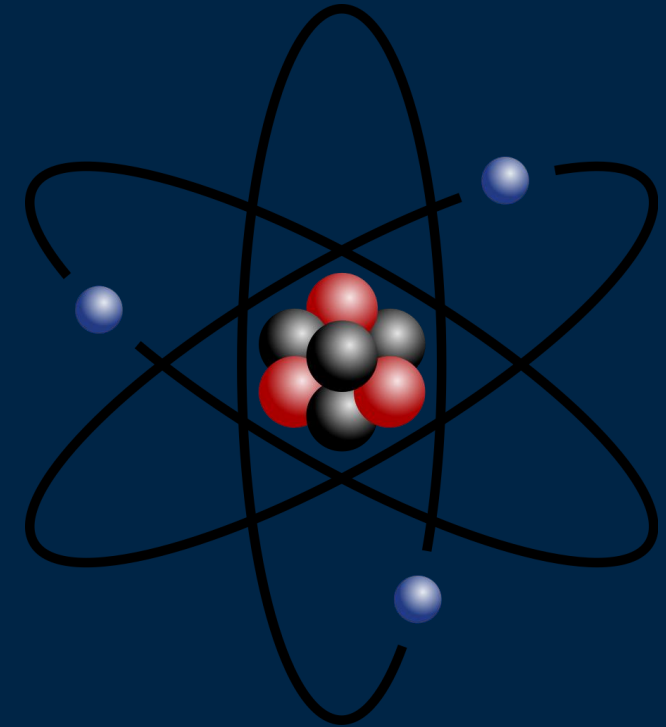


ALPHAZERO: FRA SKAK TIL KVANTESYSTEMER



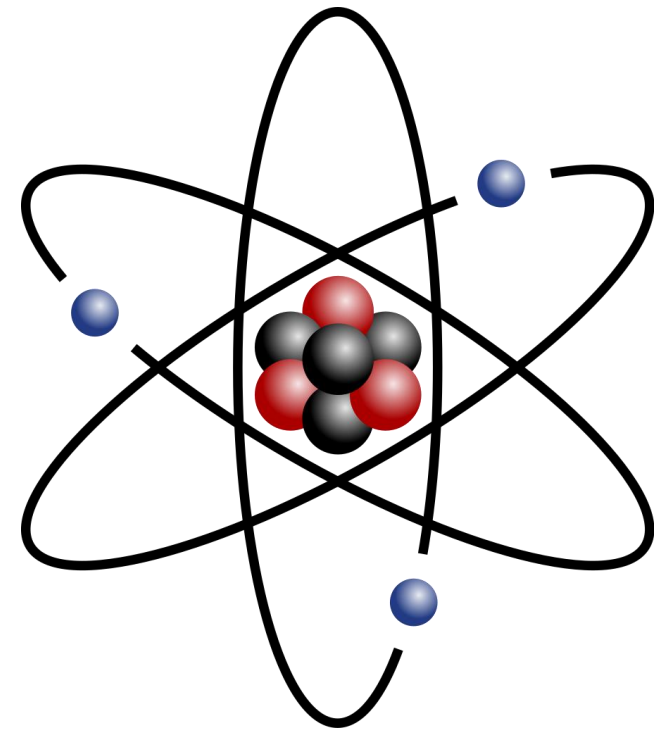
HVEM ER JEG?

- Mogens Dalgaard (dalgaard@phys.au.dk)
- PhD studerende i fysik ved Aarhus Universitet
- Arbejder med kontrol af kvantesystemer
- Undersøger specifikt Machine Learning metoder
- Vejleder: Prof. Jacob Sherson



KVANTEMEKANIK

- Beskriver nogle af de mindste byggesten i vores univers.
- Opstod tilbage i begyndelsen af 1900-tallet.
- Bygger på sandsynlighed.
- Tillader et system at være i en kombination af flere tilstande på én gang.



Klassisk:



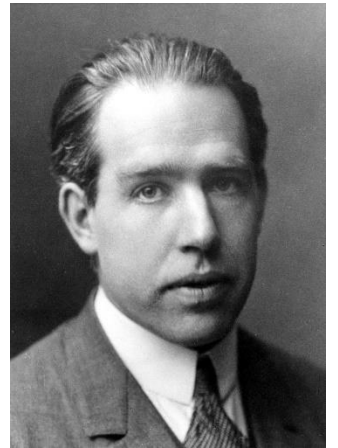
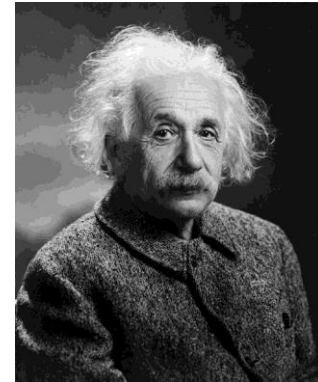
eller



Kvantemekanisk:



og



KVANTECOMPUTER

En klassisk computer bruger bits: dvs. 0 eller 1

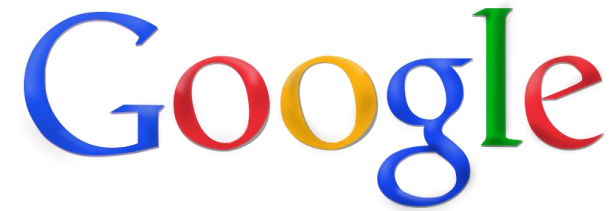
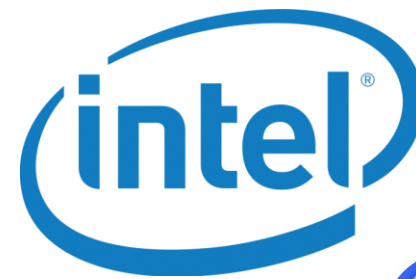
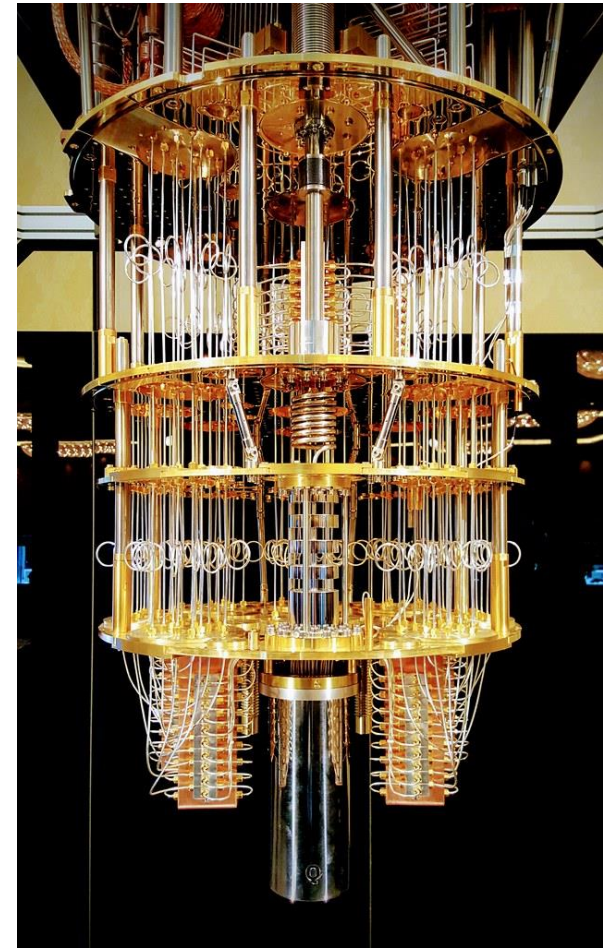
For N bits har man 2^N mulig kombinationer

Disse bliver manipuleret gennem såkaldte gate operationer.

I en kvantecomputer kan man indkode de 2^N mulige kombinationer i én tilstand. Via gate operationer kan man så lave 2^N beregninger på én gang.

Klassisk: (1 0 1 ... 0) én ad gangen

Kvant.: $\begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & 0 \\ 1 & \dots & 1 & 1 \end{pmatrix} 2^N$ ad gangen



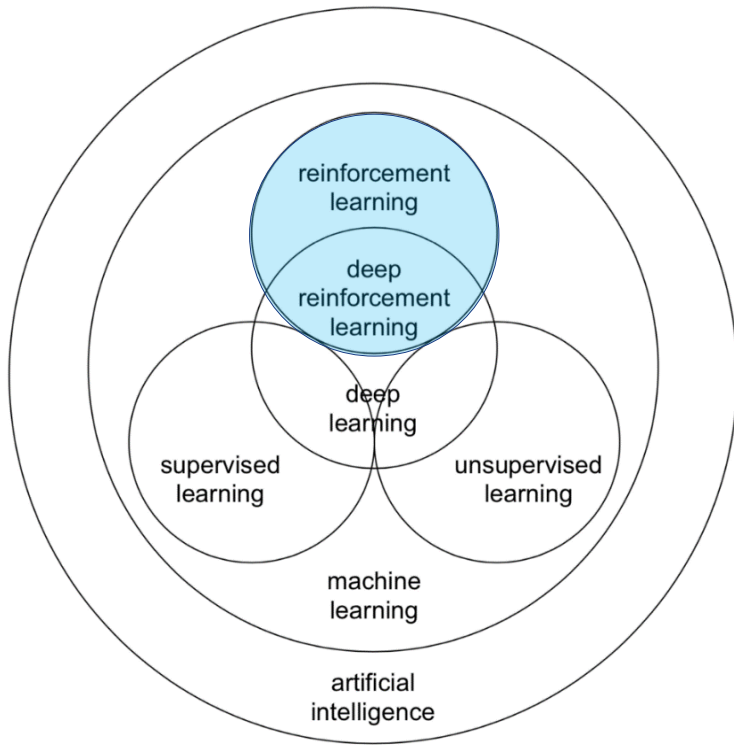
GOOGLE DEEPMIND'S ALPHAZERO

- AlphaGo Zero udkom i 2017
- Go er et kinesisk brætspil, hvor det gælder om at omringe modstanderens sten.
- 10^{170} mulige spil konfigurationer.
- 10^{43} i skak og 10^{80} atomer i universet.
- AlphaGo Zero slog både professionelle spillere og specialiseret spille-software.
- AlphaGo Zero lærte at spille "from scratch".
- I 2018 udkom efterfølgeren AlphaZero der også kunne spille skak og shogi.

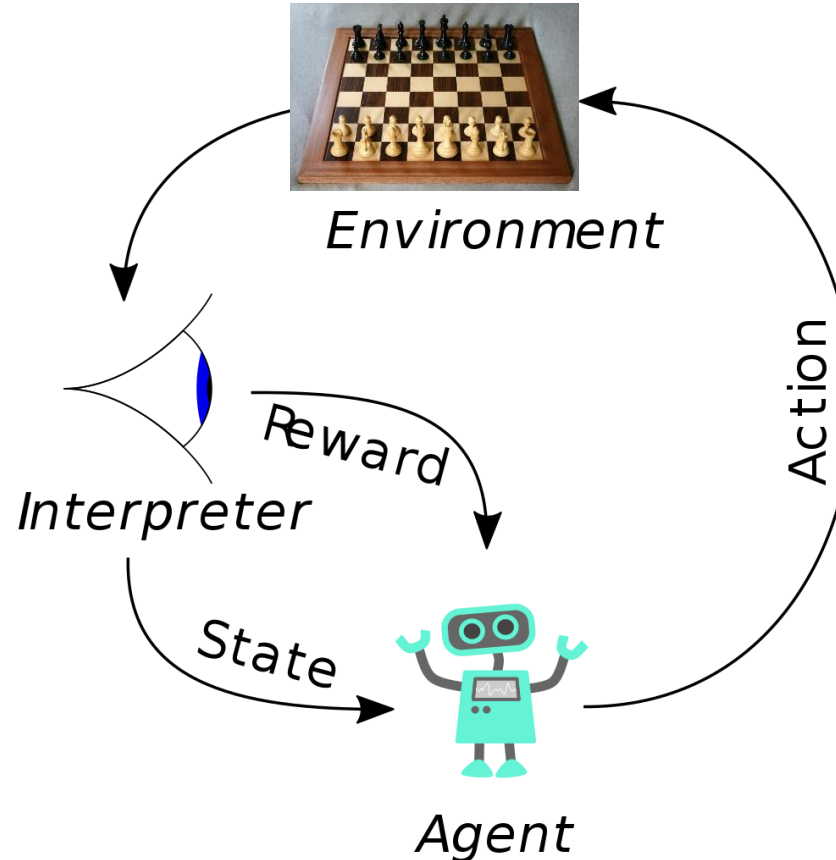


REINFORCEMENT LEARNING

I Reinforcement Learning skal en software agent lære at interagere med et miljø. Målet er at maksimere et feedback signal kaldet en reward (reelt tal).



Yuxi Li, Deep Reinforcement Learning, arXiv, 2018



VORES OPGAVE:

Opbyg gradvist en puls-sekvens, der realiserer en gate operation mellem to kvante-bits

I sidste ende regner vi en belønning F som ligger mellem 0 og 1.

Hvor $F = 1$ betyder 100% succes

Trin 1:

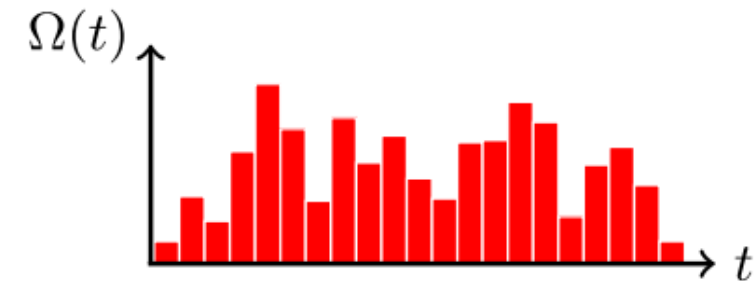


Trin 2:

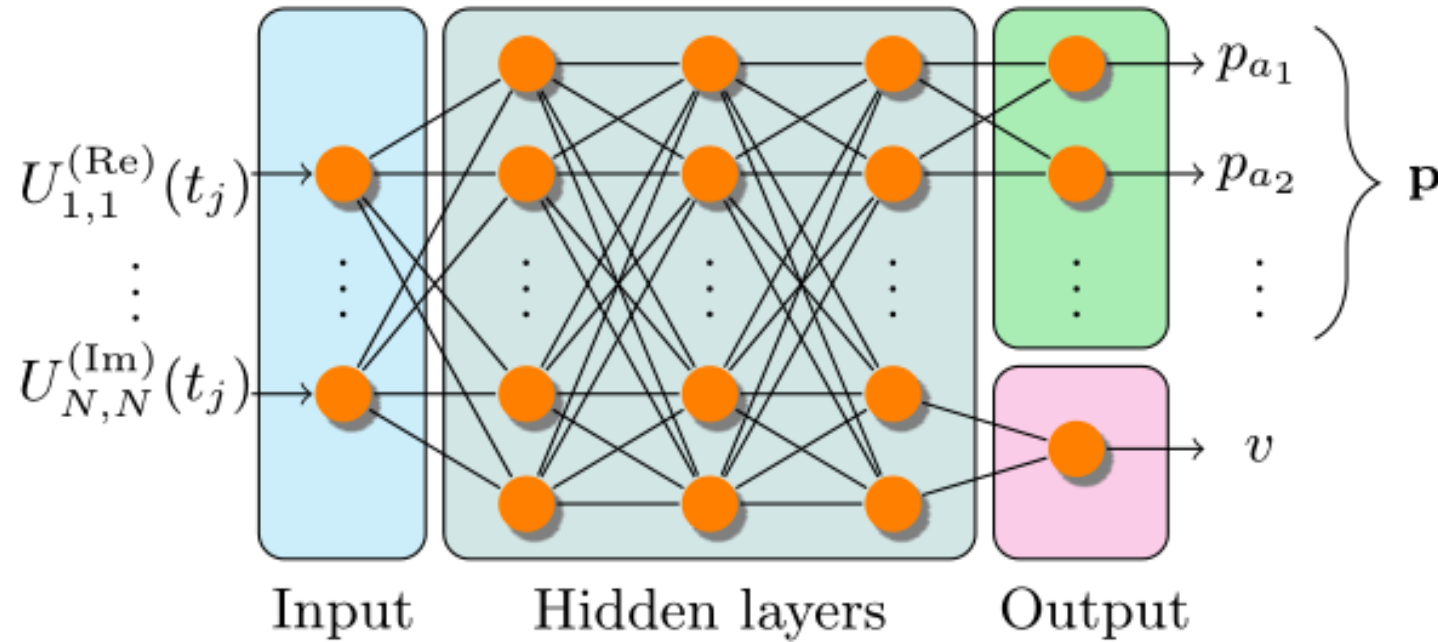


⋮

Trin N:



ALPHAZERO: DYBT NEURALT NETVÆRK



Vores kvantesystem er matematisk beskrevet ved en række tal $U_{1,1}, \dots, U_{NN}$

$p_1, p_2, \dots$ er “sandsynligheder” for en given action.

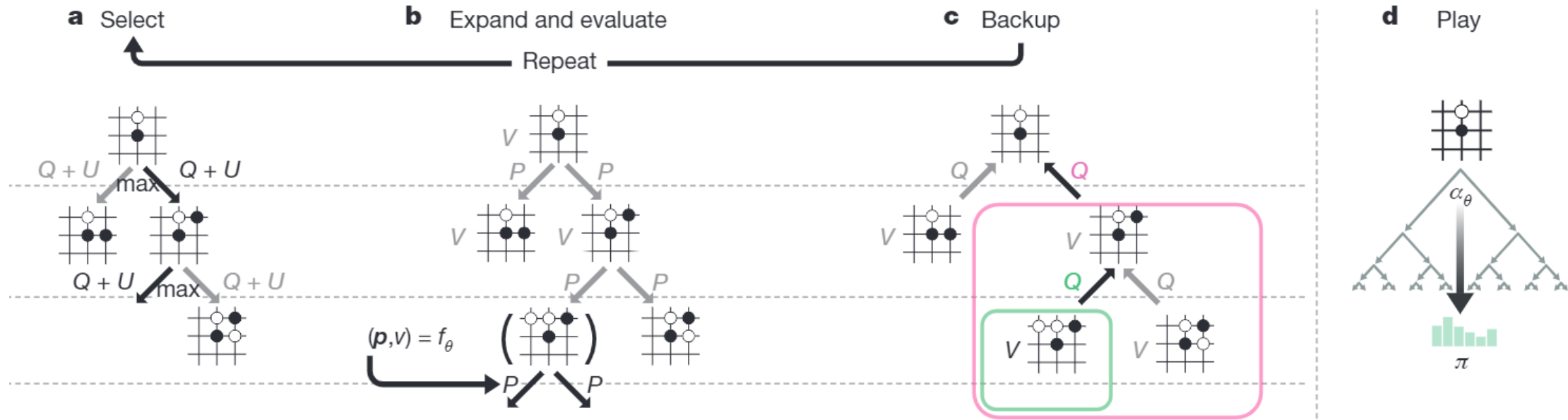
v er estimate af den endelig score.

ALPHAZERO- TRÆSØGNING

Q : Gennemsnitlig estimeret belønning (reward)

U : En usikkerhed der sikrer mere udforskning

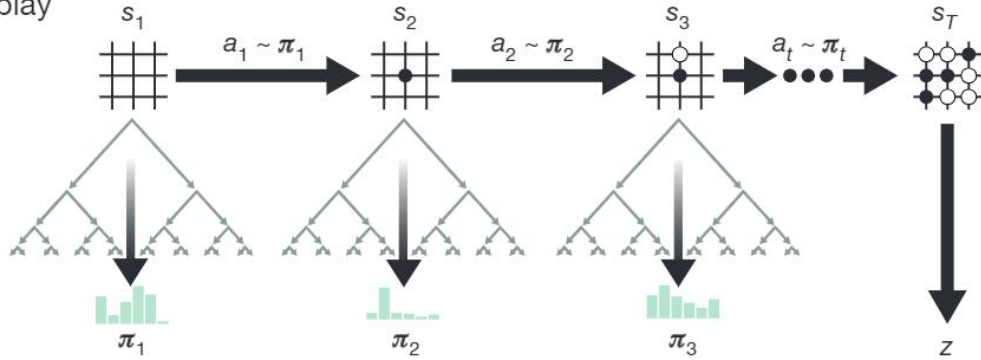
π : en politik, dvs. sandsynlighed for at tage en given handling (action).



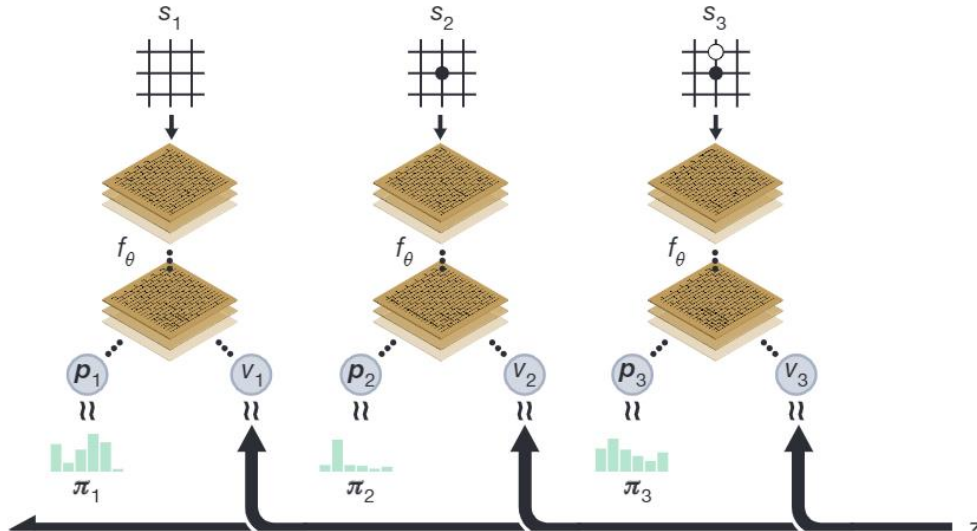
D. Silver et al., Nature 550, 354–359 (2017).

ALPHAZERO: SPIL OG TRÆNING

a Self-play



b Neural network training



For hver skridt/træk regner AlphaZero en politik π (vha. træsoegningen), som den trækker en handling fra.

Den gemmer så politikken π samt den endelige score z .

Det neurale netværk trænes så, således at (p, v) nærmer sig (π, z) .

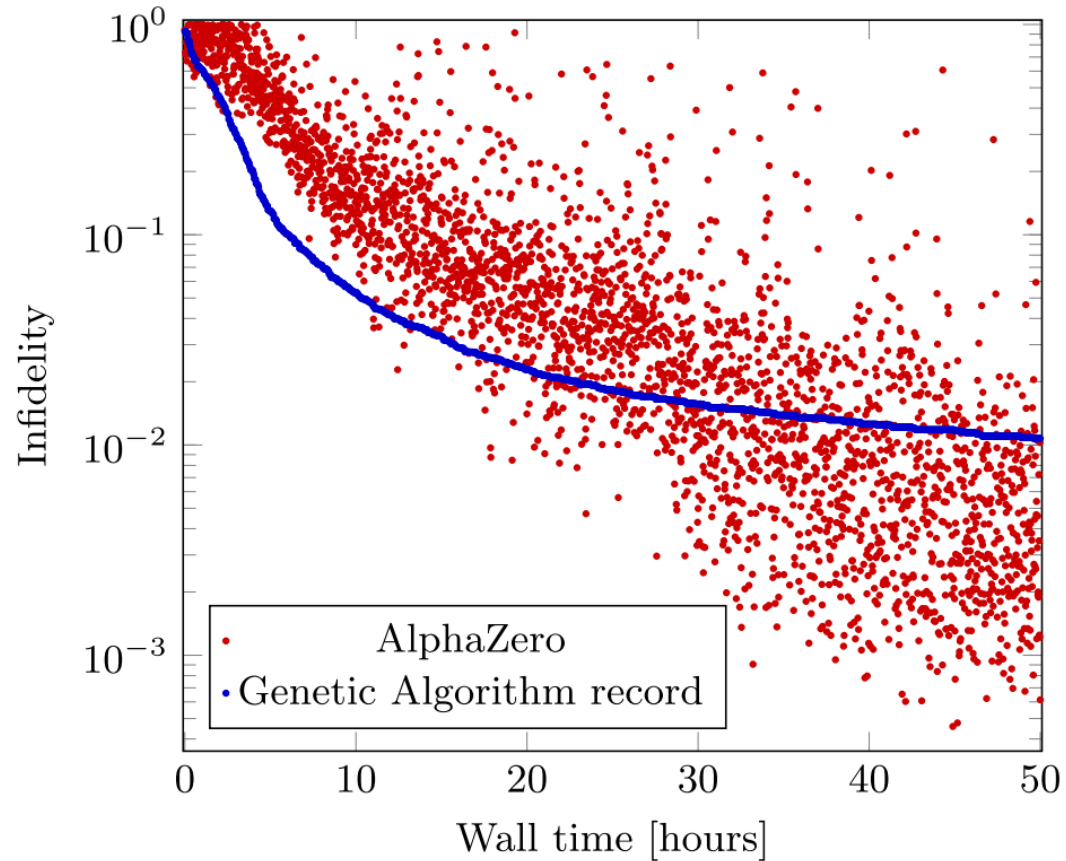
D. Silver et al., Nature 550, 354–359 (2017).

I TILFÆLDE AF AT I FALDT I SØVN:

AlphaZero bruger en træ søgning til at se flere skridt frem af!

VORES RESULTATER

Infidelity: En fejl-rate (lavere er bedre).



nature > npj quantum information > articles > article

MENU ▾ npj | Quantum Information

Article | [Open Access](#) | Published: 14 January 2020

Global optimization of quantum dynamics with AlphaZero deep exploration

Mogens Dalgaard, Felix Motzoi, Jens Jakob Sørensen & Jacob Sherson [✉](#)

npj Quantum Information **6**, Article number: 6 (2020) | [Cite this article](#)

5959 Accesses | **54** Altmetric | [Metrics](#)

VORES RESULTATER

Vi testede AlphaZero på tre forskellige problemer inden for kvantekontrol.

På to af problemerne fik vi en klar forbedring med AlphaZero relativt til konventionelle metoder.

Men på det sidste problem, blev AlphaZero slået af en kvanteoptimeringsalgoritme.

Vi lavede da en Hybrid-algoritme der kombinerede AlphaZero med denne kvanteoptimeringsalgoritme. Her blev AlphaZero løsninger efterfølgende optimeret.

Dette gav en kæmpe forbedring i både kvantitet og kvalitet af løsningerne.

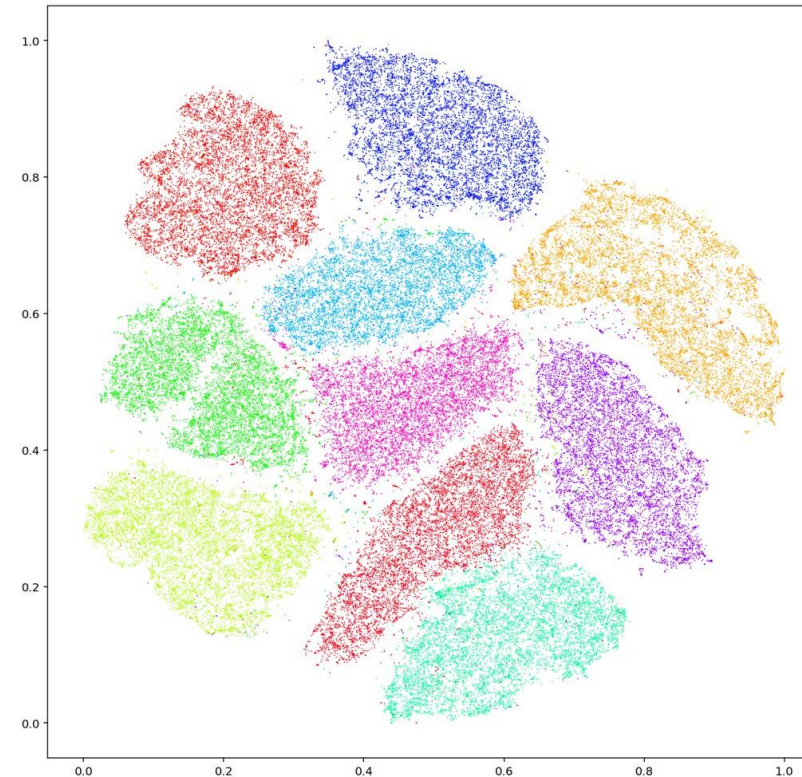
VORES RESULTATER

Vi analyserede efterfølgende vores resultater med en t-SNE algoritme.

En t-SNE algoritme laver en afbildning ned i et to-dimensionelt plan, således at hvis 2 inputs minder om hinanden vil de være repræsenteret af 2 punkter, der ligger tæt.

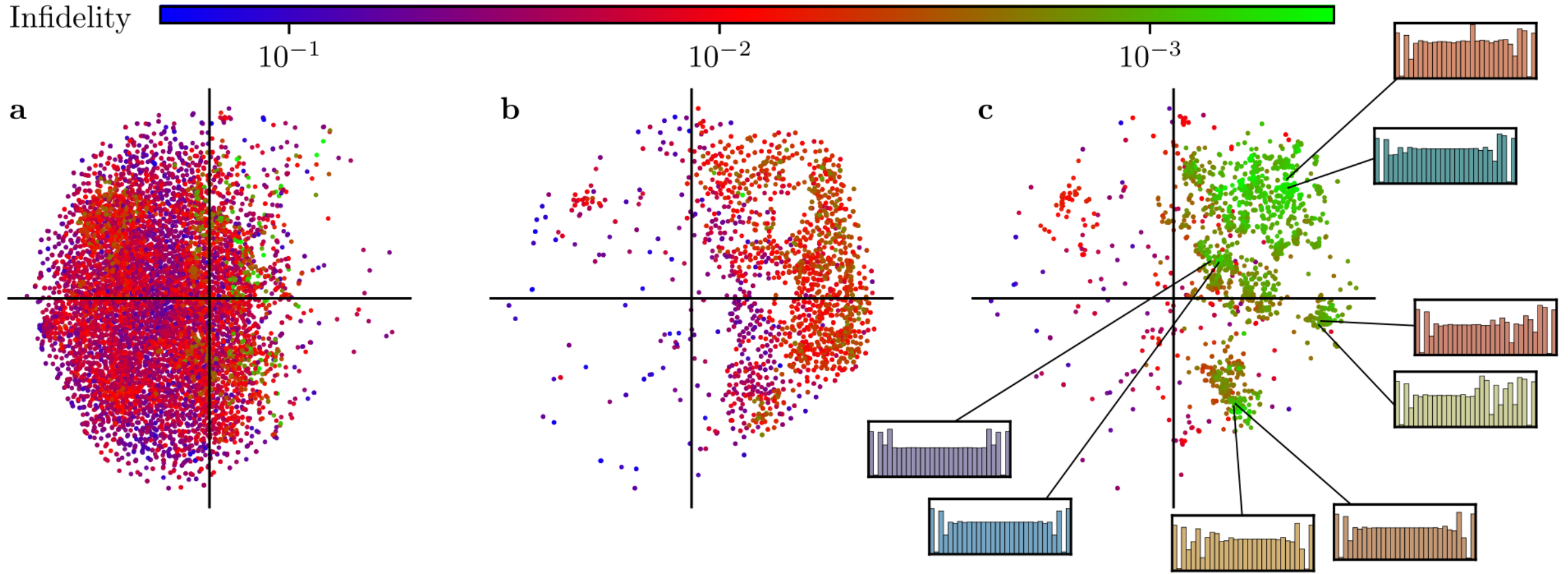
Eksempel:

Genkend håndskrevne tal fra 0 til 9



HYBRID ALGORITMEN

t-SNE: a) Kvanteoptimeringsalgoritme, b) AlphaZero , c) Hybrid



KONKLUSION

- AlphaZero er en rigtig god algoritme, der kan lære at løse forskellige problemer (fra at spille skak til at styre et kvantesystem).
 - Hybrid-løsninger hvor man kombinerer machine learning-metoder med metoder uden for feltet, kan være bedre en ren machine learning.
-
- Tip til jer: Hold øje med kvantecomputere, det kan være fremtiden!

VIDERE ARBEJDE

- Siden vores publikation er vi blevet kontaktet af flere forskningsgrupper.
- Derved håber vi at vores teoretiske arbejde snart kan anvendes eksperimentelt.
- I vores videre arbejde vil vi desuden fokuserer på mere realistiske modeller.

Tak fordi at i lyttede...