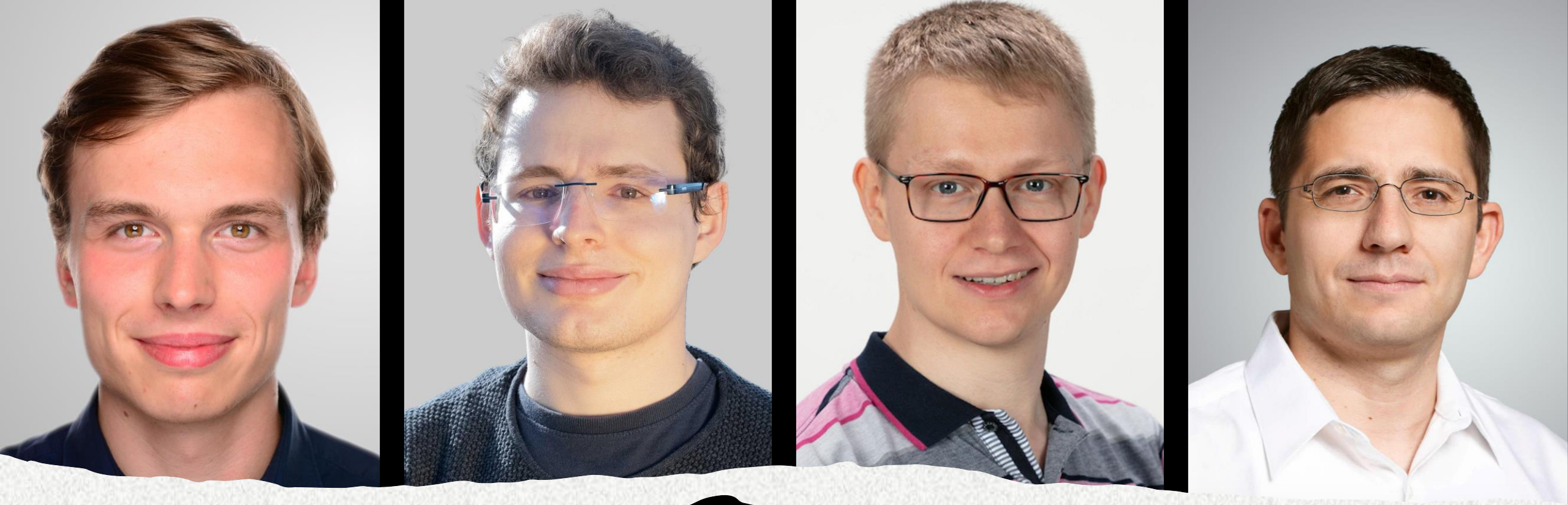


Abstract Interpretation of Fixpoint Iterators with Applications to Neural Networks

- **Department of Computer Science, ETH Zurich, Switzerland**
- Mark Niklas Muller, Marc Fischer, Robin Staab, Martin Vechev

Outline

- **Background of Deep Learning**
- **DEQs**
- **Mono DEQs**
- **Combine Abstraction interpolation with DEQs**
- **Result**

Outline

Background of Deep Learning

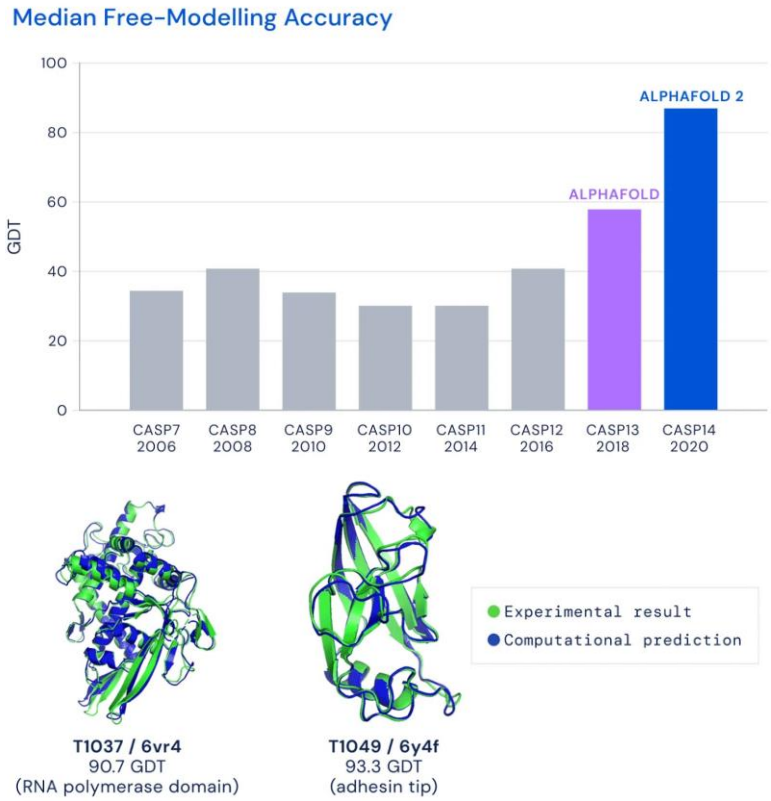
DEQs

Mono DEQs

Combine Abstraction interpolation with DEQs

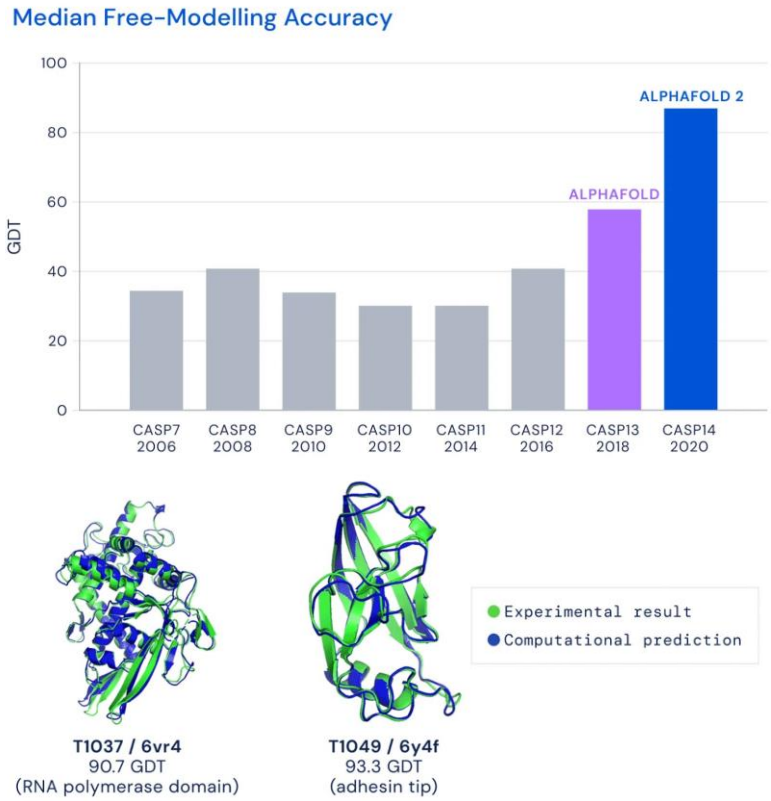
Result

The deep learning revolution (recent examples)



AlphaFold: Jumper et al., 2021

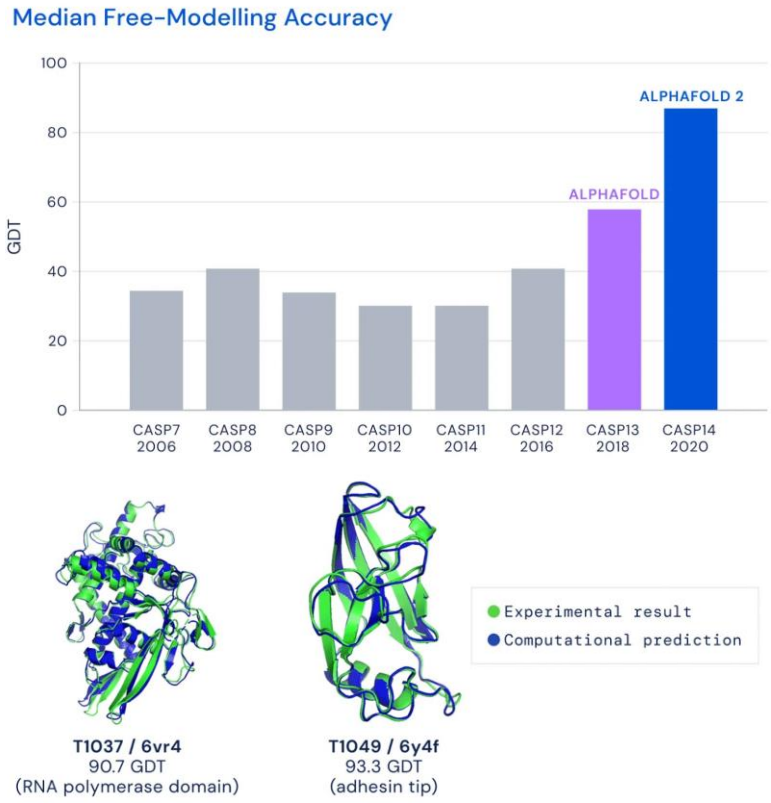
The deep learning revolution (recent examples)



PaLM: Chowdhery et al., 2022

AlphaFold: Jumper et al., 2021

The deep learning revolution (recent examples)



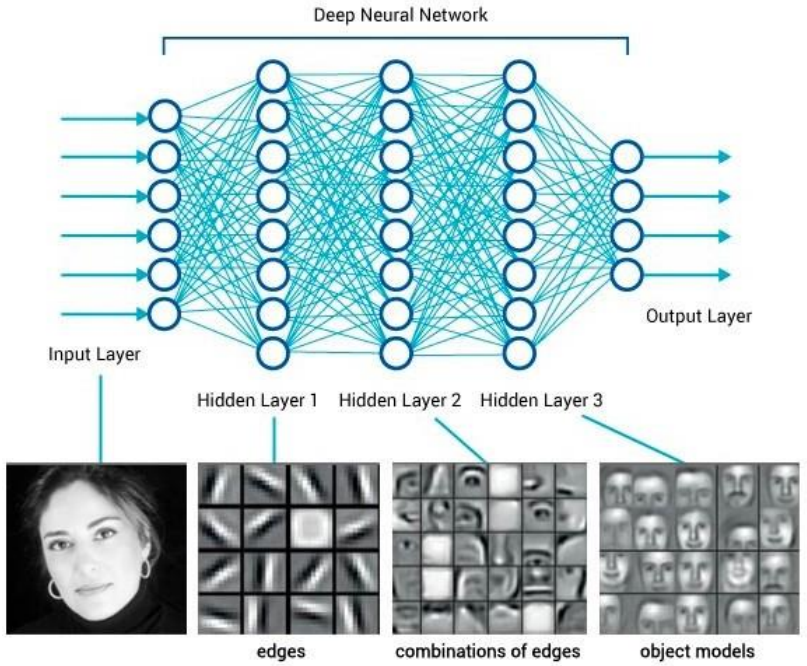
AlphaFold: Jumper et al., 2021

PaLM: Chowdhery et al., 2022

DALL-E 2: Ramesh et al., 2022

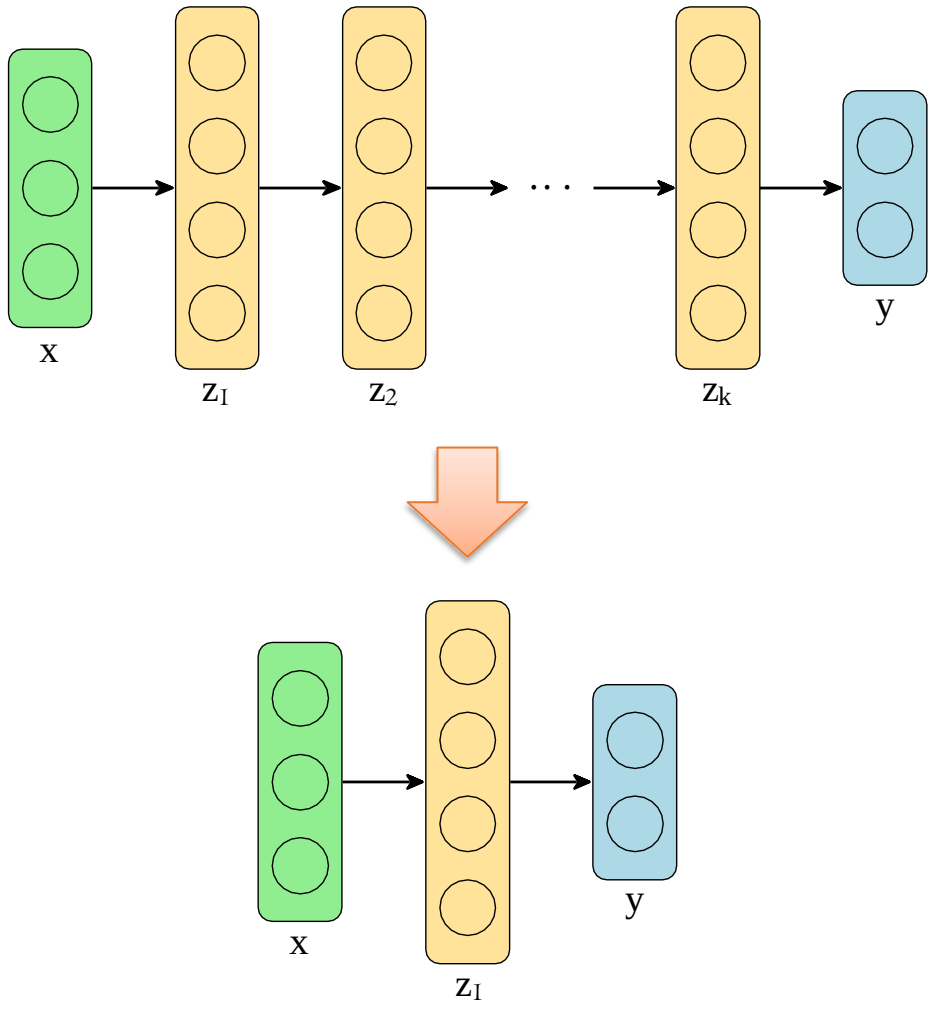
Deep Learning

The story we all tell: deep learning algorithms build hierarchical models of input data, where the earlier layers create “simple” features and hidden layers create hig h-level abstractions of the data.



Can we lightweight the model?

Deep Learning



DEQs can replace traditional depth in dee p networks with a single (implicit) layer.

- Simpler architectural design.
- Vastly reduced memory requirements.
- Matches or exceeds accuracy of co mparable fixed-depth networks.

Offers a new perspective on what deep networks are “really” computing.

Outline

Background of Deep Learning

DEQs

Mono DEQs

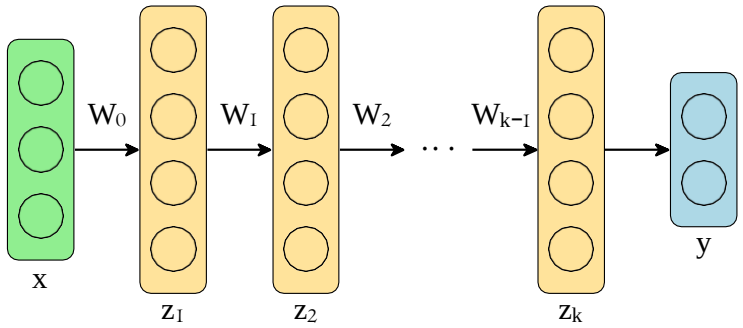
Combine Abstraction interpolation with DEQs

Result

From deep network to DEQs

Traditional network

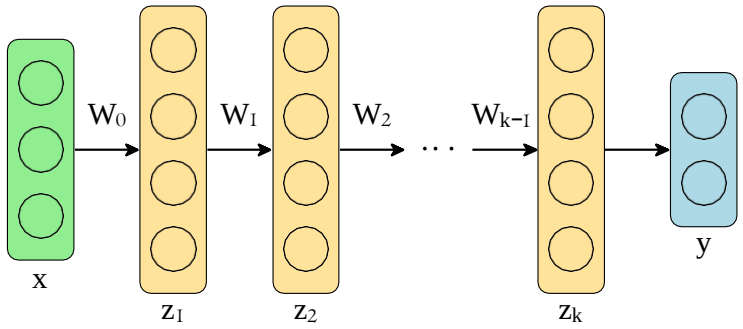
$$x_{i+1} = \sigma(W_i x_i + b_i)$$



From deep network to DEQs

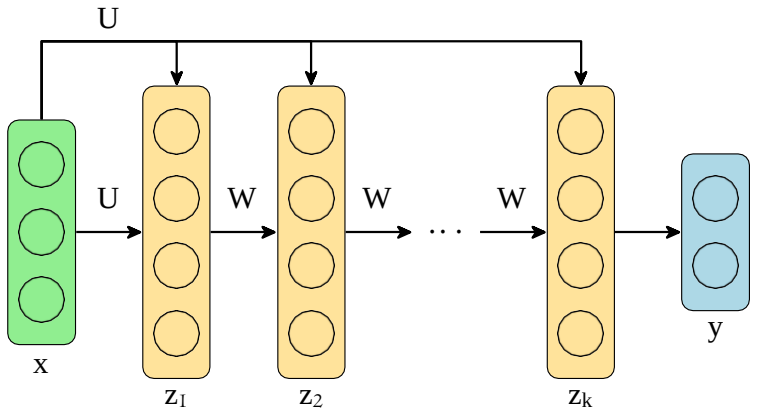
Traditional network

$$x_{i+1} = \sigma(W_i x_i + b_i)$$



Weight-tied, input-injected network

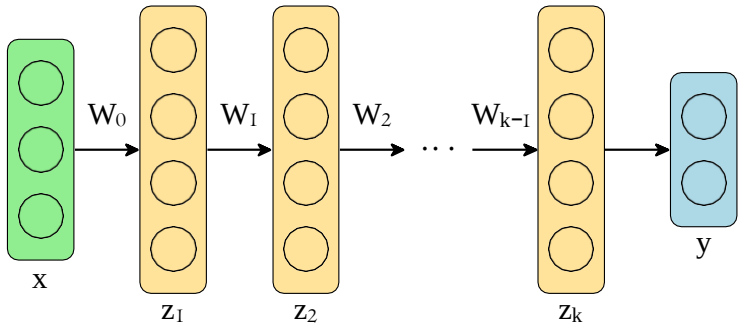
$$x_{i+1} = \sigma(W x_i + U x + b)$$



From deep network to DEQs

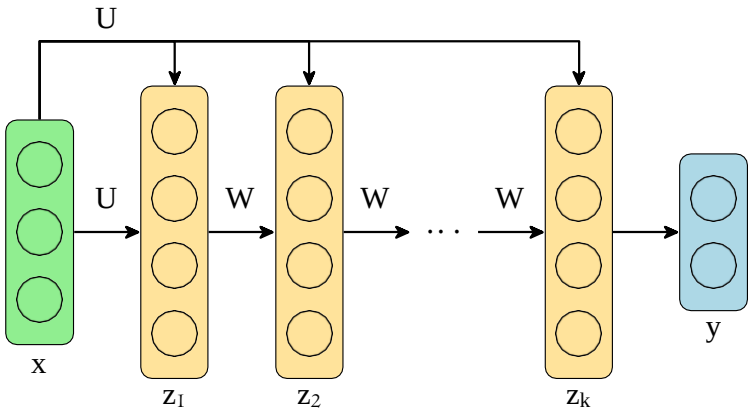
Traditional network

$$x_{i+1} = \sigma(W_i x_i + b_i)$$



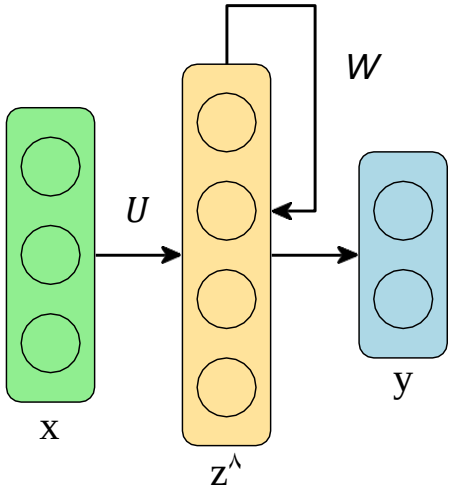
Weight-tied, input-injected network

$$x_{i+1} = \sigma(W x_i + U x + b)$$



Deep Equilibrium (DEQ) model

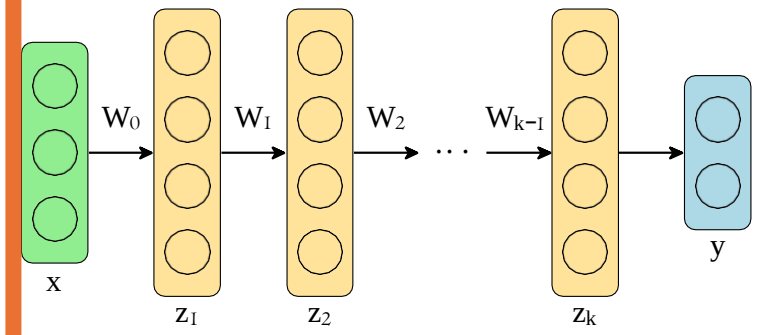
$$x^\star = \sigma(W x^\star + U x + b)$$



From deep network to DEQs

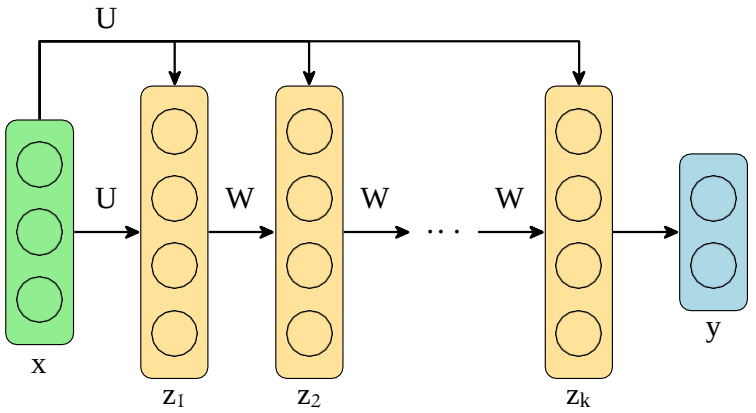
Traditional network

$$x_{i+1} = \sigma(W_i x_i + b_i)$$



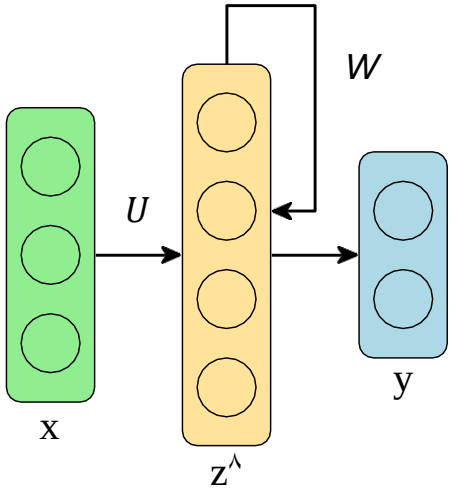
Weight-tied, input-injected network

$$x_{i+1} = \sigma(W x_i + U x + b)$$



Deep Equilibrium (DEQ) model

$$x^\star = \sigma(W x^\star + U x + b)$$

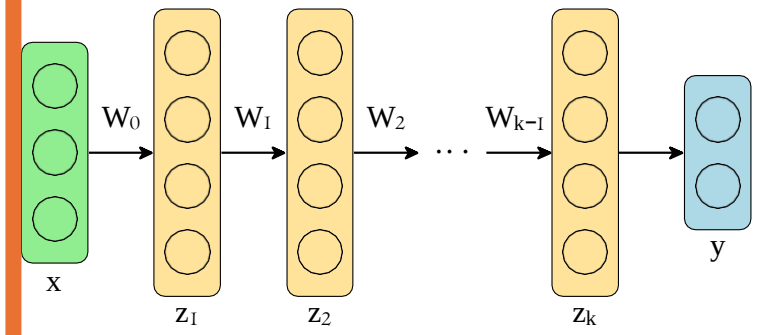


Threshold based

From deep network to DEQs

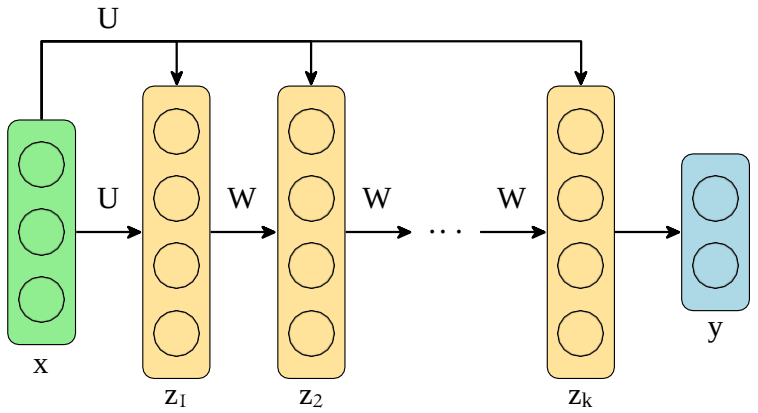
Traditional network

$$x_{i+1} = \sigma(W_i x_i + b_i)$$



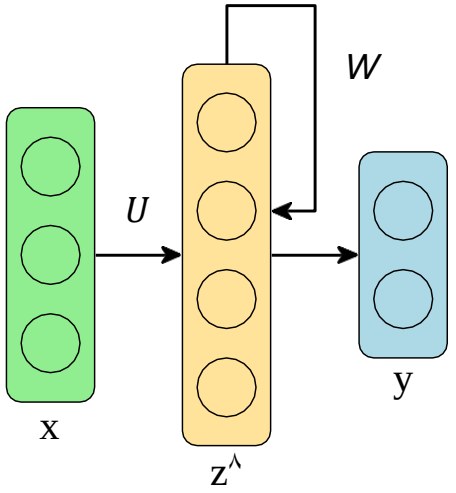
Weight-tied, input-injected network

$$x_{i+1} = \sigma(W x_i + U x + b)$$



Deep Equilibrium (DEQ) model

$$x^\star = \sigma(W x^\star + U x + b)$$



Threshold based

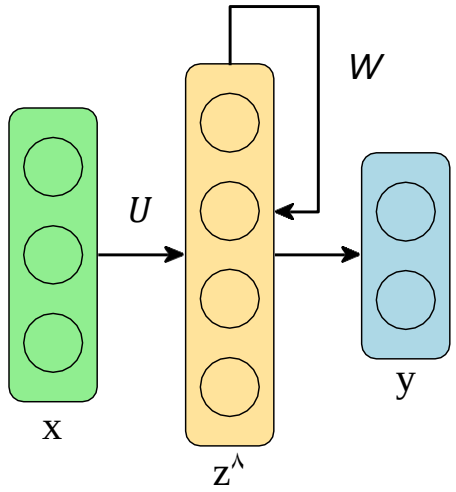
Fixed point based

Fixed point in DEQ

- Fixed poit in AI
  The End of learning
  Consistence input and output

Learning and adjust the weight until
F(x) = y always consistence.

Deep Equilibrium (DEQ) model



Outline

Background of Deep Learning

DEQs

Mono DEQs

Combine Abstraction interpolation with DEQs

Result

Mono DEQs

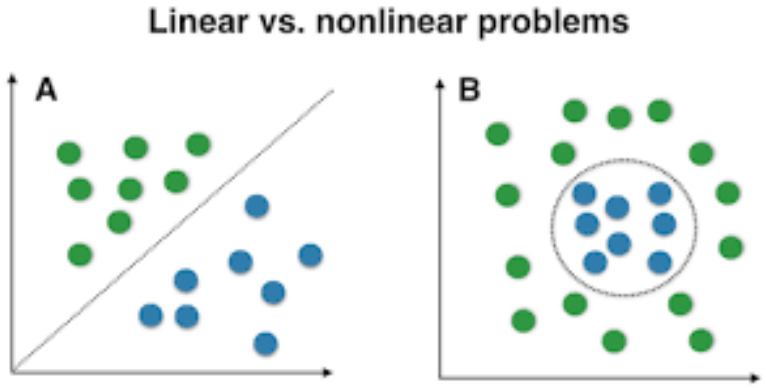
- What can we say about the fixed point?

Mono DEQs

- What can we say about the fixed point?
- Finding linear model’s fixed point is easy.

Mono DEQs

- What can we say about the fixed point?
- Finding linear model's fixed point is easy.
- In general, a non-linear dynamical system can be difficult to establish existence, uniqueness and stability of the fixed point iteration.



Mono DEQs

- What can we say about the fixed point?
- Finding linear model's fixed point is easy.
- In general, a non-linear dynamical system can be difficult to establish existence, uniqueness and stability of the fixed point iteration.

Finding fixed point in non-linear is hard.

Mono DEQs

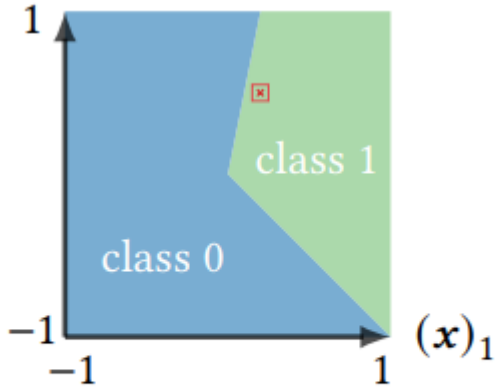
- Finding mono DEQs's fixed point is hard

Mono DEQs

- Finding mono DEQs's fixed point is hard
- Prior Research (Kleene) introduced abstract interpretation to get fixed point

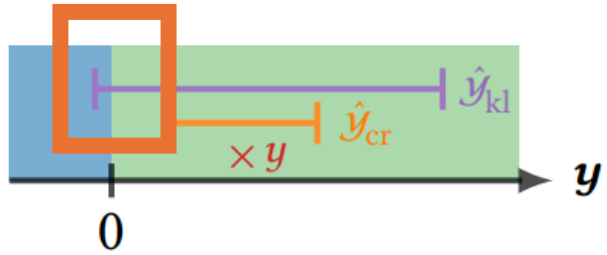
Mono DEQs

- Finding mono DEQs's fixed point is hard
- Prior Research (Kleene) introduced abstract interpretation to get fixed point
  - slow & inaccurate



(a) Classification of concrete inputs in $[-1, 1]^2$ by h.

Inaccurate



(c) Abstractions of the output y and resulting classification.

Mono DEQs

- Finding mono DEQs's fixed point is hard
- Prior Research (Kleene) introduced abstract interpretation to get fixed point
  - slow & inaccurate
- Domain are not suitable for DEQs

	Iteration	Inclusion	Precision
Box	✓	✓	✗
(Hybrid) Zonotope	✗	✗	✓
Polyhedra	✗	✗	?
CH-Zonotope	✓	✓	✓

Outline

Background of Deep Learning

DEQs

Mono DEQs

Combine abstraction interpretation with DEQs

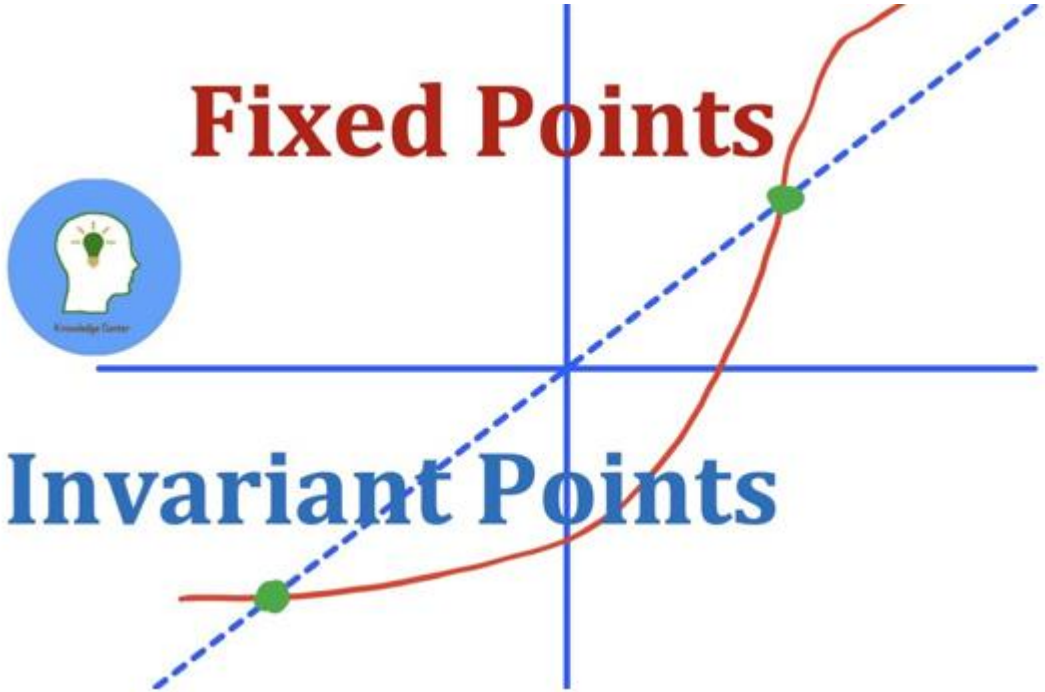
Result

Abstract interpretation in Mono DEQs

They want to evaluate and know the fixed point

Apply Abstract interpretation
➔ Static analysis technique

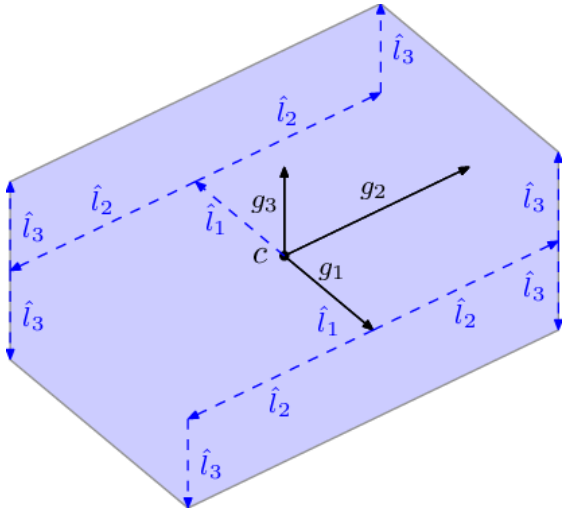
➔ The fixed point was predicted using abstract interpretation.



Abstract interpretation in Mono DEQs

- They apply the zonotope for fixed point.
  - But the neural network is complex.

- Introduce **CH-zonotope.**
  - Apply new domain based on zonotope
  - When dealing with complex problems, it is advantageous.
  - efficient propagation of abstract elements
  - fast (abstract) inclusion checks



Outline

Background of Deep Learning

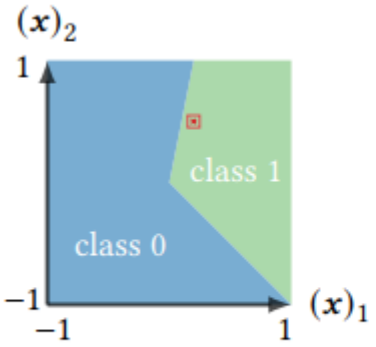
DEQs

Mono DEQs

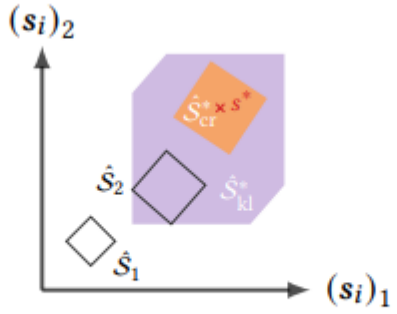
Combine Abstraction interpretation with DEQs

Result

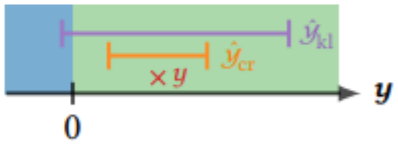
Result



(a) Classification of concrete inputs in $[-1, 1]^2$ by h.



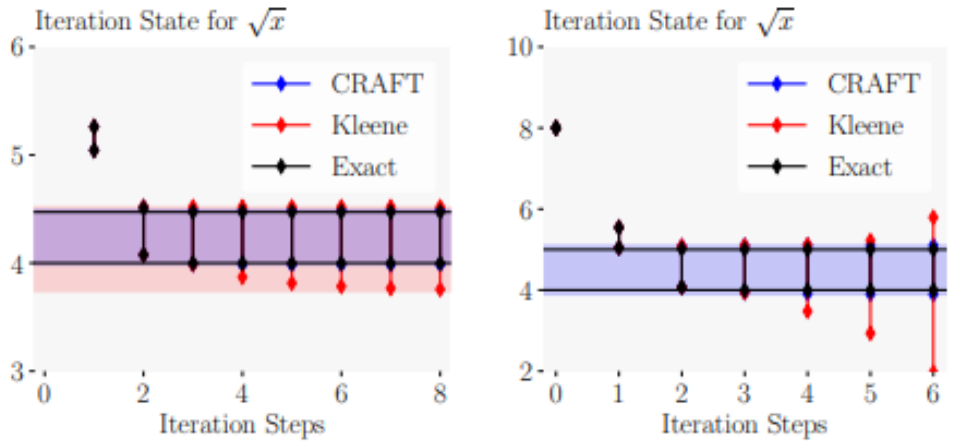
(b) Abstractions of iteration steps s_i.



(c) Abstractions of the output y and resulting classification.

The tool they developed, named Craft,
exhibits better performance than the existing Kleene iteration.

Result



CRAFT provides more accurate and stable convergence results compared to Kleene iteration

Conclusion

- Interpretation the range of correct fixed points with CH-zonotope
  - Convergence Detection
  - Early Stopping Criterion
  - Prevention of Overfitting
  - Resource Efficiency

Thank you