

Bandittisme og kvanteagenter

Boye Gravningen Sjo

NTNU

Masterpresentasjon 2023

Banditter

Spillet

Gitt k «armer», tidshorisont T .

Ukjent belønningsfordelinger p_1, \dots, p_k .

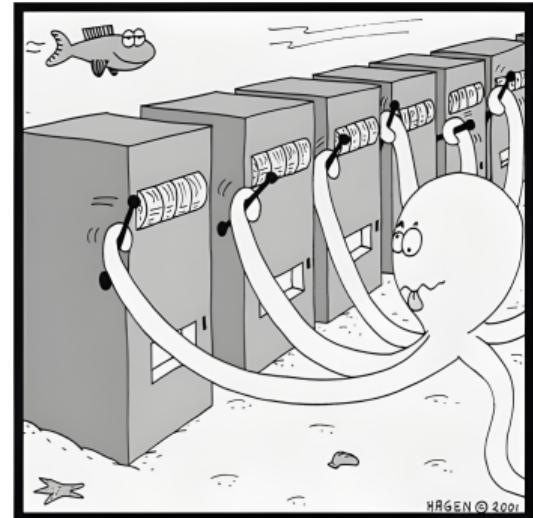
Antatt p_a tilhører en gitt klasse \mathcal{P} .

Regler for hver runde $t = 1, \dots, T$:

velg arm $a_t \in \{1, \dots, k\}$

motta belønning $X_a \sim p_a$.

Mål maksimer $\mathbb{E} \left[\sum_{t=1}^T X_t \right]$.



Utnyttelse versus utforskning

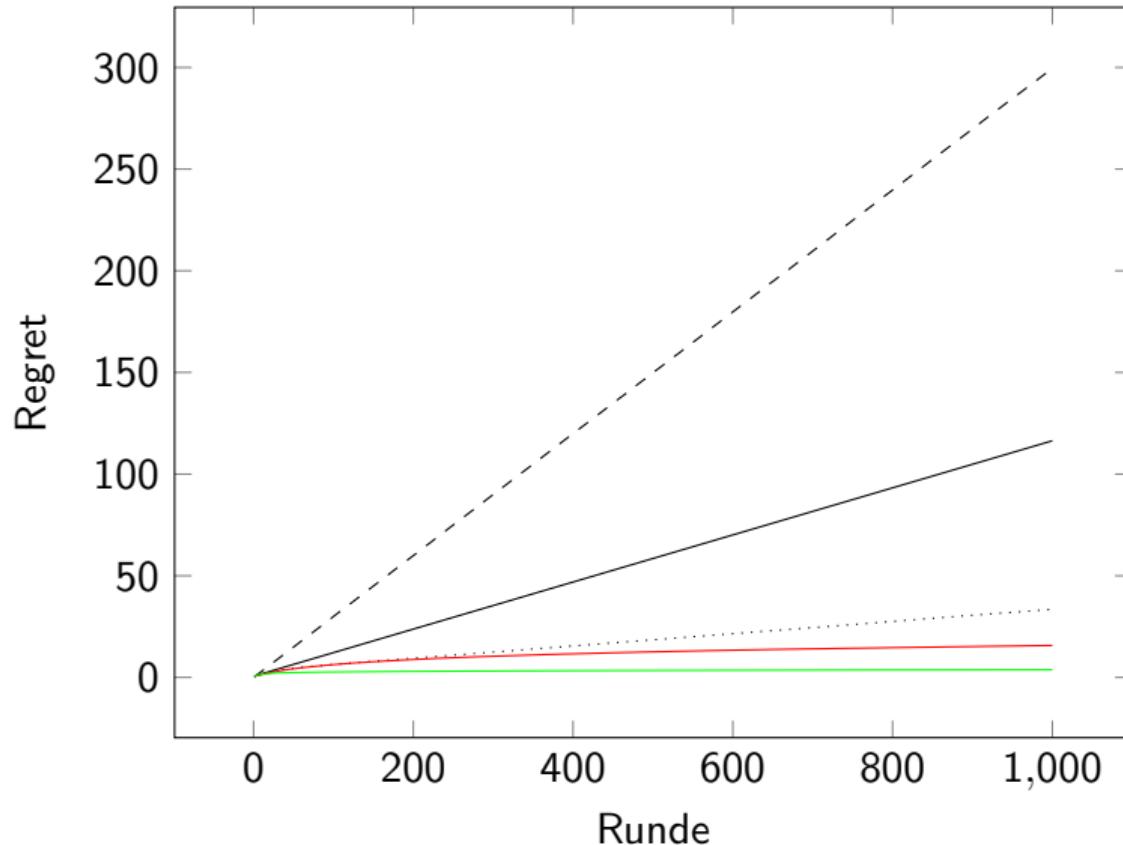
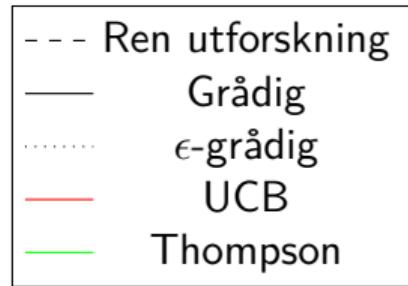
- Dra i den beste arm (kan være feil)
- Prøve andre armer (kaster bort runder)

Anvendelser

- Reklame på nett
- Kliniske tester
- Design av nettsider
- Investeringer
- Anbefalingssystemer
- Dynamisk prising
- Oppgavedelegering
- &c &c &c ...



Klassiske strategier



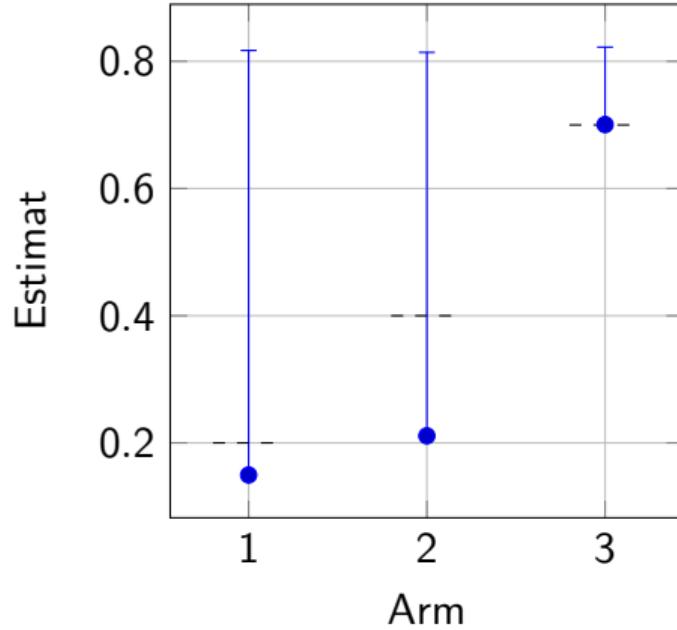
Kvantemaskiner



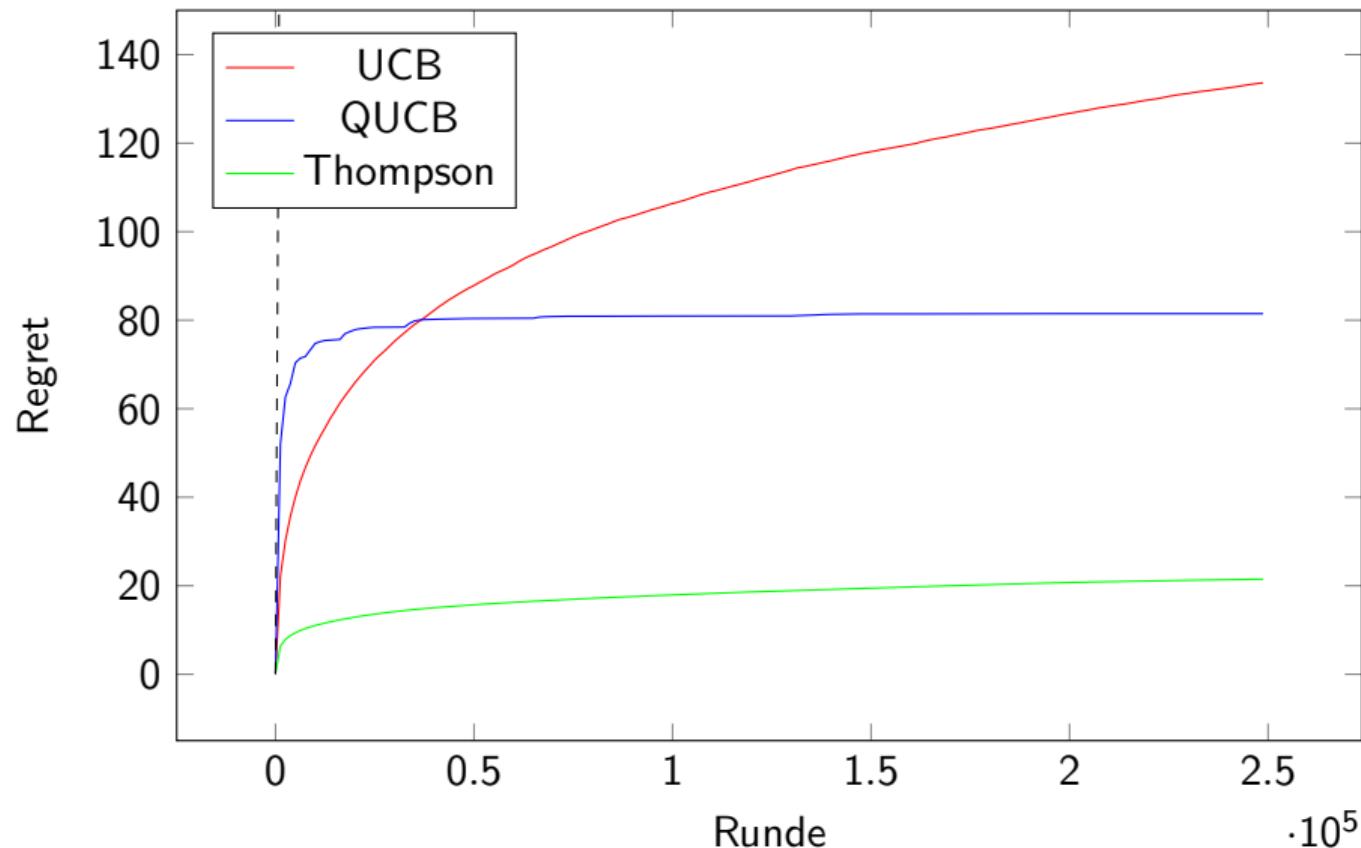
Kvantealgoritmer

Kvantealgoritme	Problem	Kompleksitet	Beste klassiske
QFT	Fouriertransformasjon	$O(\log(N))$	$O(N \log(N))$
Shor	Faktorisering	$O(\log(N)^3)$	Subeksponensiell
HHL	Lineære ligningssystemer	$O(\log(N)\kappa^2)$	$O(N\kappa)$
Grover	Ustrukturert søk	$O(\sqrt{N})$	$O(N)$
QMC	Estimere $\mathbb{E}[X]$	$O(1/\epsilon)$	$O(1/\epsilon^2)$

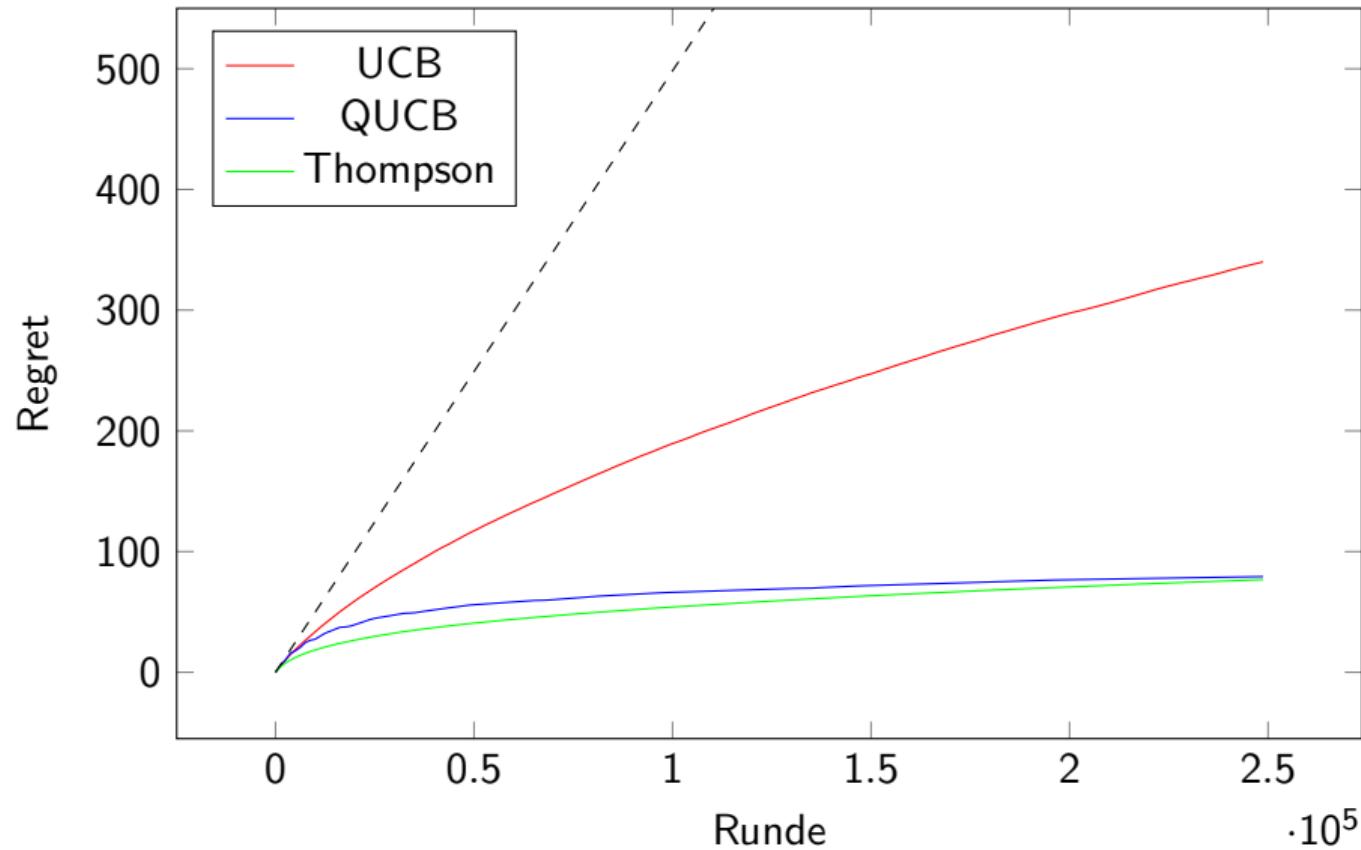
- Ide: bruk QMC til å estimere $\mathbb{E}[X_i]$.
- Fordel: konfidensintervaller blir mindre
- Resultat: $O(\log(T))$ -algoritme



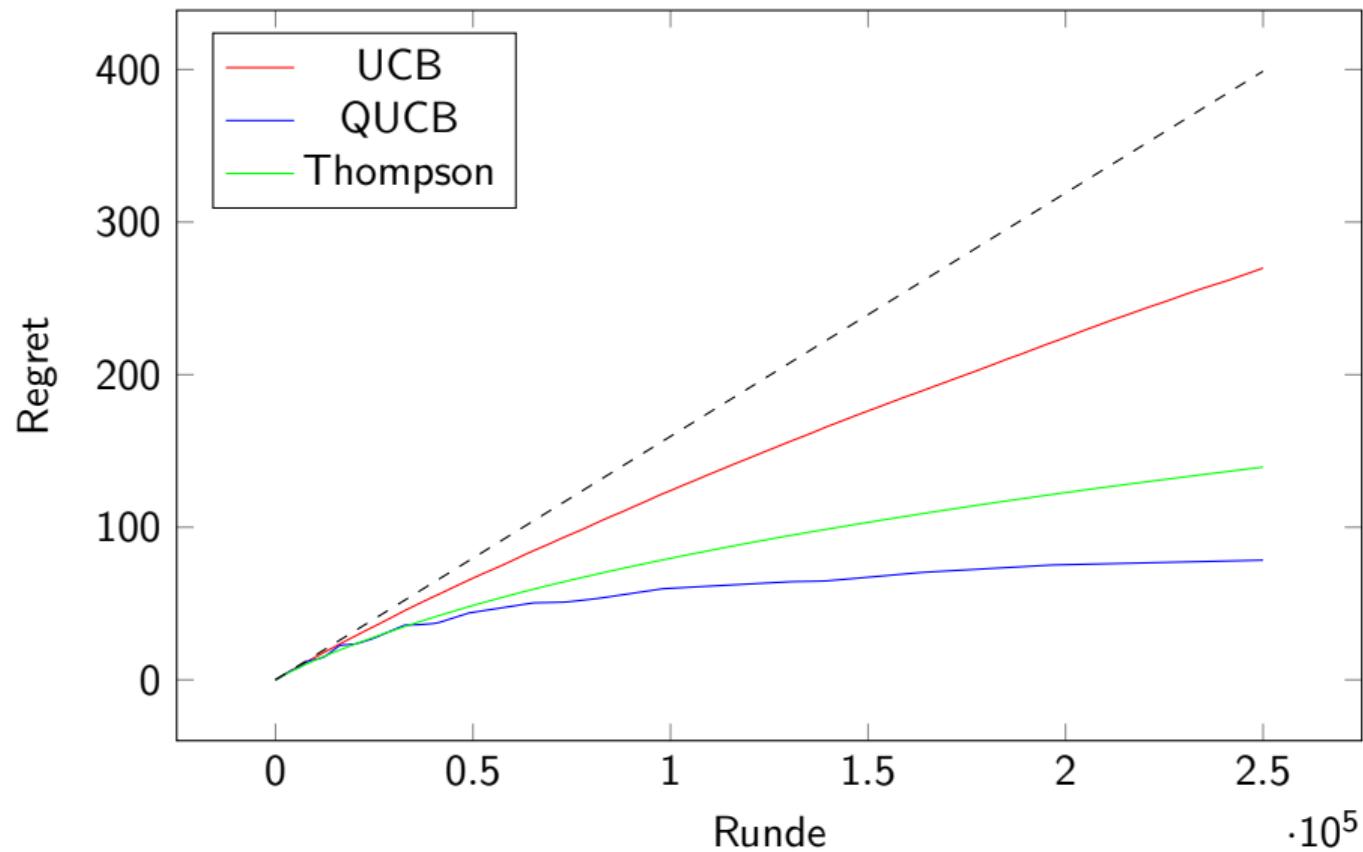
Simuleringer: uniform prior



Simuleringer: korrelert prior



Simuleringer: meget korrelert prior

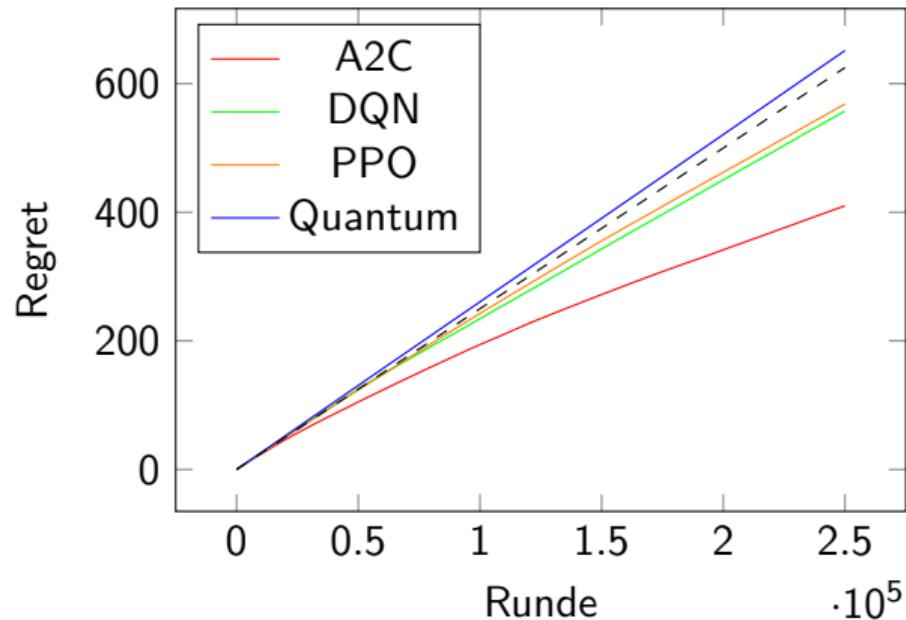


Konklusjoner

- Bedre i *noen* tilfeller
- Juks med kvante
- Vanskelig å anvende
- QUCB kan videreutvikles
- Kvantealgoritmer for andre bandittvarianter?

Vedlegg: maskinlæring og banditter

- Forsterket læring
- Hvordan uttrykke bandittproblemet som et RL-problem?
- Fungerer kanskje, men krever i så fall anpassing



Vedlegg: kvantovermakt

