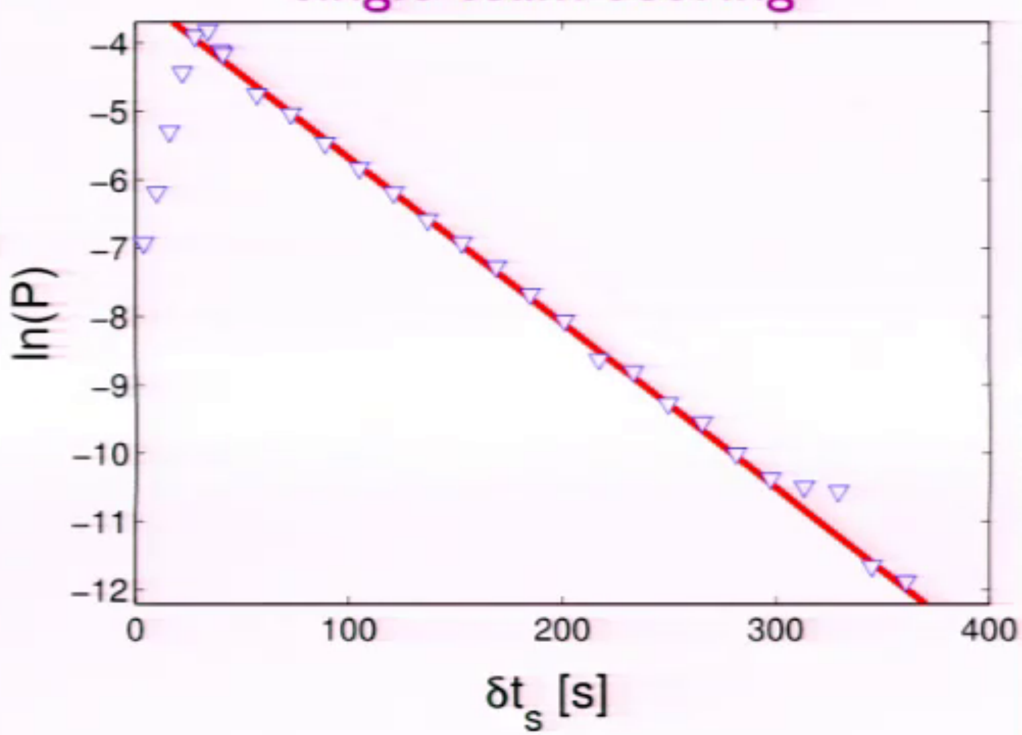
Random Time Intervals Between Scores

either team scoring



$$P(\delta t) \sim e^{-\lambda \delta t}$$

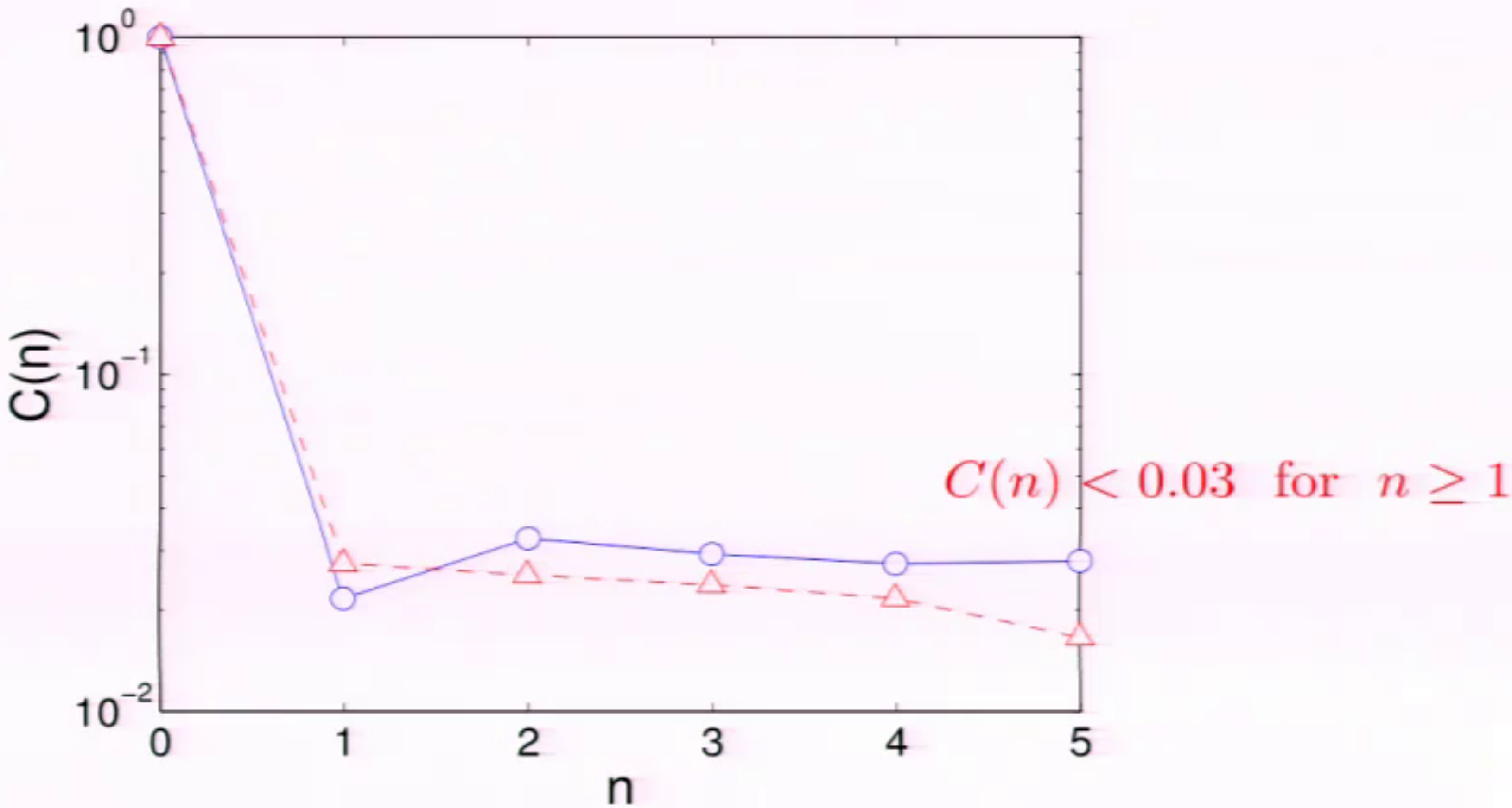
single team scoring



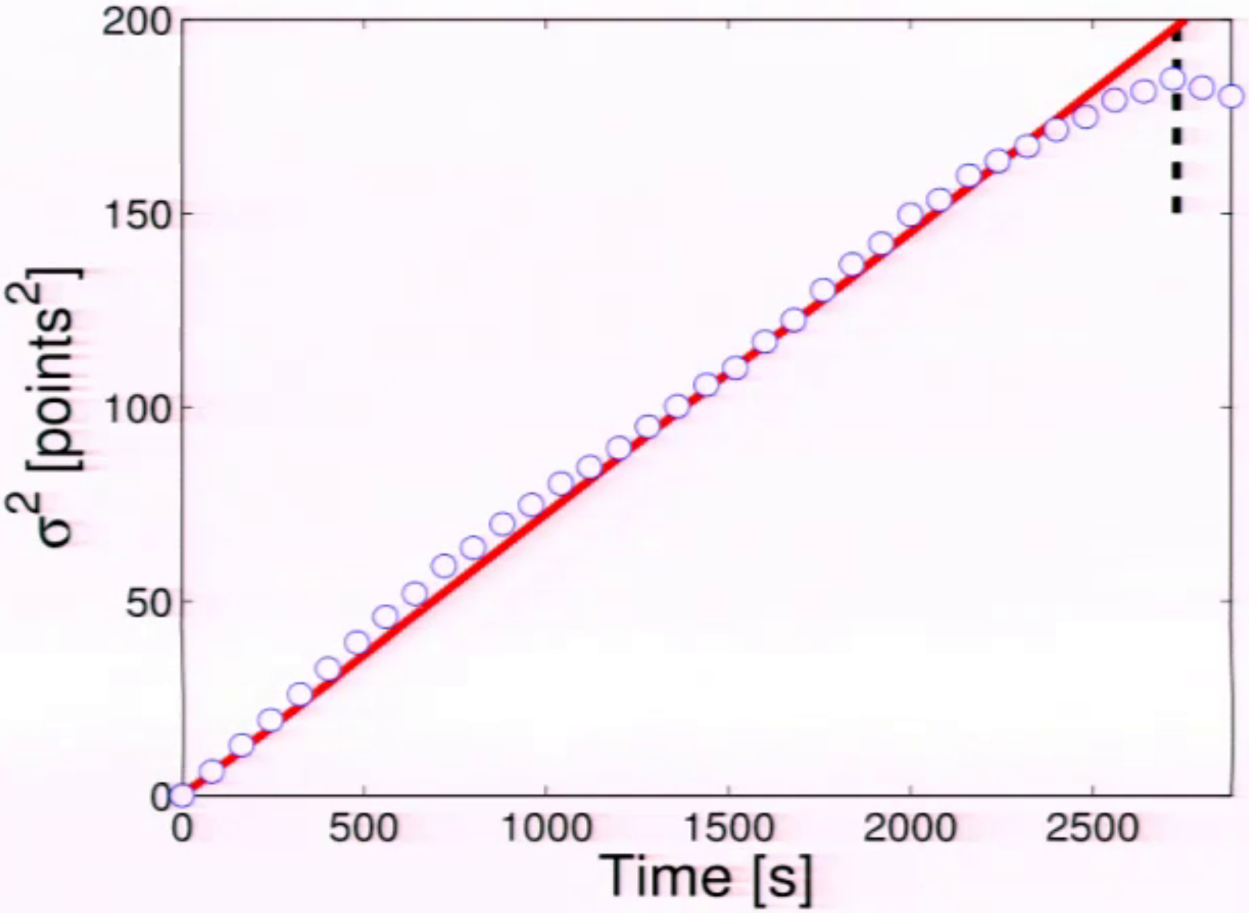
No "Hot Hand" Gilovich, Vallone & Tversky (1985)

Correlation Between Successive Scores

$$C(n) \equiv \frac{\sum_k (t_k - \langle t \rangle)(t_{k+n} - \langle t \rangle)}{\sum_k (t_k - \langle t \rangle)^2}$$



Linear Growth in the Score Difference Variance



Statistics of Lead Changes

- How many lead changes occur?
- How long does one team lead?
- When does the last lead change occur?
- When is a lead "safe"?

Number of Lead Changes

evenly matched teams: $m \sim \sqrt{N}$

$$\int^{N} dN' \, P(0, N') \sim \sqrt{N}$$

evenly matched teams;
antipersistence p: $m \sim \sqrt{N(1-p)/p}$

Antipersistence: After team A scores

team A scores next with probability p=0.35
team B scores next with probability 1-p=0.65

Distribution of Number of Lead Changes

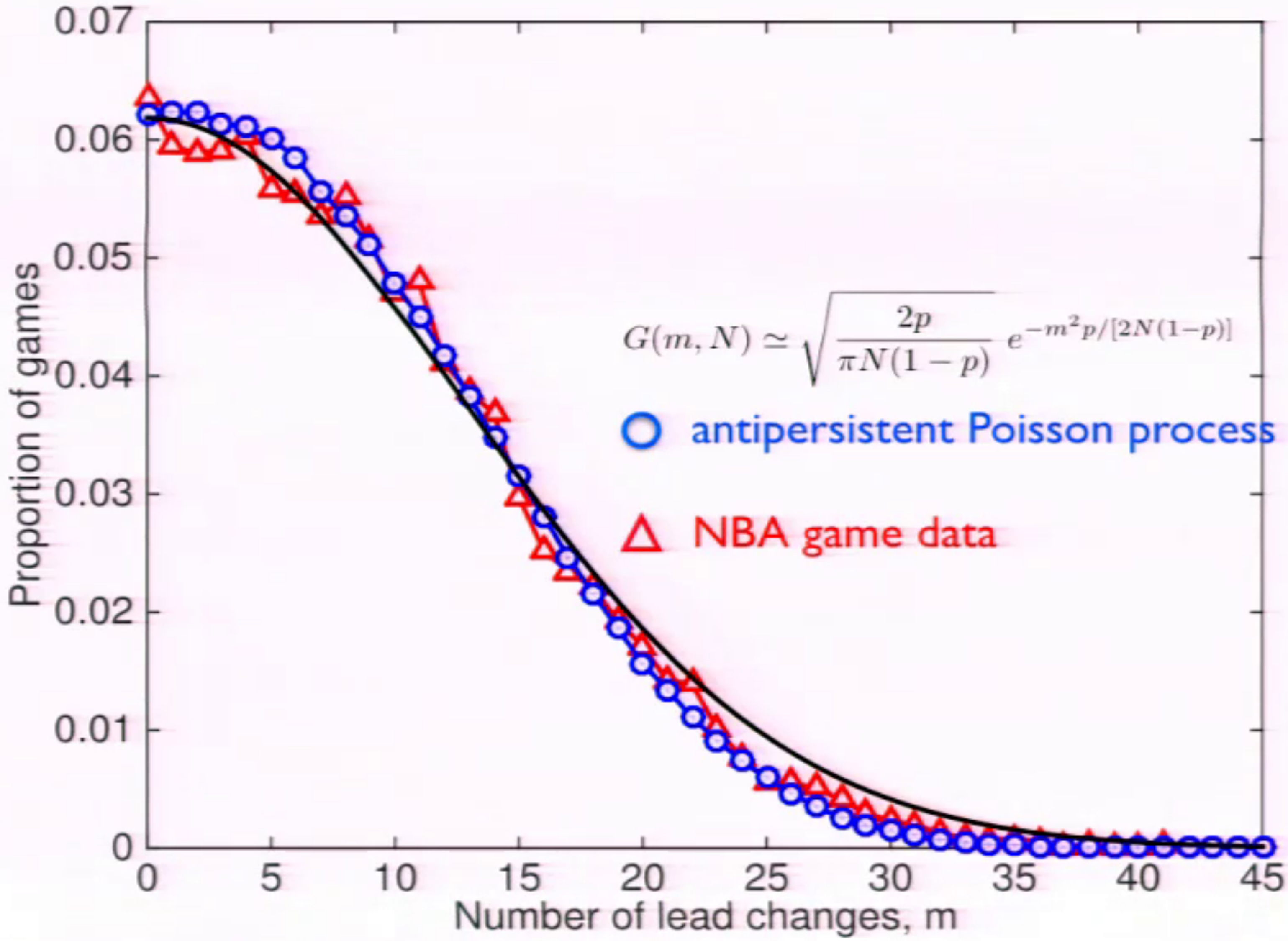
no antipersistence:

$$G(m,N) \simeq \sqrt{\frac{2}{\pi N}}\, e^{-m^2/2N}$$

with antipersistence:

$$G(m,N) \simeq \sqrt{\frac{2p}{\pi N(1-p)}}\, e^{-m^2 p/[2N(1-p)]}$$

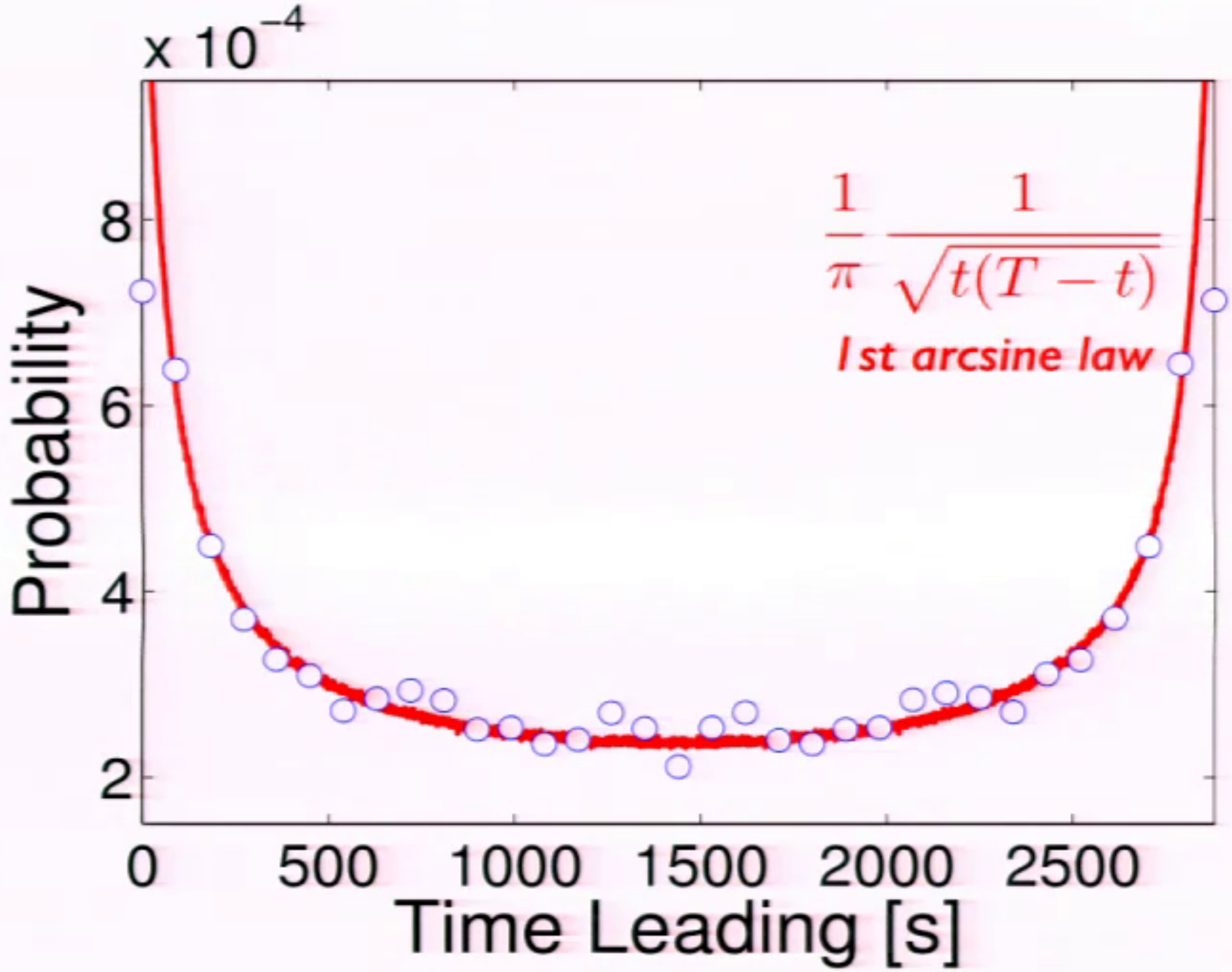
Distribution of Number of Lead Changes



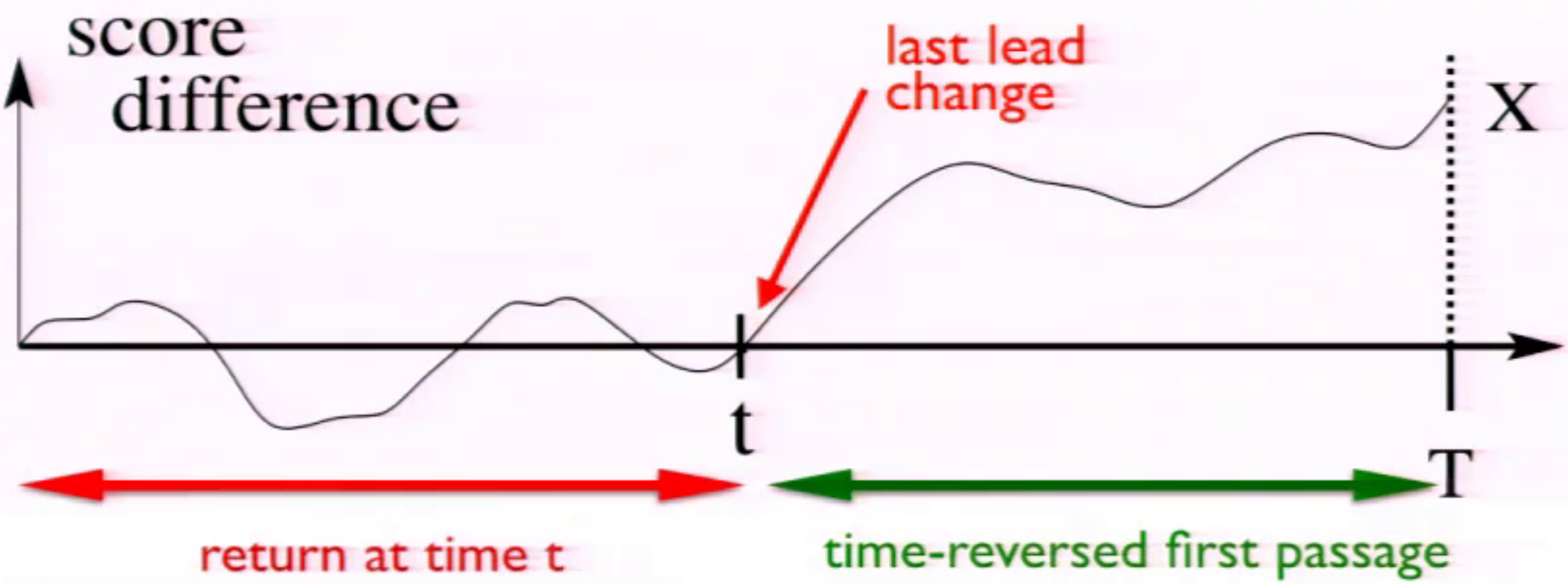
How Long Does One Team Lead?

Chicago-Denver 11/26/2010

How Long Does One Team Lead?



When Does the Last Lead Change Occur?



$$\frac{1}{\sqrt{4\pi D t}}$$

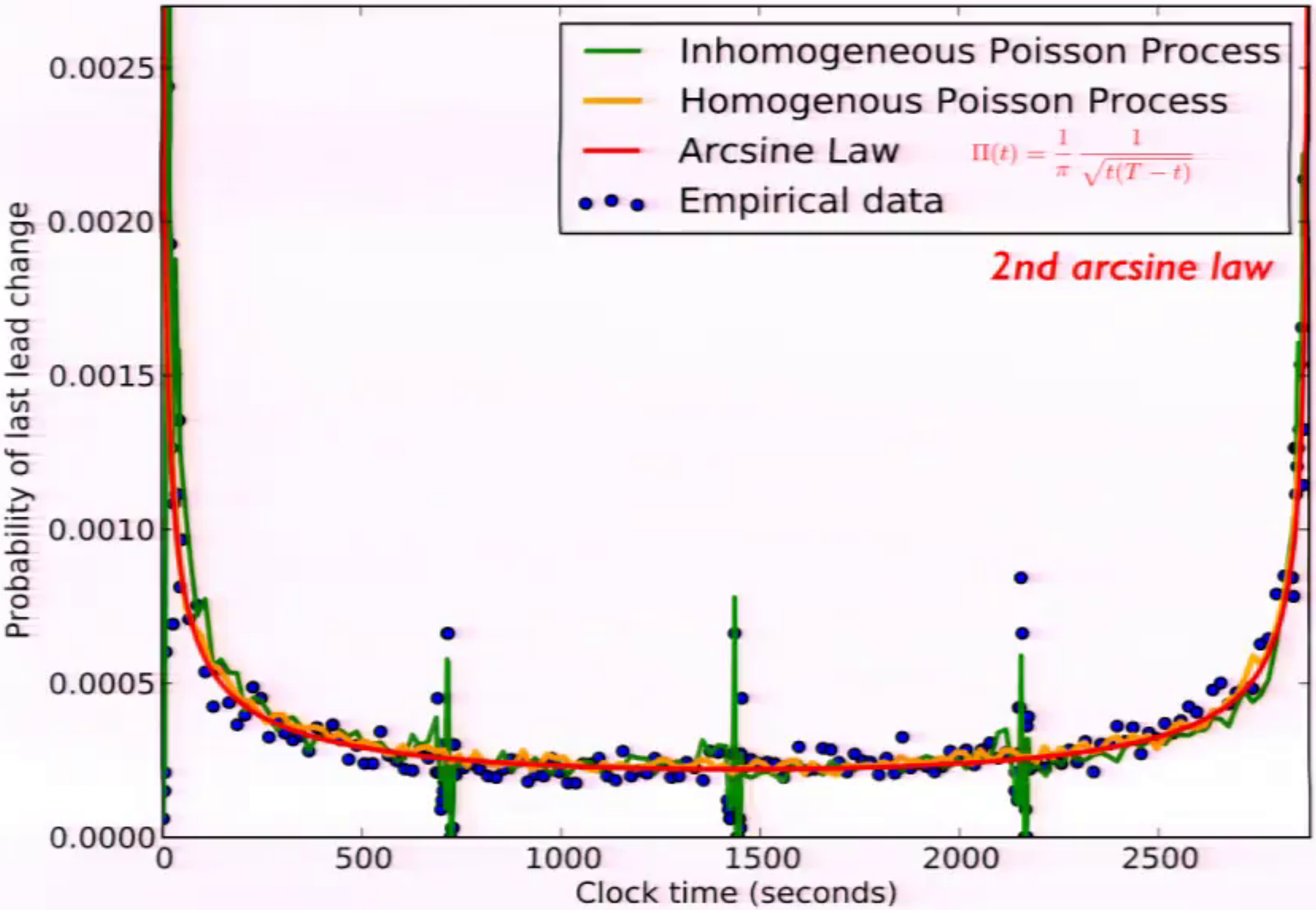
$$\frac{x}{\sqrt{4\pi D (T-t)^3}}\, e^{-x^2/[4D(T-t)]}$$

When Does the Last Lead Change Occur?

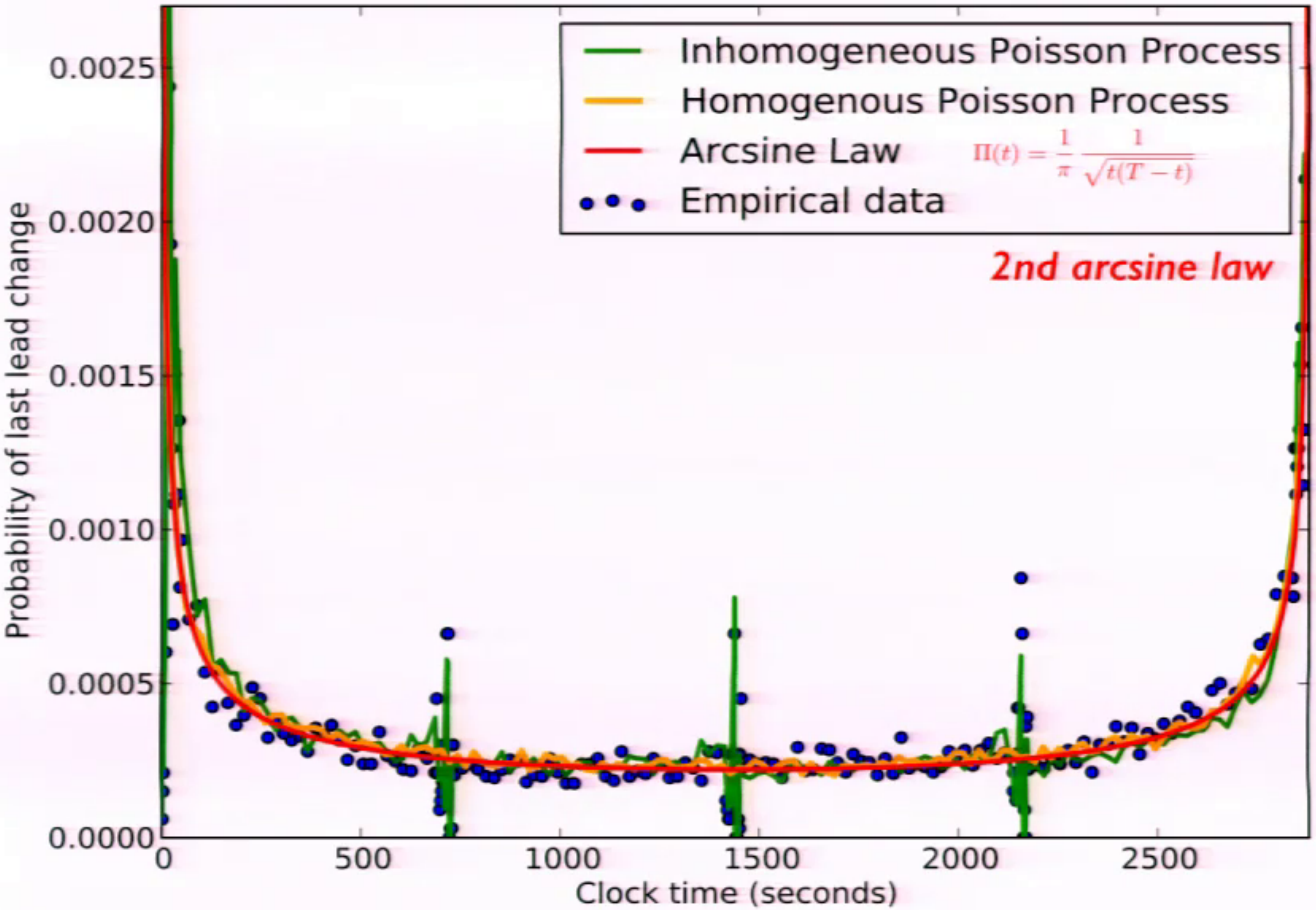
$\Pi(t)$ = Probability that last lead change occurs at time t

$$\begin{aligned}
\Pi(t) &= 2\int_0^\infty dx\, P(x{=}0,t)\, F(x,T{-}t) \\
&= 2\int_0^\infty \frac{dx}{\sqrt{4\pi Dt}}\, \frac{x}{\sqrt{4\pi D(T{-}t)^3}}\, e^{-x^2/[4D(T-t)]} \\
&= \frac{1}{\pi}\, \frac{1}{\sqrt{t(T-t)}}
\end{aligned}$$

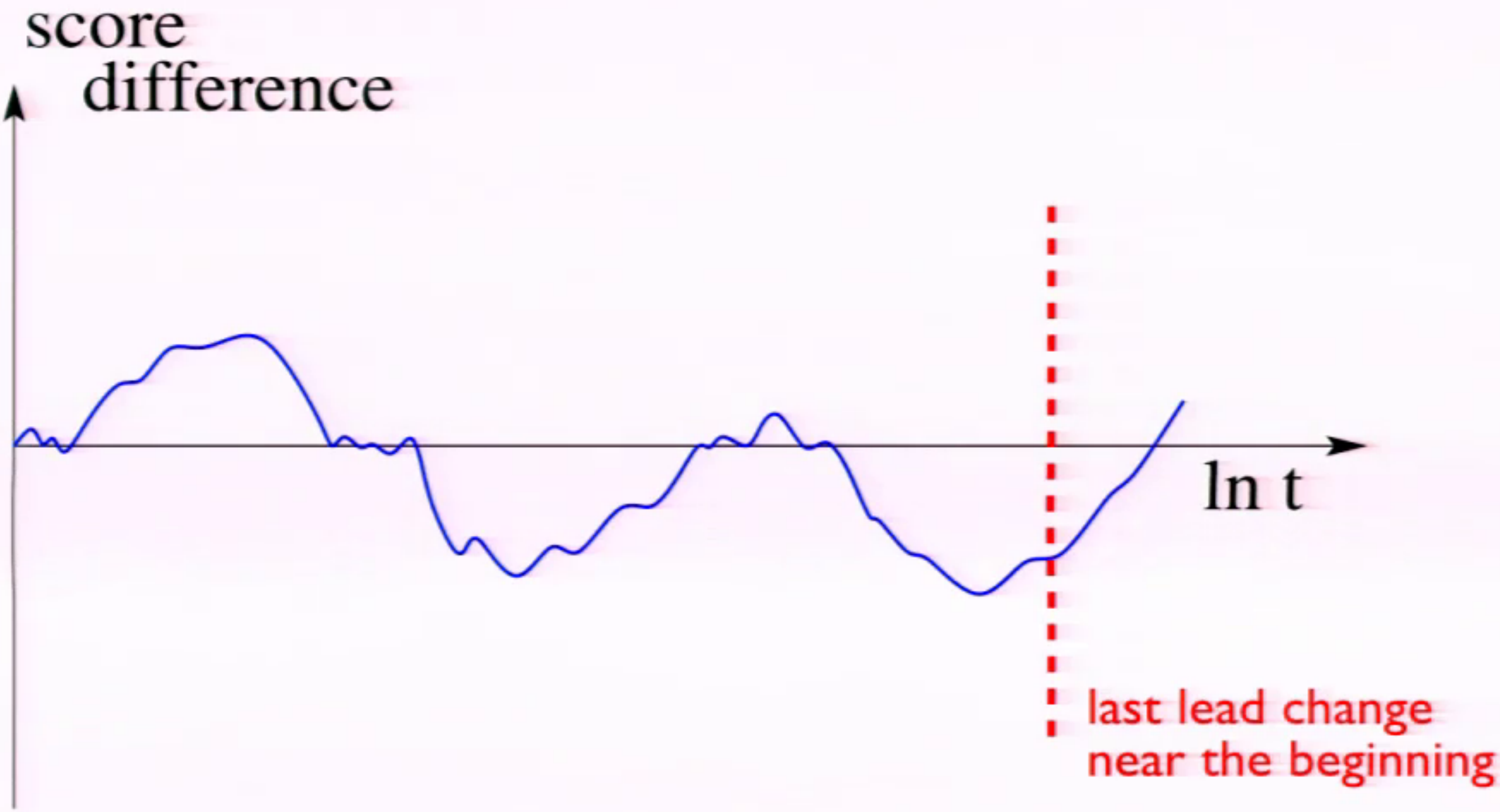
Probability When Last Lead Occurs



Probability When Last Lead Occurs

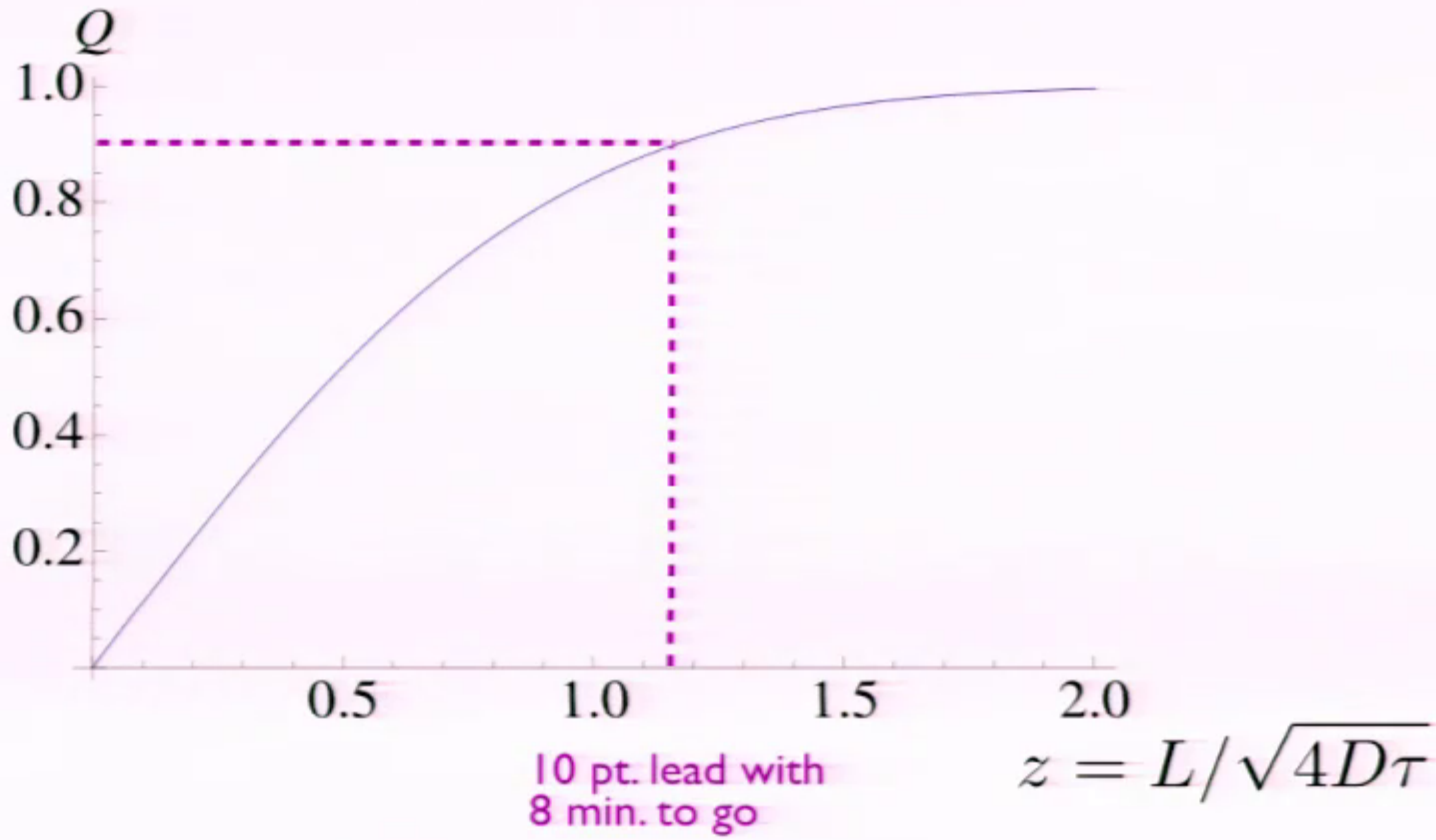


Why?

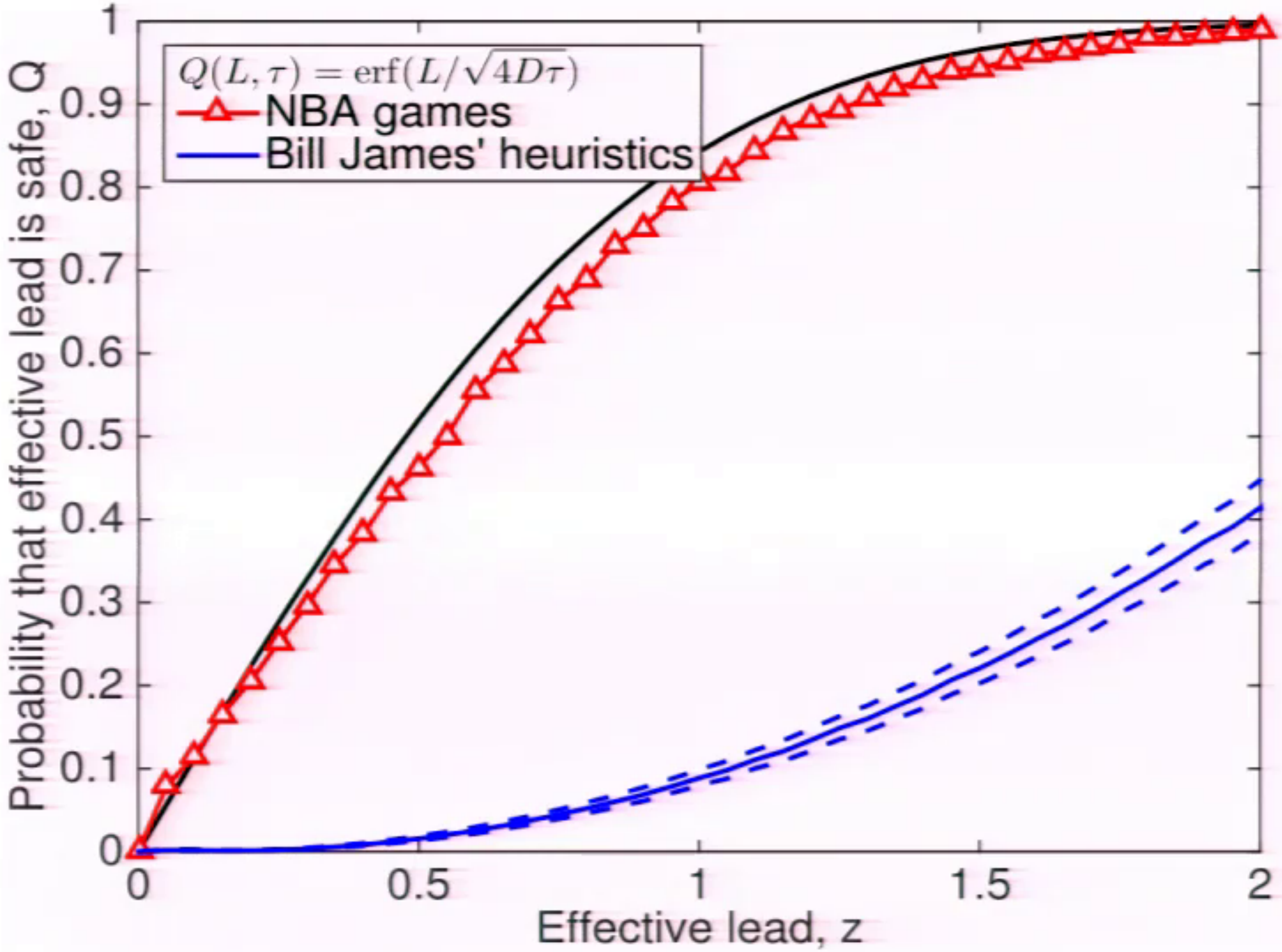


When Is a Lead Safe?

$$Q(L,\tau) = \mathrm{erf}(L/\sqrt{4D\tau})$$



When Is a Lead Safe?



Bill James
Inventor of *Sabermetrics*

Statistics of Basketball Scoring and Lead Changes

in collaboration with Alan Gabel, Marina Kogan, Aaron Clauset

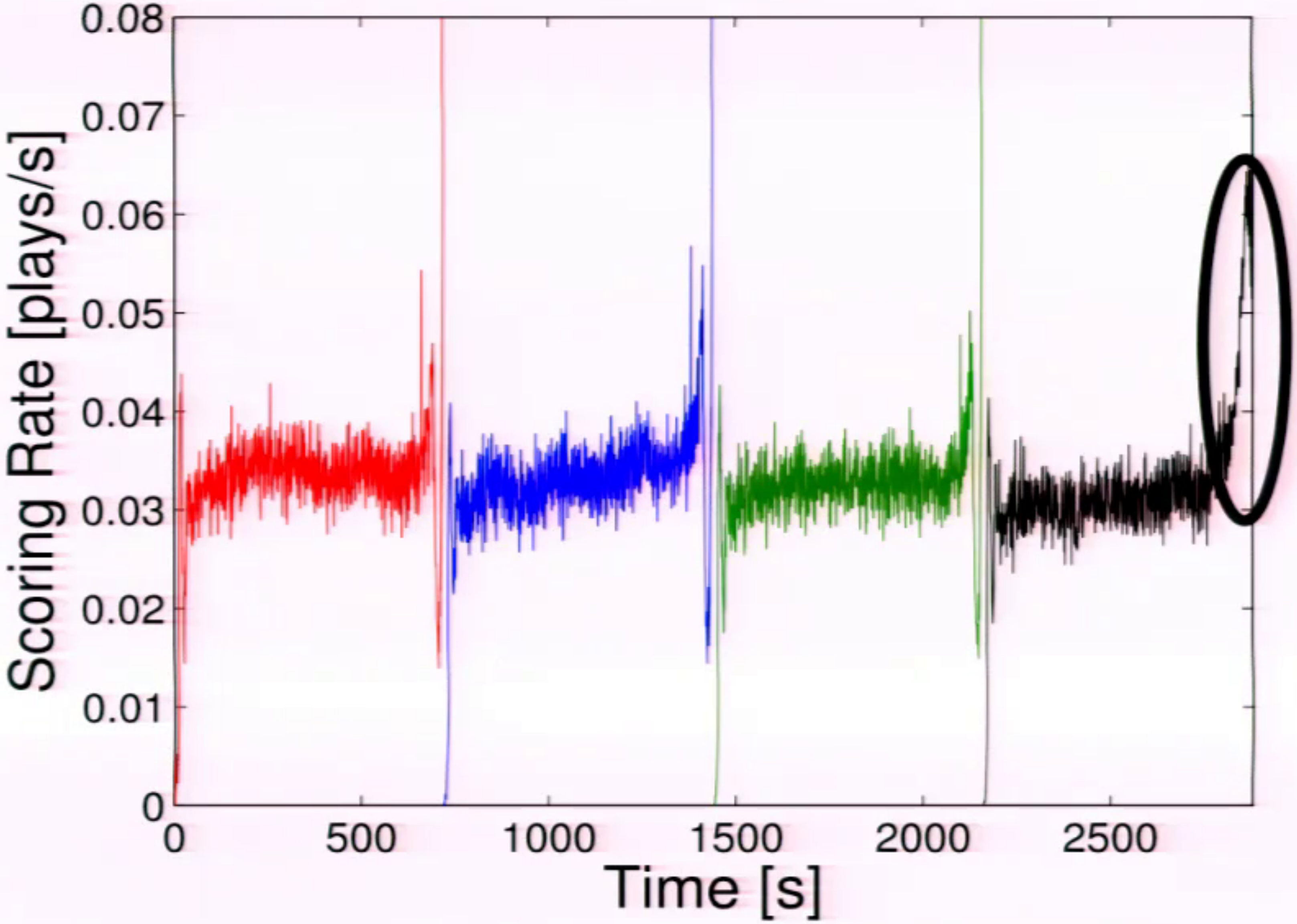
2015 SIAM Conference on Applications of Dynamical Systems

Can basketball scoring be described as a random walk?

Yes!

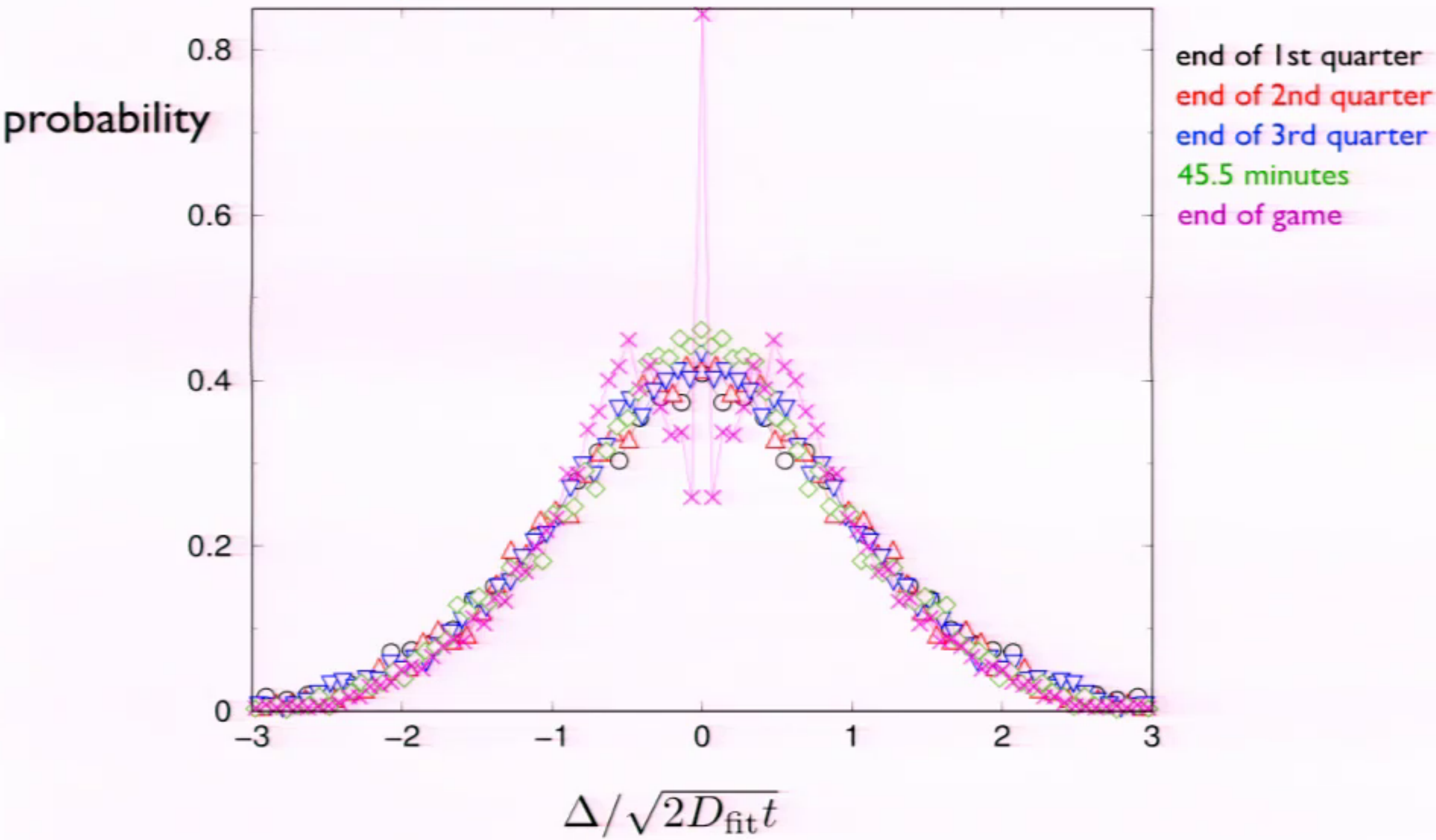
Some Buts……….

0. ***Last 2.5 minutes:*** anomalous.



Scoring Rate [plays/s]
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
500
1000
1500
2000
2500
Time [s]

Almost Gaussian Score Difference Distribution
except in the last 2.5 minutes



Some Buts..........

0. ***Last 2.5 minutes:*** anomalous.

1. ***Bias:*** Teams typically not that unevenly matched.

2. ***Antipersistence:*** After team A scores, same team scores next with probability 0.35.

3. ***Return to zero:*** winning team coasts, losing team desperate.