

Near-extremal points of a random walk and variations around Odlyzko's algorithm for the search of its maximum

Grégory Schehr

Labo. de Physique Théorique et Modèles Statistiques (LPTMS)
CNRS-Université Paris-Sud

Séminaire Philippe Flajolet, December 4, 2014

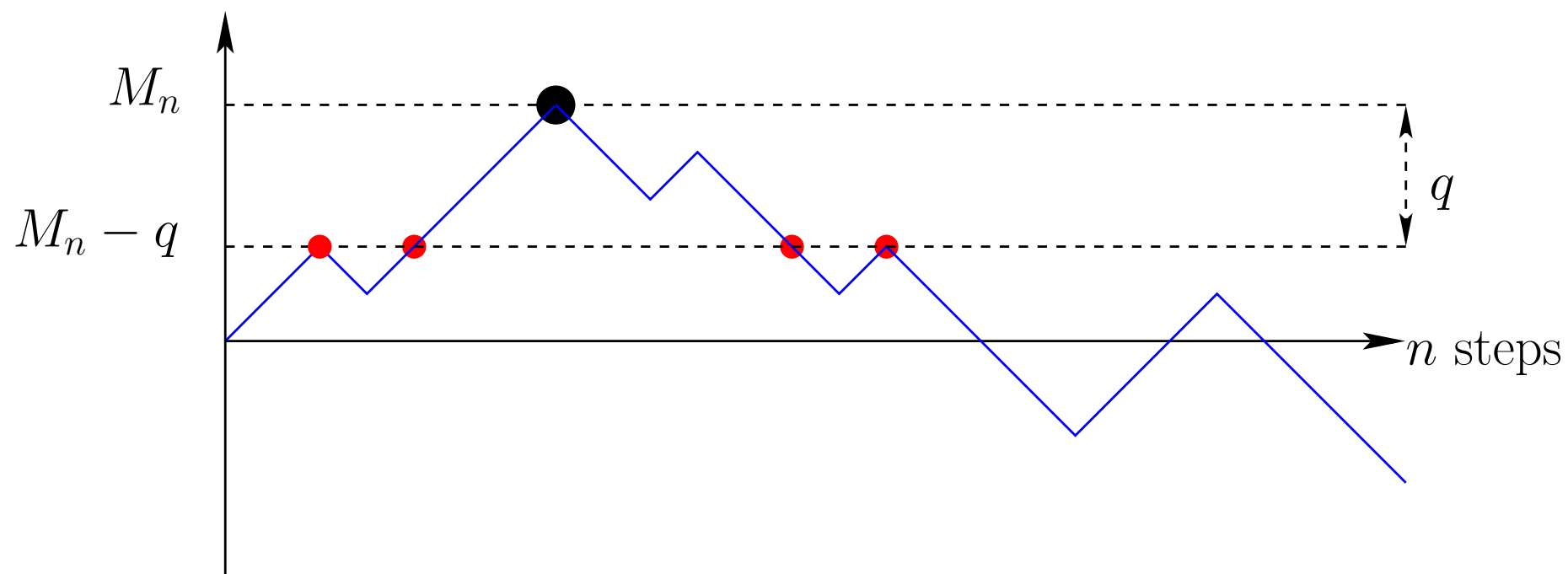
- Anthony Perret (LPTMS, Orsay)
- Alain Comtet (LPTMS, Orsay)
- Satya N. Majumdar (LPTMS, Orsay)

A. Perret, A. Comtet, S. N. Majumdar, G. S., Phys. Rev. Lett. **111**, 240601 (2013)
& preprint arXiv:1502.01218

Near-extremal points of random walks

$$X_0 = 0 ,$$

$$X_i = X_{i-1} + \epsilon_i , \quad \epsilon_i = \begin{cases} +1 , & \text{w. proba. } 1/2 \\ -1 , & \text{w. proba. } 1/2 \end{cases}$$

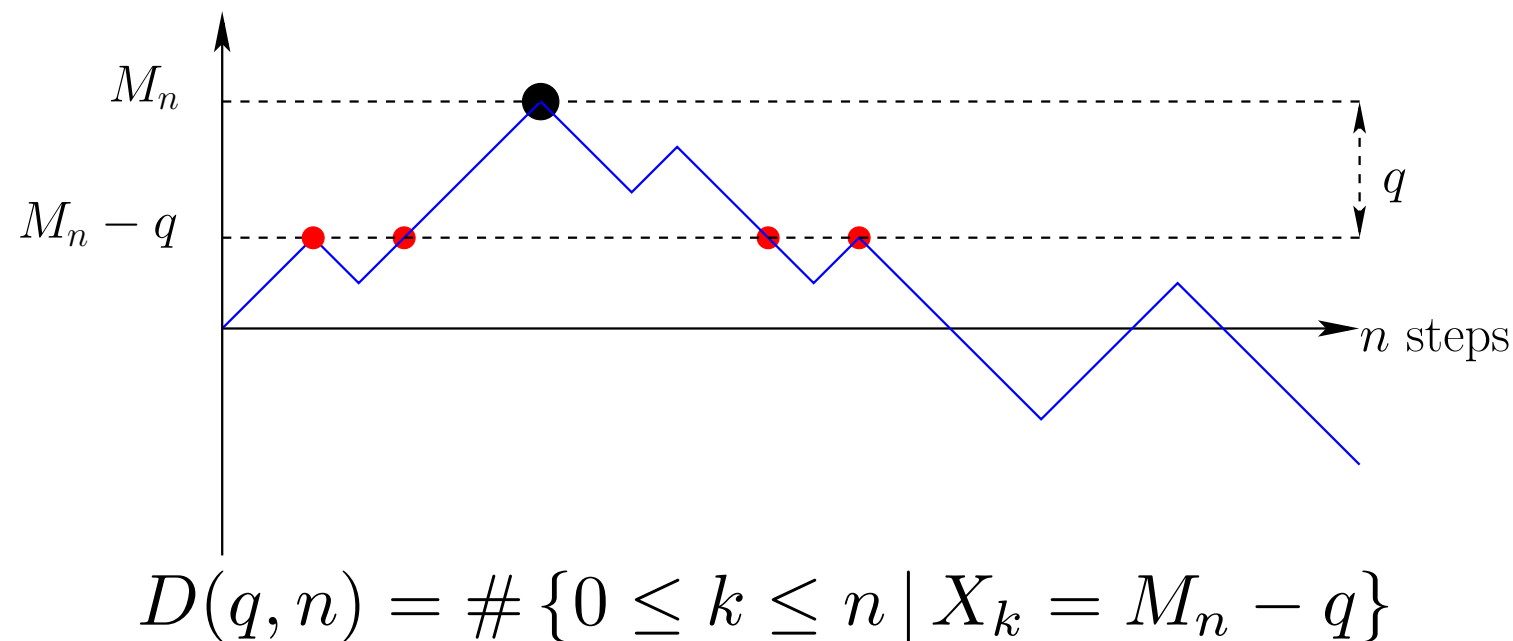


■ Number of near-extremal points

$$M_n = \max_{0 \leq k \leq n} X_k$$

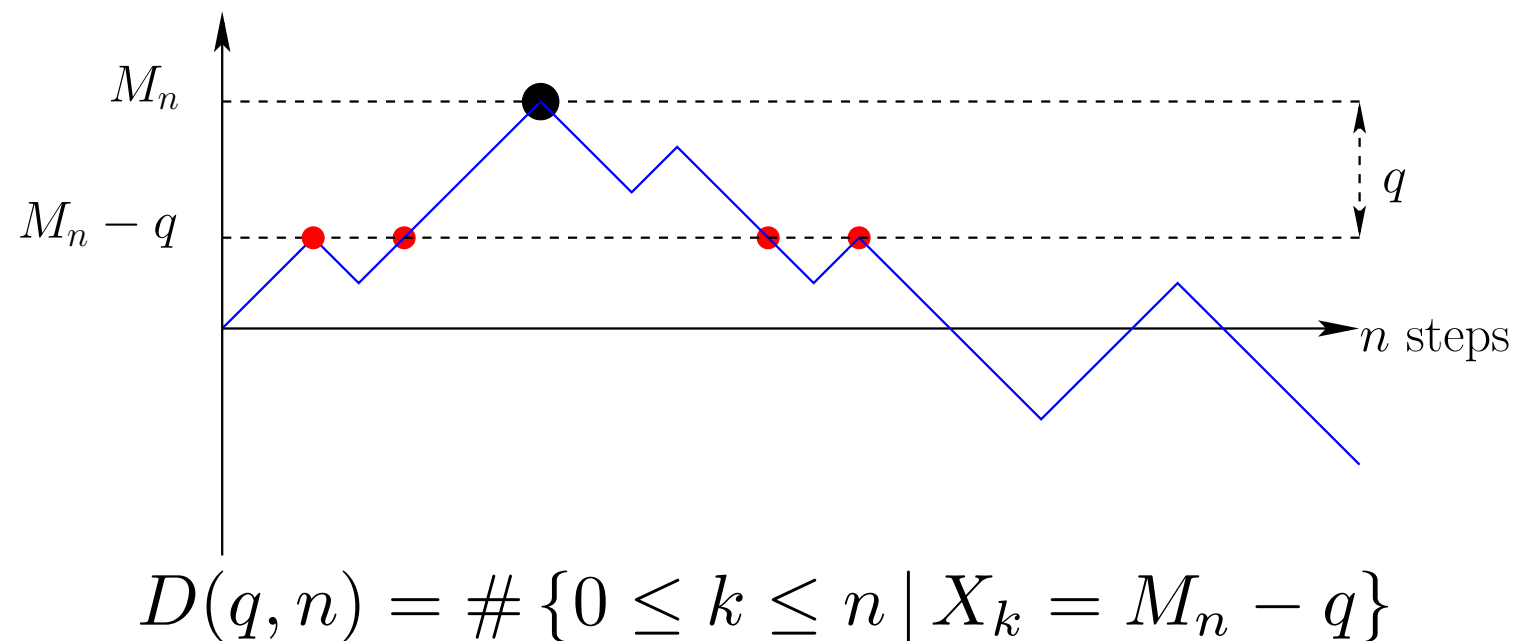
$$D(q, n) = \# \{0 \leq k \leq n \mid X_k = M_n - q\}$$

Related questions in the literature on random walks



- Local time of RW (and Brownian motion)
- Frequently and rarely visited sites : Erdős, Revesz,..., Toth
- Number of times a random walk is at its maximum: Csaki, Odlyzko

Motivations



- “Crowding” near the maximum: **is the maximum lonely at the top ?**

Sabhapandit, Majumdar '07

- Plays an important role in the analysis of the **optimal** algorithm to find the maximum of a random walk

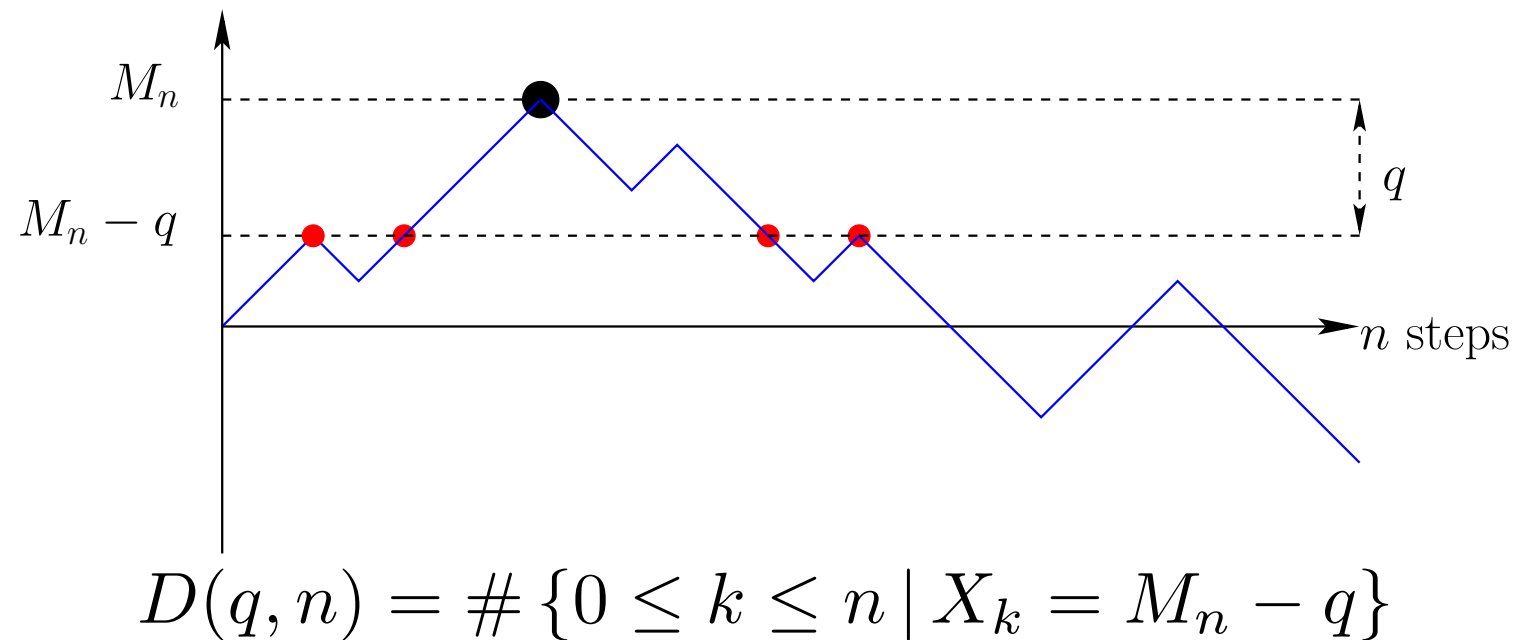
Odlyzko '95

Hwang '97, Chassaing '99

Chassaing, Marckert, Yor '99

- Functionals of the maximum of RW and Brownian motion

Local time of Brownian motion close to its maximum



- Asymptotic limit $n \rightarrow \infty$

$$\frac{1}{\sqrt{n}} D(q = \lfloor r \sqrt{n} \rfloor, n) \xrightarrow{\text{law}} \rho(r) = \int_0^1 \delta(x_{\max} - x(\tau) - r) d\tau$$

$x(\tau)$: **Brownian motion**

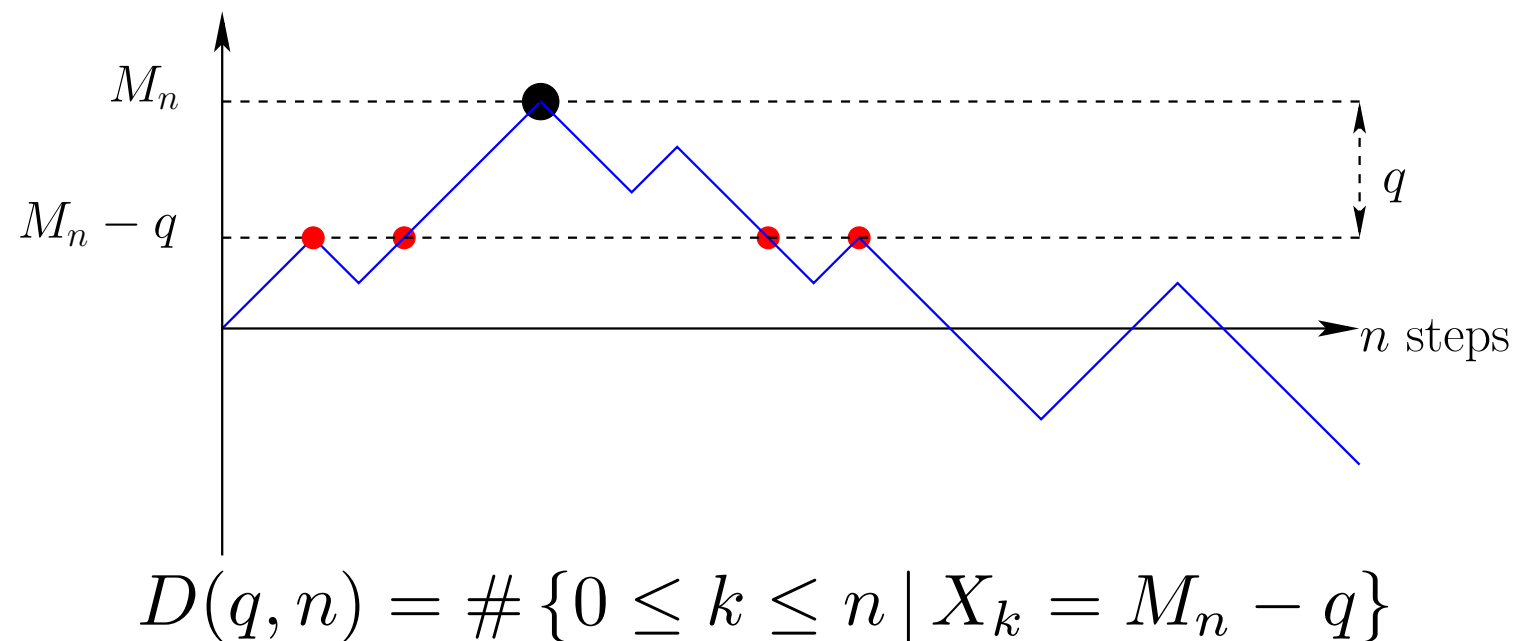
$$x_{\max} = \max_{0 \leq \tau \leq 1} x(\tau)$$

- Q: full statistics of the **density of near-extremes** $\rho(r)$?

$$\mathbb{E}[\rho^k(r)] = 8k! \sum_{l=0}^{k-1} (-1)^l \binom{k-1}{l} [(2l+1)\Phi^{(k+1)}((2l+1)r) + (k-2(l+1))\Phi^{(k+1)}(2(l+1)r)]$$

$$\Phi^{(0)}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \Phi^{(j+1)}(x) = \int_x^\infty \Phi^{(j)}(u) du$$

Motivations



- ☑ “Crowding” near the maximum: **is the maximum lonely at the top ?**

Sabhapandit, Majumdar '07

- Plays an important role in the analysis of the **optimal** algorithm to find the maximum of a random walk

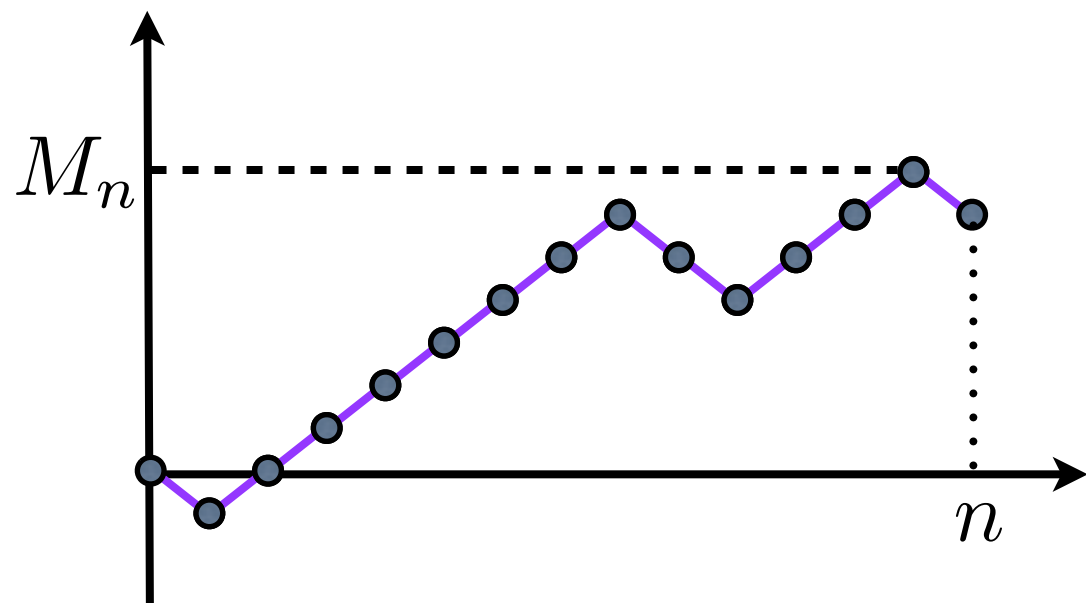
Odlyzko '95

Hwang '97, Chassaing '99

Chassaing, Marckert, Yor '99

- Applications to “functionals” of the RW and Brownian motion

Search algorithm for the maximum of a RW



■ $a \in A_n$: algorithm that finds M_n

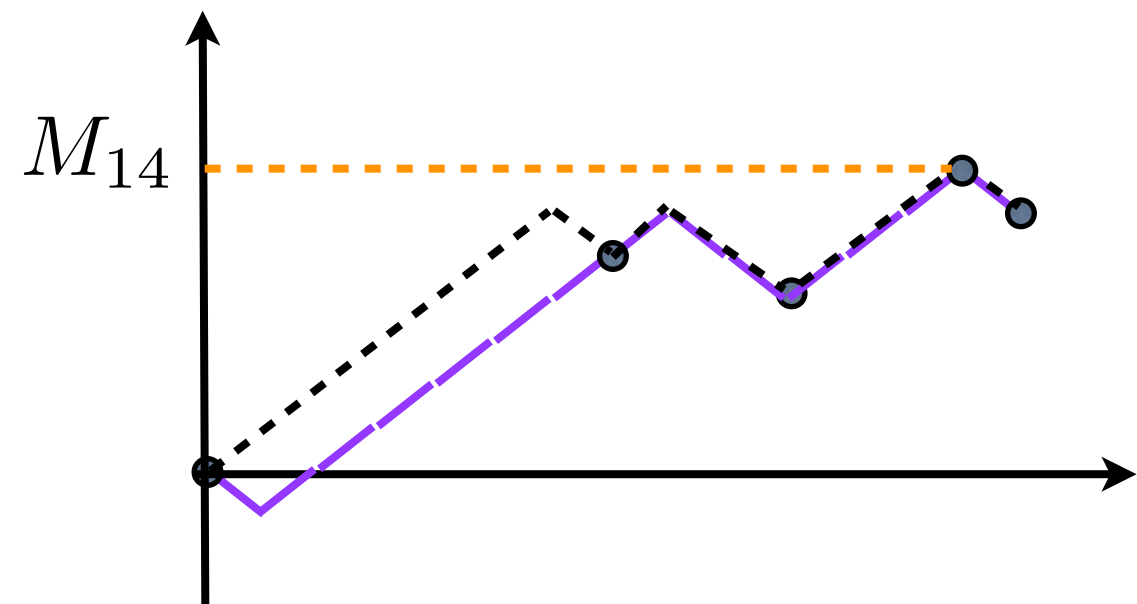
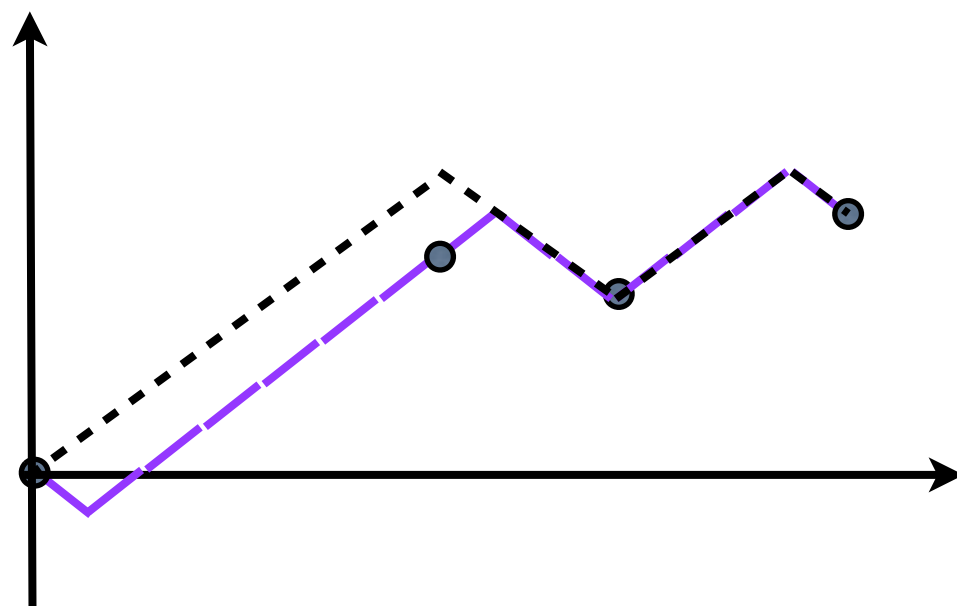
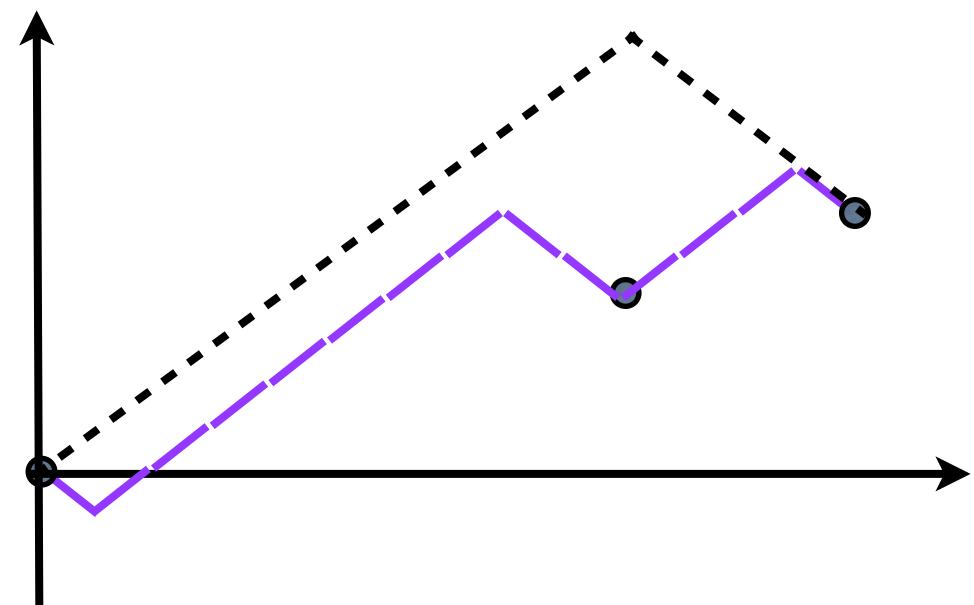
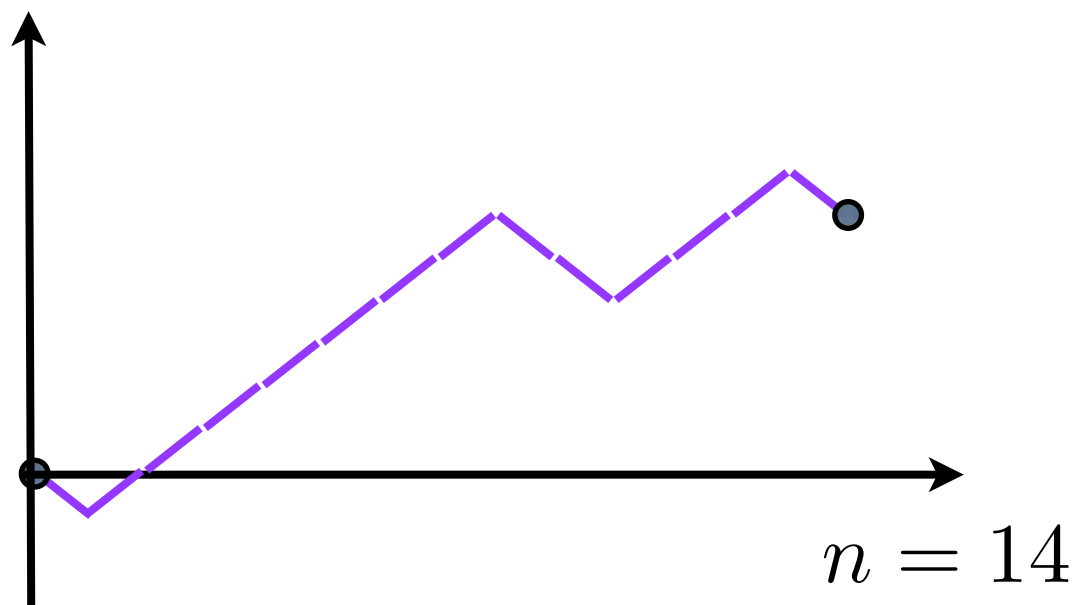
■ Cost of the algorithm:

$C(a)$ = number of probes needed

■ Q: what is the optimal algorithm ?

- The **simplest** algo. probes all the positions: its cost is n
- Because of the **correlations** between the positions of the random walker, one can usually do much better

Searching for the maximum of a RW: exploiting correlations



The maximum is found in 4 probes ($4 < 14$) !

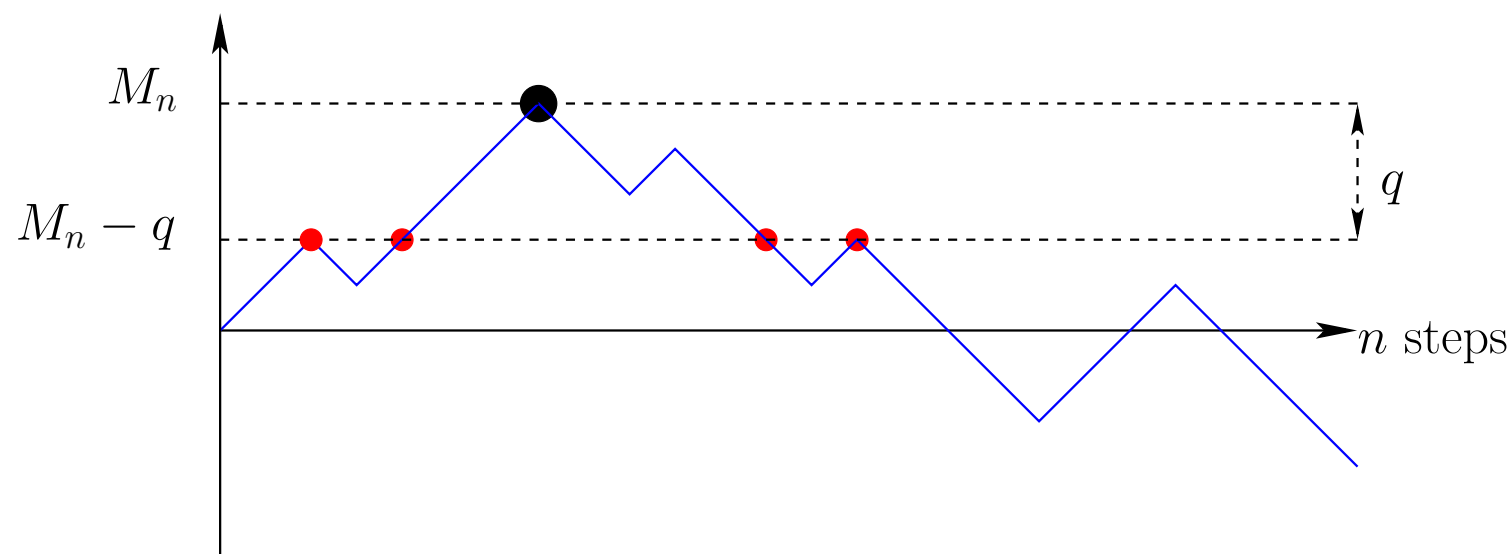
Searching for the maximum of a RW: optimal algorithm

■ Average case optimality

$$\min_{a \in A_n} \mathbb{E}(C(a)) = c_0 \sqrt{n} + o(\sqrt{n})$$

Odlyzko '95

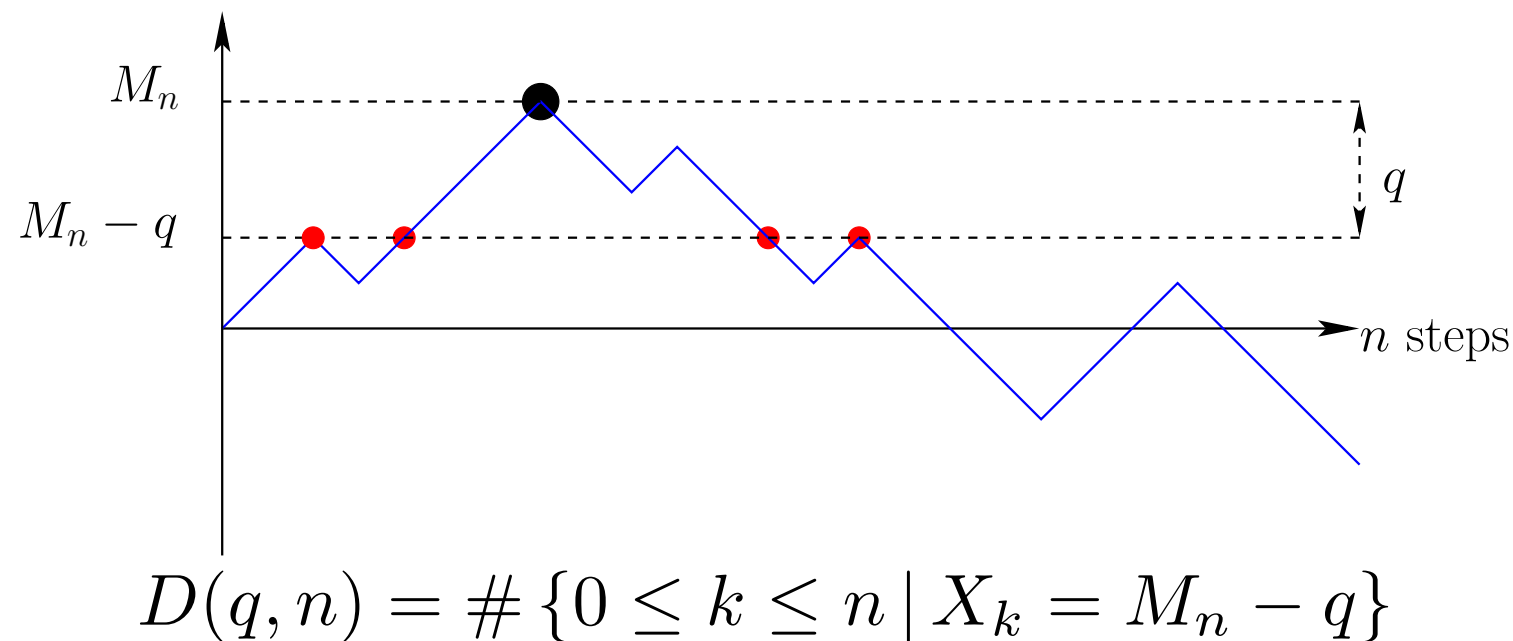
“In particular we need to prove that random walks do not spend much time close to their maxima.”



$$D(q, n) = \# \{0 \leq k \leq n \mid X_k = M_n - q\}$$

$$c_0 = \lim_{n \rightarrow \infty} \frac{1}{2\sqrt{n}} \sum_{q=0}^n \frac{\mathbb{E}[D(q, n)]}{q+1}$$

Motivations



- ❑ “Crowding” near the maximum: **is the maximum lonely at the top ?**

Sabhapandit, Majumdar '07

- ❑ Plays an important role in the analysis of the **optimal** algorithm to find the maximum of a random walk

Odlyzko '95

Hwang '97, Chassaing '99

Chassaing, Marckert, Yor '99

- Applications to “functionals” of the RW and Brownian motion

Searching for the maximum of a RW: optimal algorithm

■ Average case optimality

$$\min_{a \in A_n} \mathbb{E}(C(a)) = c_0 \sqrt{n} + o(\sqrt{n}) \quad \text{Odlyzko '95}$$

$$c_0 = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} \frac{dy}{y} \int_0^1 \frac{1}{\sqrt{w}} \exp\left(-\frac{y^2}{2w}\right) \operatorname{erf}\left(\frac{y}{\sqrt{2(1-w)}}\right) dw$$

after some manipulations...

$$c_0 = \sqrt{\frac{8}{\pi}} \log 2 \quad \text{Hwang '97, Chassaing '99}$$

■ Connection with a functional of the maximum of Brownian motion

Chassaing '99

Chassaing, Marckert, Yor '99

$x(\tau)$: **Brownian motion**

$$x_{\max} = \max_{0 \leq \tau \leq 1} x(\tau)$$

$$c_0 = \mathbb{E}(I) , \quad I = \frac{1}{2} \int_0^1 \frac{d\tau}{x_{\max} - x(\tau)}$$

Searching for the maximum of a RW: optimal algorithm

$$\min_{a \in A_n} \mathbb{E}(C(a)) = c_0 \sqrt{n} + o(\sqrt{n})$$

Odlyzko '95

$$c_0 = \mathbb{E}(I) , \quad I = \frac{1}{2} \int_0^1 \frac{d\tau}{x_{\max} - x(\tau)} \qquad c_0 = \sqrt{\frac{8}{\pi}} \log 2$$

- Odlyzko described an algorithm $\text{Od}(n)$ which is quasi-optimal

$$\mathbb{E}[C(\text{Od}(n))] = c_0 \sqrt{n} + o(\sqrt{n})$$

- $\text{Od}(n)$ is quasi-optimal **in distribution** (not only on average)
and it was shown that

Chassaing, Marckert, Yor '99

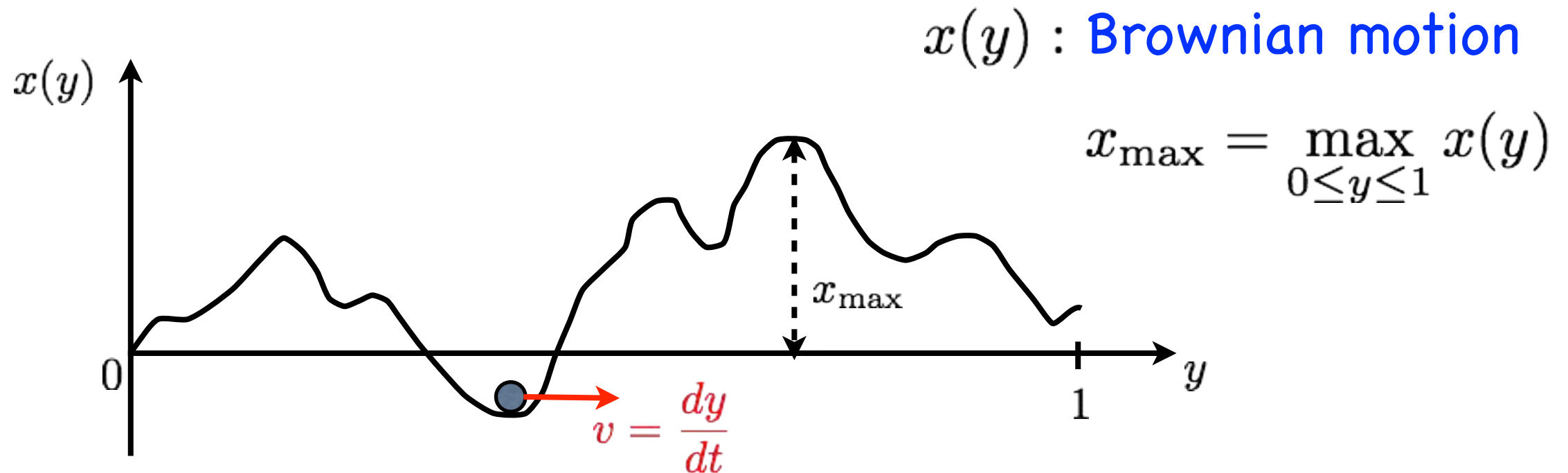
$$\lim_{n \rightarrow \infty} \Pr \left(\frac{C(\text{Od}(n))}{\sqrt{n}} \geq x \right) = \Pr(I \geq x)$$

➡ relevance of a **functional of the maximum** of Brownian motion

$$I = \int_0^1 V(x_{\max} - x(\tau)) d\tau , \quad V(x) = \frac{1}{2x}$$

Functionals of the maximum of BM in physics

- Largest exit time of classical particles moving ballistically through a disordered **Brownian** potential



the **slowest particle** that crosses the sample is such that

$$\frac{1}{2} \left(\frac{dy}{dt} \right)^2 + x(y) = x_{\max}$$

and the **largest time** to cross the sample is

$$\mathcal{T}_{\max} = \frac{1}{\sqrt{2}} \int_0^1 \frac{dy}{\sqrt{x_{\max} - x(y)}}$$

An interesting family of functionals of the maximum of BM

$$T_{\alpha}(t) = \int_0^t (x_{\max} - x(\tau))^{\alpha} d\tau$$

- For $\alpha = -1$ it describes the cost of Odlyzko's algorithm
- For $\alpha = -\frac{1}{2}$ it describes the largest time to cross a Brownian barrier
- For $\alpha = +1$ it describes an area or "Airy" type of random variable

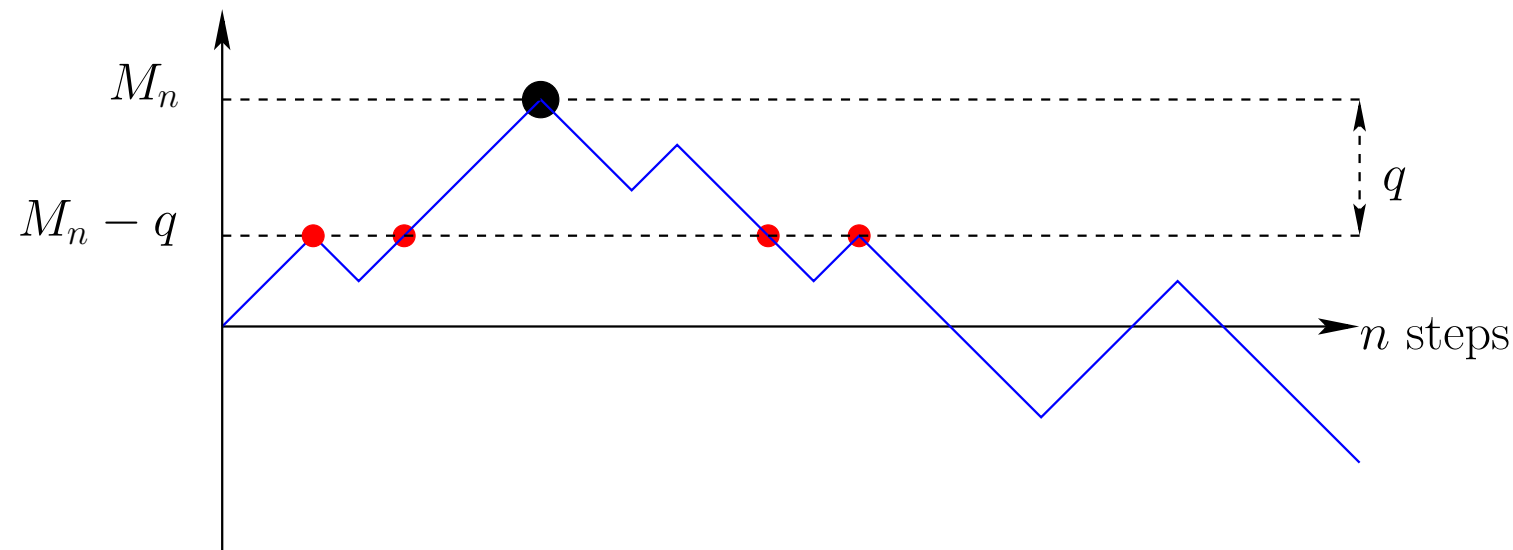


In this work we develop tools to study the statistics of such functionals of the maximum of BM

Outline

- Path counting method (based on propagators of BM)
- Feynman-Kac approach
- Applications of the Feynman-Kac approach
- Conclusion

Average number of near-extremal points for RW



$$D(q, n) = \# \{0 \leq k \leq n \mid X_k = M_n - q\}$$

- Odlyzko obtained an exact formula for $\mathbb{E}(D(q, n))$

Odlyzko '95

$$\mathbb{E}(D(q, n)) = \sum_{m=0}^n (A(n, m, q) + B(n, m, q))$$

$$A(n, m, q) = 2^{-n} \binom{m}{\lfloor \frac{m+q+1}{2} \rfloor} \sum_{j=0}^q \binom{n-m}{\lfloor \frac{n-m+j}{2} \rfloor}$$

$$B(b, m, q) = 2^{-n} \binom{m}{\lfloor \frac{n-m-q}{2} \rfloor} \sum_{j=0}^{q-1} \binom{m}{\lfloor \frac{m-q+1}{2} \rfloor + j}$$

What about the asymptotic limit $n \rightarrow \infty$?

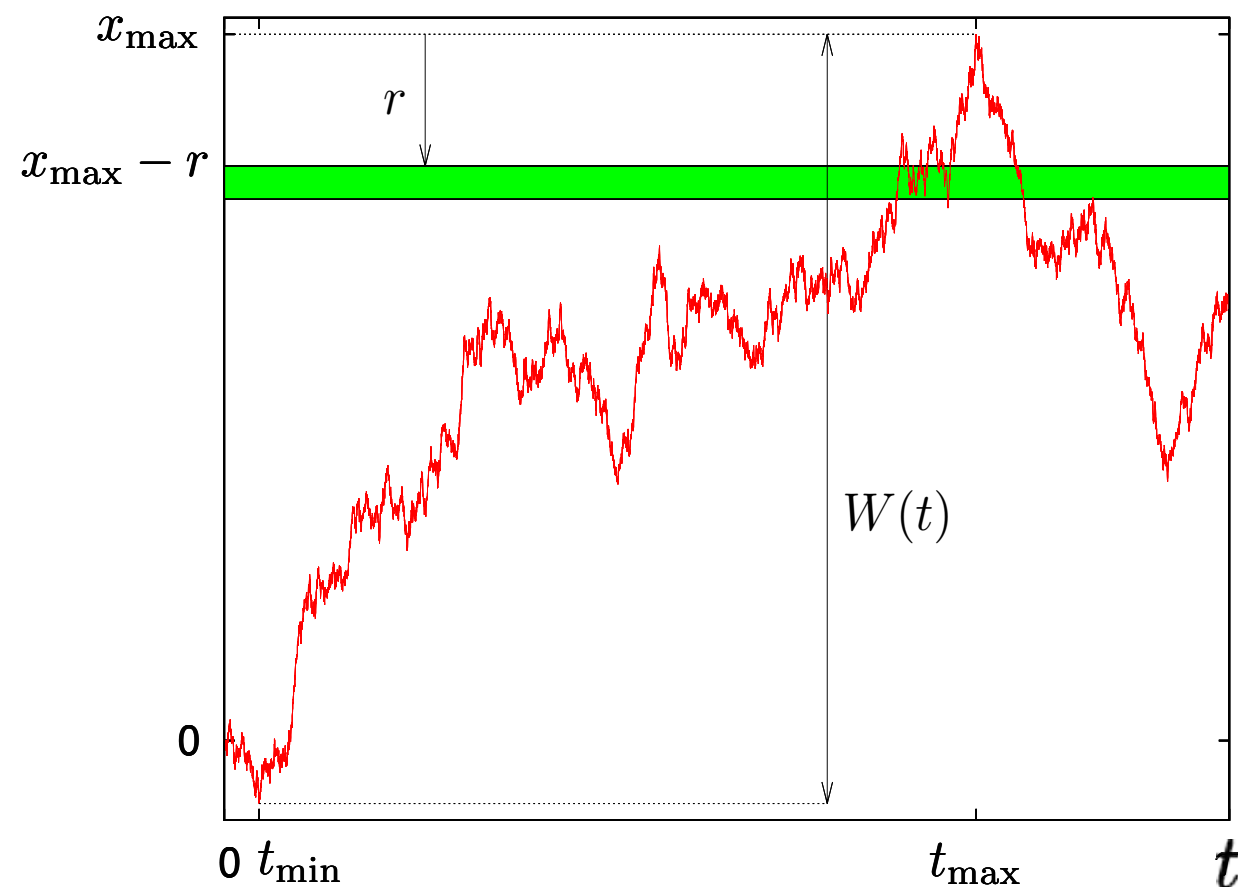
Density of near-extremes for Brownian motion

- Asymptotic limit $n \rightarrow \infty$

$$\frac{1}{\sqrt{n}} D(q = \lfloor r \sqrt{n} \rfloor, n) \xrightarrow{\text{law}} \rho(r) = \int_0^1 \delta(x_{\max} - x(\tau) - r) d\tau$$

$x(\tau)$: **Brownian motion**

$$x_{\max} = \max_{0 \leq \tau \leq 1} x(\tau)$$



$\rho(r)dr$: time spent by the BM in
 $[x_{\max} - r - dr, x_{\max} - r]$

$$\int_0^\infty \rho(r) dr = 1$$

$$\mathbb{E}[\rho(r)] = ?$$

Average density of near-extremes for Brownian motion

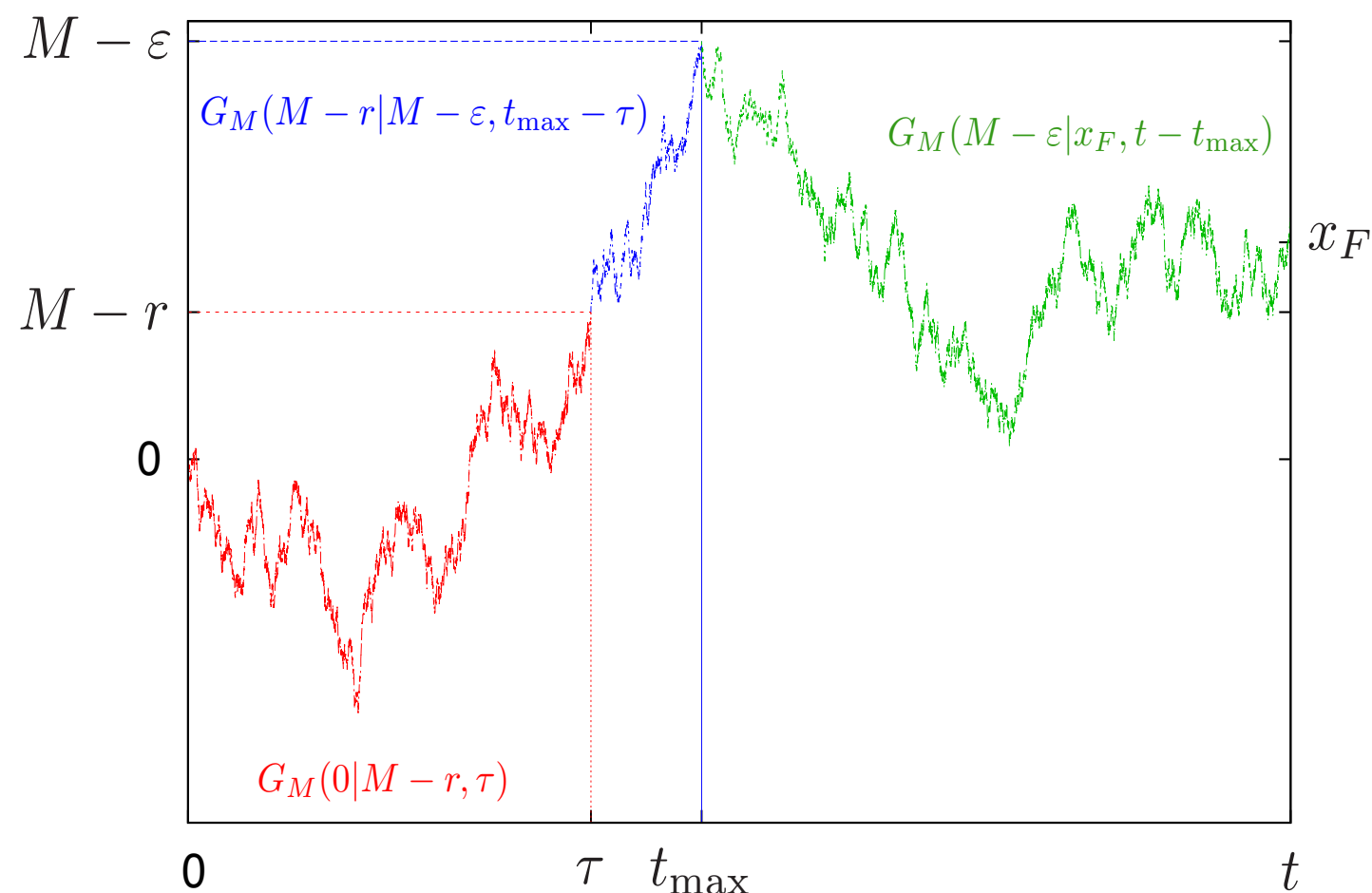
■ Propagator of Brownian motion

$$G_M(\alpha|\beta, t)d\beta = \text{Pr} . [x(t) \in [\beta, \beta + d\beta] \mid x(0) = \alpha \ \& \ x(\tau) < M \ \forall \tau \in [0, t]]$$

$$G_M(\alpha|\beta, t) = \frac{1}{\sqrt{2\pi t}} \left(e^{-\frac{(\beta-\alpha)^2}{2t}} - e^{-\frac{(2M-\beta-\alpha)^2}{2t}} \right)$$

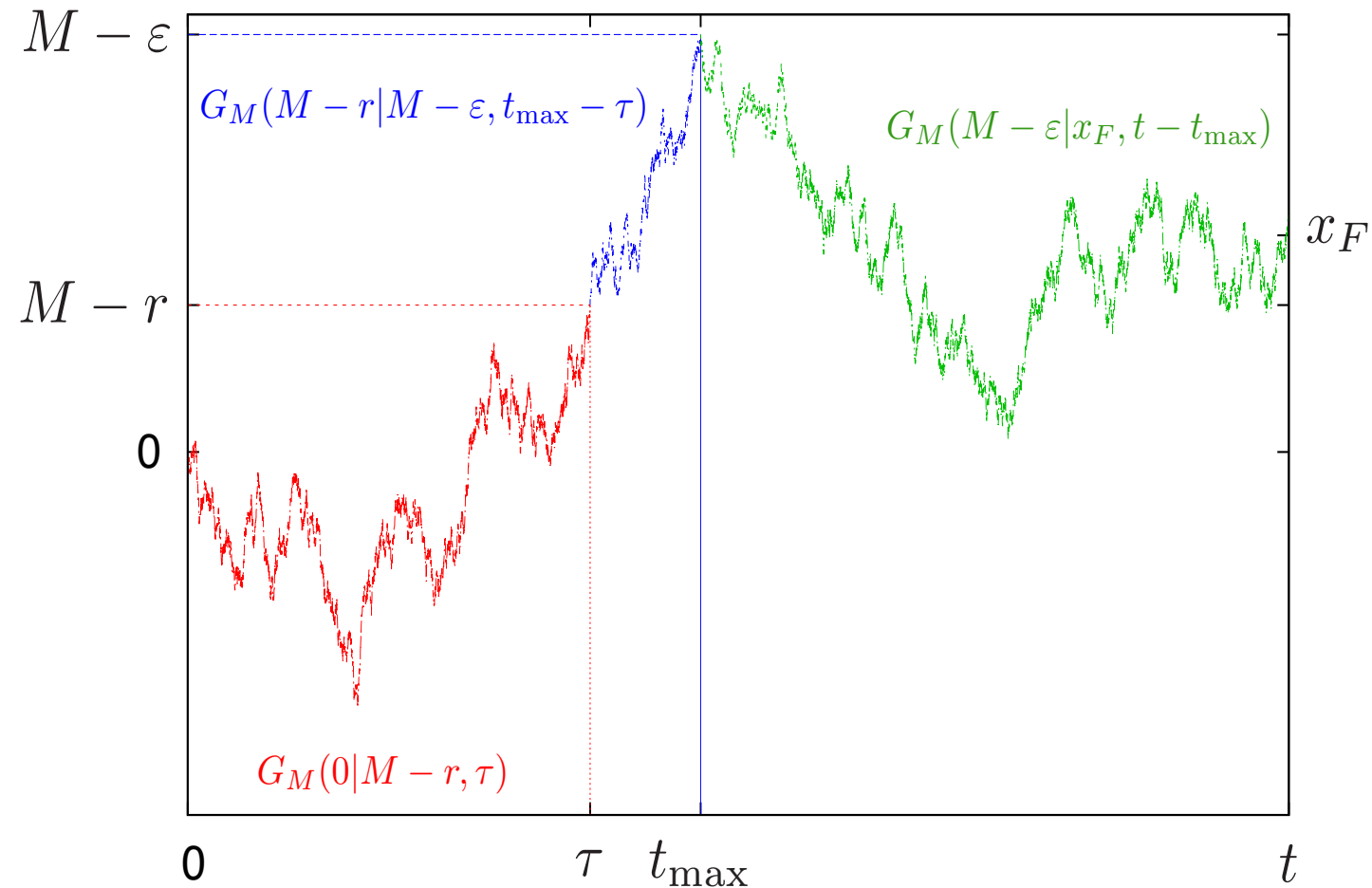
■ Average density of near-extremes: counting paths method

$$\mathbb{E}[\rho(r, t)] = \int_0^t \mathbb{E}[\delta(x_{\max} - x(\tau) - r)]d\tau$$



Average density of near-extreme: counting paths

$$\mathbb{E}[\rho(r, t)] = \int_0^t \mathbb{E}[\delta(x_{\max} - x(\tau) - r)] d\tau$$



$$\mathbb{E}[\rho(r, t)] = \lim_{\epsilon \rightarrow 0} \frac{2}{Z(\epsilon)} \int_0^\infty dM \int_0^t dt_{\max} \int_{-\infty}^M dx_F \int_0^{t_{\max}} d\tau \quad \textcolor{red}{G_M(0|M-r, \tau)} \textcolor{blue}{G_M(M-r|M-\epsilon, t_{\max} - \tau)} \\ \times \textcolor{green}{G_M(M-\epsilon|x_F, t - t_{\max})},$$

$$Z(\epsilon) = \int_0^\infty dM \int_0^t dt_{\max} \int_{-\infty}^M dx_F G_M(0|M-\epsilon, t_{\max}) G_M(M-\epsilon|x_F, t - t_{\max})$$

Average density of near-extremes: counting paths

$$\mathbb{E}[\rho(r, t)] = \int_0^t \mathbb{E}[\delta(x_{\max} - x(\tau) - r)] d\tau$$

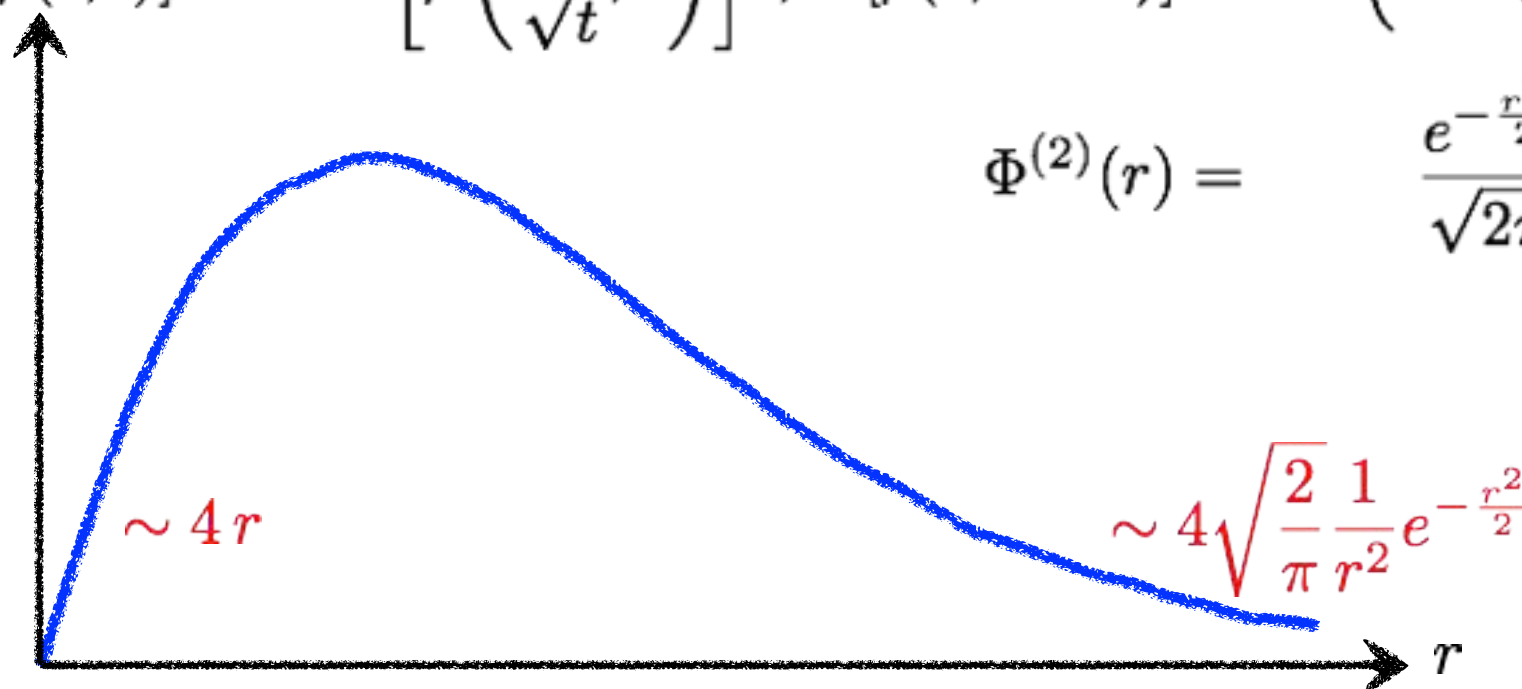
- Use of Laplace transform with respect to time t

$$\int_0^\infty dt e^{-st} \mathbb{E}[\rho(r, t)] = 8 \frac{e^{-\sqrt{2s}r} - e^{-2\sqrt{2s}r}}{(2s)^{3/2}}$$

- And finally...

$$\mathbb{E}[\rho(r, t)] = \sqrt{t} \mathbb{E} \left[\rho \left(\frac{r}{\sqrt{t}}, 1 \right) \right], \quad \mathbb{E}[\rho(r, t = 1)] = 8 \left(\Phi^{(2)}(r) - \Phi^{(2)}(2r) \right),$$

$$\Phi^{(2)}(r) = \frac{e^{-\frac{r^2}{2}}}{\sqrt{2\pi}} - \frac{r}{2} \operatorname{erfc} \left(\frac{r}{\sqrt{2}} \right)$$



Application to the average cost of Odlyzko's algorithm

- Average cost of the optimal algorithm for the search of the maximum of RW

$$\mathbb{E}[C(\text{Od}(n))] = c_0 \sqrt{n} + o(\sqrt{n})$$

$$c_0 = \mathbb{E}(I) , \quad I = \frac{1}{2} \int_0^1 \frac{d\tau}{x_{\max} - x(\tau)}$$

- Using the average density of near-extremes of Brownian motion

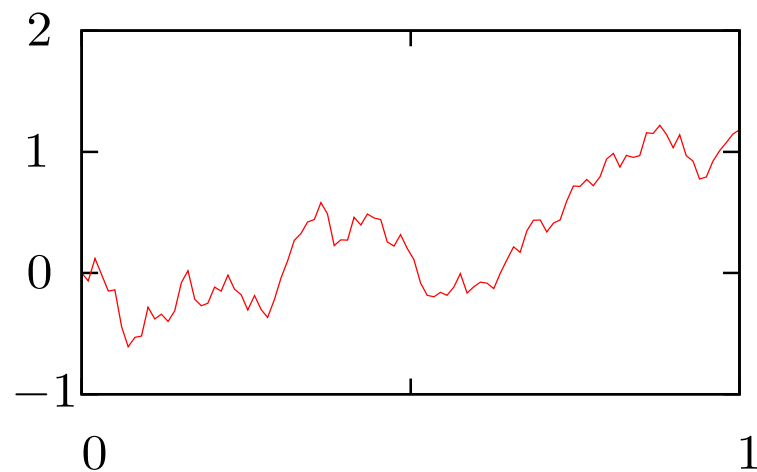
$$\mathbb{E}[\rho(r)] = \int_0^1 \mathbb{E}[\delta(x_{\max} - x(\tau) - r)] d\tau$$


$$c_0 = \frac{1}{2} \int_0^\infty \frac{\mathbb{E}[\rho(r)]}{r} dr = \sqrt{\frac{8}{\pi}} \log 2$$

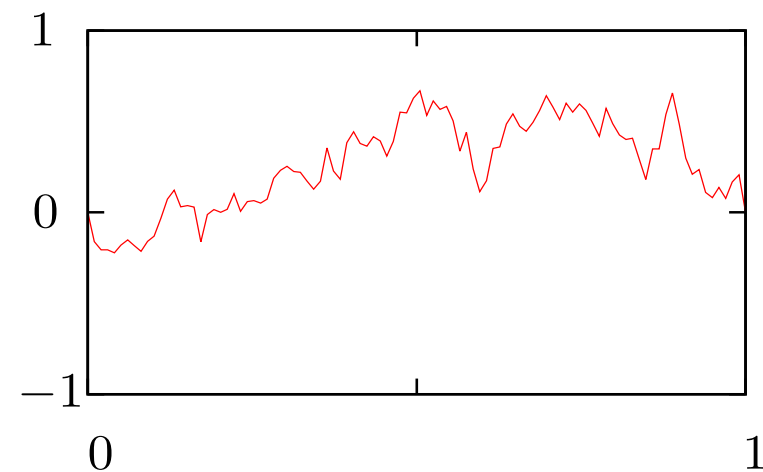
- For more general functionals of the maximum of BM

$$T_\alpha = \int_0^1 (x_{\max} - x(\tau))^\alpha d\tau , \quad \mathbb{E}[T_\alpha] = \int_0^\infty r^\alpha \rho(r) dr = \frac{(2 - 2^{-\alpha}) \Gamma(\frac{1+\alpha}{2})}{(2 + \alpha) \sqrt{\pi}}$$

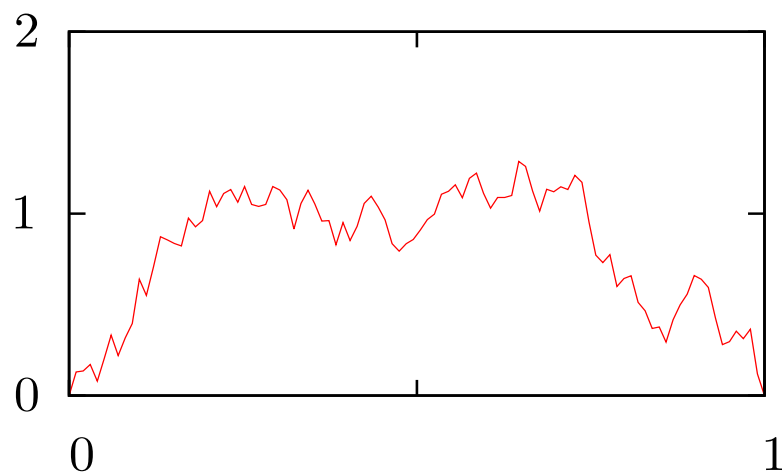
Average density of near-extremes of constrained BM



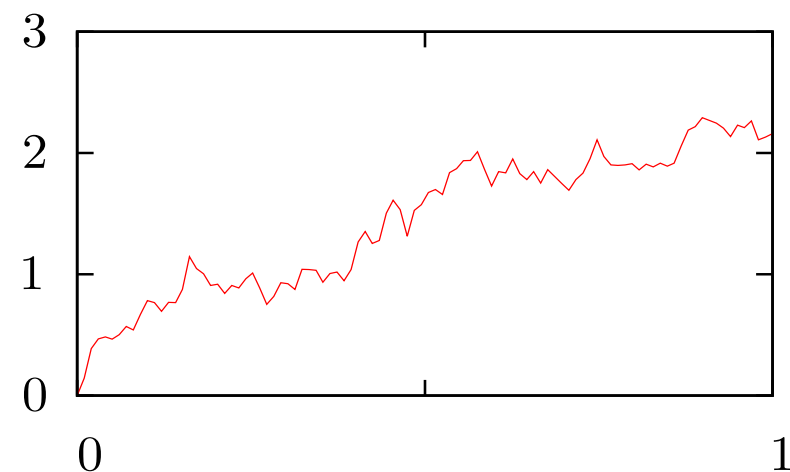
(a) Brownian motion



(b) Brownian bridge



(c) Brownian excursion



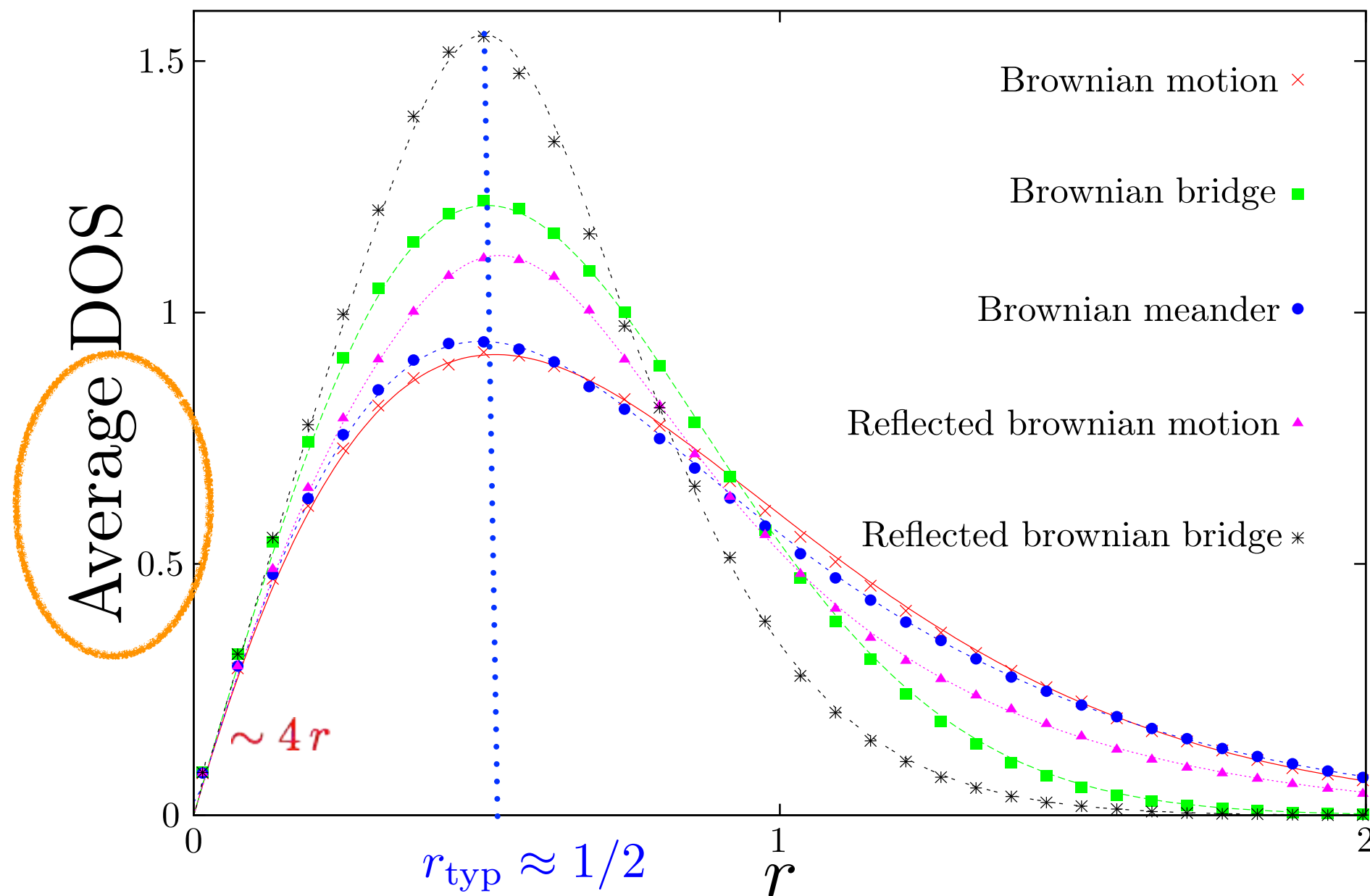
(d) Brownian Meander

For instance, for the Brownian meander:

$$\mathbb{E}[\rho_{Me}(r)] = \sqrt{2\pi} \left(\sum_{n=1}^{\infty} \frac{4n(-1)^n}{2n^2 + 3(-1)^n - 5} \operatorname{erfc}(nr) - \operatorname{erfc}(2r) \right)$$

A. Perret, A. Comtet, S. N. Majumdar, G. S., 2013 & 2014

Average density of near-extremes of constrained BM



A. Perret, A. Comtet, S. N. Majumdar, G. S., 2013 & 2014

What about higher moments of such functionals ?

Outline

- ☑ Path counting method (based on propagators of BM)

- ☐ Feynman-Kac approach

- ☐ Applications of the Feynman-Kac approach

- ☐ Conclusion

Exponential functionals of the maximum of BM

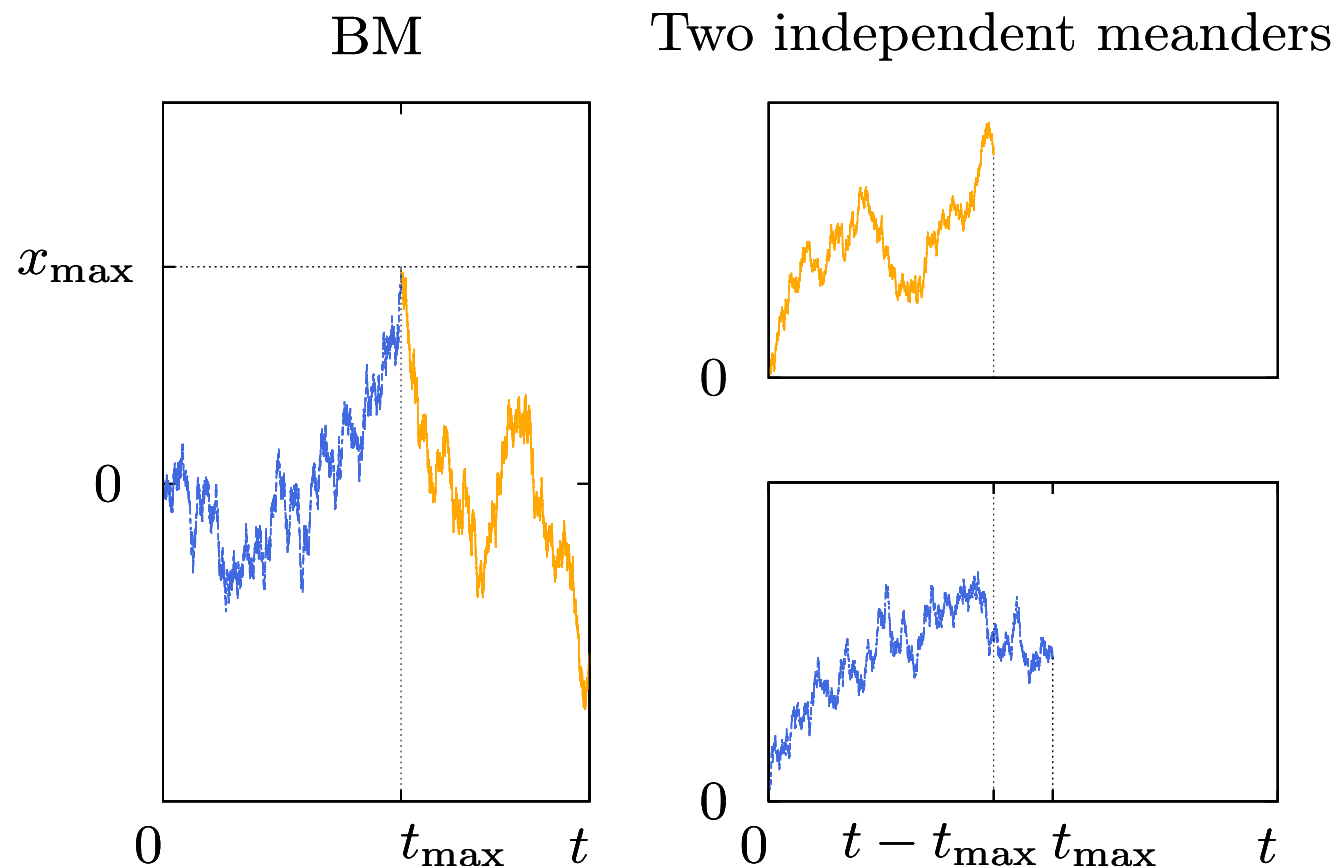
- Functional of the maximum of Brownian motion :

$$\mathcal{O}_{\max}(t) = \int_0^t V(x_{\max} - x(\tau)) d\tau$$

- Goal: compute the Laplace transform of the PDF of $\mathcal{O}_{\max}(t)$:

$$\mathbb{E} \left[\exp \left(-\lambda \int_0^t V(x_{\max} - x(\tau)) d\tau \right) \right]$$

- Decompose the path into **two independent meanders**

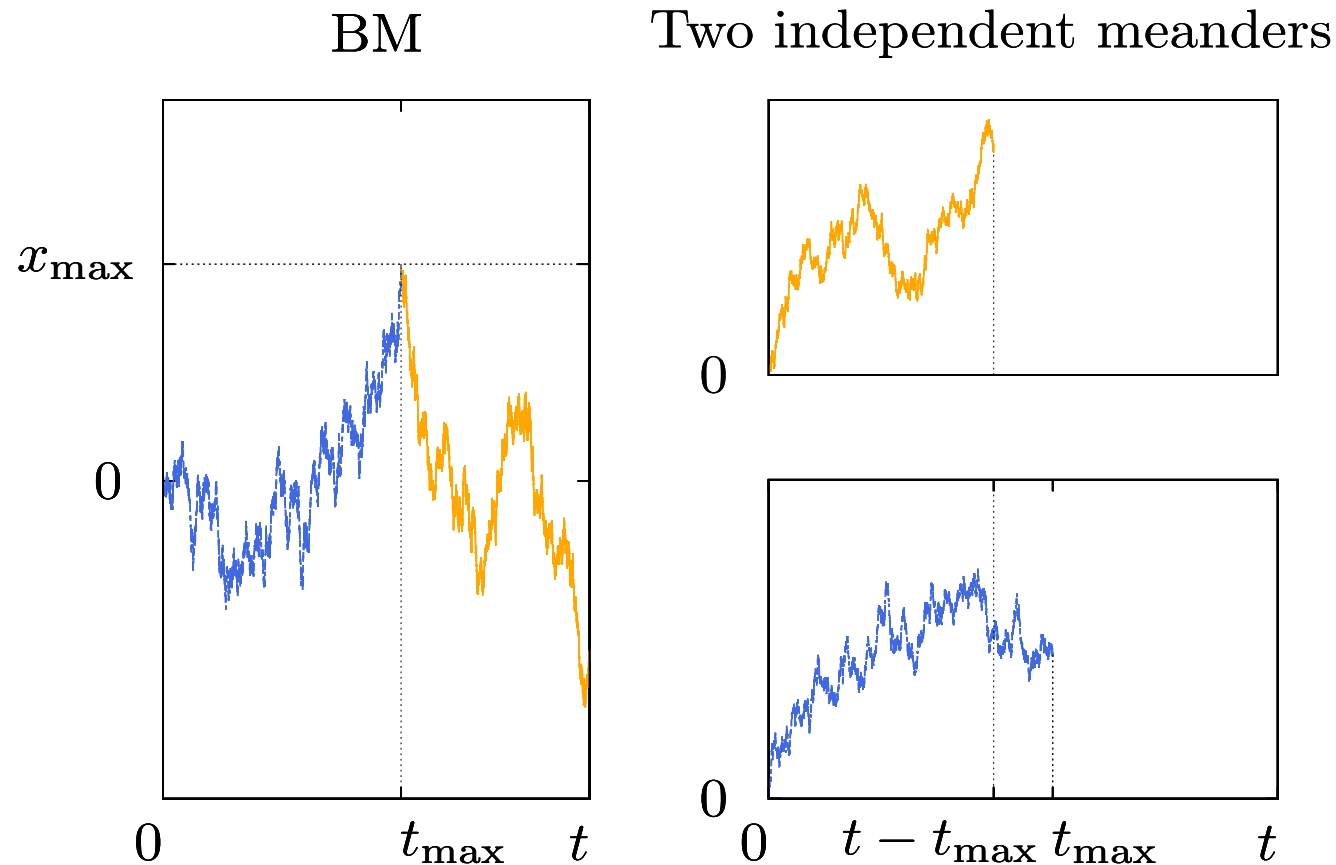


$$\Pr. (t_{\max} \leq T) = \int_0^T \frac{dx}{\pi \sqrt{x(t-x)}}$$

Lévy 's arcsine law

Reducing to functionals of the Brownian meander

- Decompose the path into **two independent meanders**



$$\Pr. (t_{\max} \leq T) = \int_0^T \frac{dx}{\pi \sqrt{x(t-x)}}$$

Lévy's arcsine law

$$\mathbb{E}\left[e^{-\lambda \int_0^t d\tau V(x_{\max} - x(\tau))}\right] = \int_0^t dt_{\max} \varphi(t_{\max}) \varphi(t - t_{\max})$$

$$\varphi(\tau) = \frac{1}{\sqrt{\pi\tau}} \mathbb{E}_+ \left[\exp \left(-\lambda \int_0^\tau du V(x_{\text{Me}}(u)) \right) \right]$$



Back to functionals of the Brownian meander

Feynman-Kac formula

$$\mathbb{E}[e^{-\lambda \int_0^t d\tau V(x_{\max} - x(\tau))}] = \int_0^t dt_{\max} \varphi(t_{\max}) \varphi(t - t_{\max})$$

$$\varphi(\tau) = \frac{1}{\sqrt{\pi\tau}} \mathbb{E}_+ \left[\exp \left(-\lambda \int_0^\tau du V(x_{\text{Me}}(u)) \right) \right]$$

- Laplace transform with respect to time t

$$\int_0^\infty e^{-st} \mathbb{E} \left(e^{-\lambda \int_0^t V(x_{\max} - x(\tau)) d\tau} \right) dt = [\tilde{\varphi}(s)]^2 ,$$

$$\tilde{\varphi}(s) = \int_0^\infty e^{-st} \varphi(t) dt$$

which can be computed as

$$\tilde{\varphi}(s) = \frac{\sqrt{2}}{W} \int_0^\infty dy_F u'_s(0) v_s(y_F)$$

where $u_s(x), v_s(x)$ are two independent solutions of

$$\left(-\frac{1}{2} \frac{d^2}{dx^2} + \lambda V(x) + s \right) \psi(x) = 0$$

such that $\lim_{x \rightarrow 0} u_s(x) = 0$ and $W = u'_s(x) v_s(x) - u_s(x) v'_s(x)$

$$\lim_{x \rightarrow +\infty} v_s(x) = 0$$

Outline

☒ Path counting method (based on propagators of BM)

☒ Feynman-Kac approach

☐ Applications of the Feynman-Kac approach

☐ Conclusion

Feynman-Kac formula: application to the density of near-extremes

$$\mathcal{O}_{\max}(t) = \int_0^t V(x_{\max} - x(\tau)) d\tau$$

- The density of near-extremes $\rho(r, t) = \int_0^t \delta(x_{\max} - x(\tau) - r) d\tau$

corresponds to $V(x) = \delta(x - r)$

- Applying the general formalism to this specific case

Feynman-Kac formula

$$\mathbb{E}[e^{-\lambda \int_0^t d\tau V(x_{\max} - x(\tau))}] = \int_0^t dt_{\max} \varphi(t_{\max}) \varphi(t - t_{\max})$$

$$\varphi(\tau) = \frac{1}{\sqrt{\pi\tau}} \mathbb{E}_+ \left[\exp \left(-\lambda \int_0^\tau du V(x_{\text{Me}}(u)) \right) \right]$$

- Laplace transform with respect to time t

$$\int_0^\infty e^{-st} \mathbb{E} \left(e^{-\lambda \int_0^t V(x_{\max} - x(\tau)) d\tau} \right) dt = [\tilde{\varphi}(s)]^2 ,$$

$$\tilde{\varphi}(s) = \int_0^\infty e^{-st} \varphi(t) dt$$

which can be computed as

$$\tilde{\varphi}(s) = \frac{\sqrt{2}}{W} \int_0^\infty dy_F u'_s(0) v_s(y_F)$$

where $u_s(x), v_s(x)$ are two independent solutions of

$$\left(-\frac{1}{2} \frac{d^2}{dx^2} + \lambda V(x) + s \right) \psi(x) = 0$$

such that $\lim_{x \rightarrow 0} u_s(x) = 0$ and $W = u'_s(x) v_s(x) - u_s(x) v'_s(x)$

$$\lim_{x \rightarrow +\infty} v_s(x) = 0$$

Feynman-Kac formula: application to the density of near-extremes

$$\mathcal{O}_{\max}(t) = \int_0^t V(x_{\max} - x(\tau)) d\tau$$

- The density of near-extremes $\rho(r, t) = \int_0^t \delta(x_{\max} - x(\tau) - r) d\tau$

corresponds to $V(x) = \delta(x - r)$

- Applying the general formalism to this specific case

$$\int_0^\infty e^{-st} \mathbb{E} \left(e^{-\lambda \rho(r, t)} \right) dt = \frac{1}{s} \left(\frac{\sqrt{2s} + \lambda(1 - e^{-\sqrt{2sr}})^2}{\sqrt{2s} + \lambda(1 - e^{-2\sqrt{2sr}})} \right)^2$$

Feynman-Kac formula: applications to the density of near-extremes

From

$$\int_0^\infty e^{-st} \mathbb{E} \left(e^{-\lambda \rho(r,t)} \right) dt = \frac{1}{s} \left(\frac{\sqrt{2s} + \lambda(1 - e^{-\sqrt{2s}r})^2}{\sqrt{2s} + \lambda(1 - e^{-2\sqrt{2s}r})} \right)^2$$

we obtain the moments of arbitrary order

$$\mathbb{E}(\rho^k(r, t=1)) = 8k! \sum_{l=0}^{k-1} (-1)^l \binom{k-1}{l} [(2l+1)\Phi^{(k+1)}((2l+1)r) + (k-2(l+1))\Phi^{(k+1)}(2(l+1)r)]$$

where $\frac{e^{-\sqrt{2s}u}}{(\sqrt{2s})^{j+1}} = \int_0^\infty t^{\frac{j-1}{2}} \Phi^{(j)} \left(\frac{u}{\sqrt{t}} \right) e^{-st} dt$

These functions were studied in detail in [Chassaing & Louchard '02](#)

Feynman-Kac formula: applications to the cost of Odlyzko's algorithm

$$\mathcal{O}_{\max}(t) = \int_0^t V(x_{\max} - x(\tau)) d\tau$$

- The cost of Odlyzko's algorithm is given by $I = \frac{1}{2} \int_0^t \frac{d\tau}{x_{\max} - x(\tau)}$

which corresponds to $V(x) = \frac{1}{2x}$

- Applying the general formalism to this specific case

$$\int_0^\infty e^{-st} \mathbb{E} \left[e^{-\frac{\lambda}{2} \int_0^t \frac{d\tau}{x_{\max} - x(\tau)}} \right] dt = \frac{4}{s} \sum_{n=0}^\infty (-1)^n \frac{\lambda^n}{2^n (\sqrt{2s})^n} \sum_{k=0}^n \tilde{\zeta}(k) \tilde{\zeta}(n-k)$$
$$\tilde{\zeta}(k) = \sum_{n=1}^\infty \frac{(-1)^{n+1}}{n^k}$$

recovering, in a quite different way, the result of Chassaing, Marckert, Yor '99

Feynman-Kac formula: applications to more general functionals

$$T_\alpha(t) = \int_0^t (x_{\max} - x(\tau))^\alpha d\tau$$

corresponding to $V(x) = x^\alpha$

■ Exact results for the second moment

$$\mathbb{E}(T_\alpha^2(t)) = \frac{t^{2+\alpha}}{2^{3\alpha}\Gamma(3+\alpha)} \left(\Gamma(\alpha+1)^2(2^\alpha-1)(2^{\alpha+1}-1) + \frac{\Gamma(3+2\alpha)(4^{\alpha+1}-1)}{4(1+\alpha)^2} + \sum_{n=1}^{\infty} \frac{\Gamma(2+2\alpha+n)}{n!2^{1+n}(1+\alpha+n)} \right)$$

Outline

- ☑ Path counting method (based on propagators of BM)
- ☑ Feynman-Kac approach
- ☑ Applications of the Feynman-Kac approach
- ☐ Conclusion

Conclusion and perspectives

- Various tools to study the statistics of functionals of the maximum of Brownian motion
- Application to the full statistics of near-extremal points of long random walks
- Alternative method to study the cost of Odlyzko's optimal algorithm to find the maximum of long random walks
- Extension of these techniques to system with several random walkers ?
- What about more general stable processes ?