



Associate Professor in Mathematics

digital futures

Chair of the Working Group "Learn"

WASP | WALLENBERG AI,
AUTONOMOUS SYSTEMS
AND SOFTWARE PROGRAM

Leader of the WASP Cluster on
Mathematical Foundations of AI

AI Foundations

Kathlen Kohn



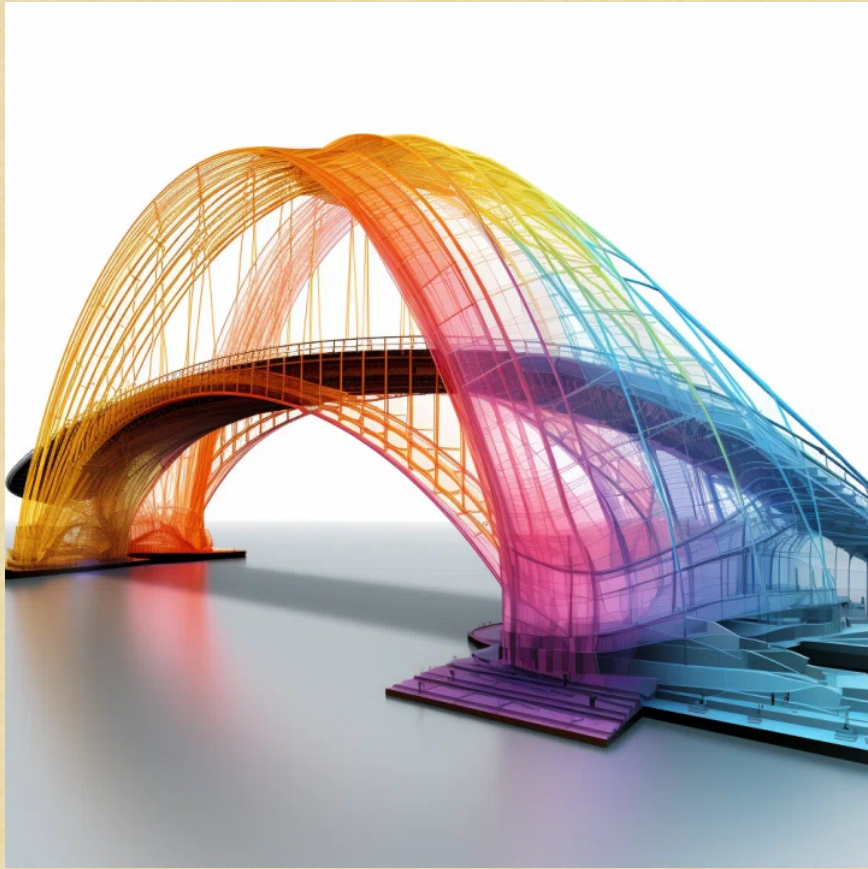
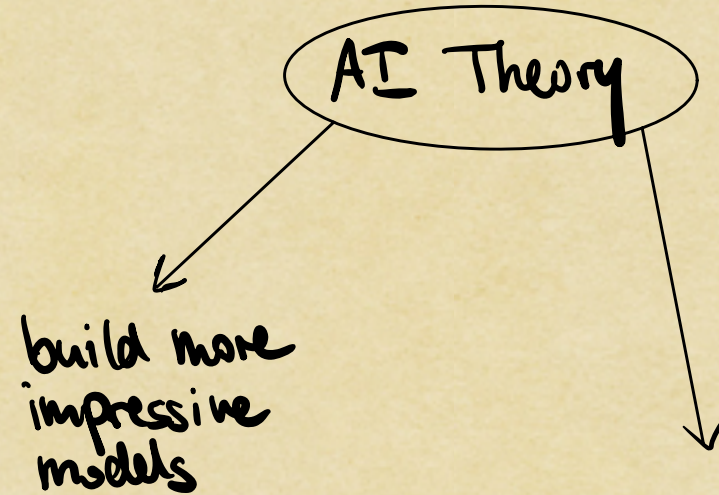
ELLIS Scholar
in Geometric Deep Learning
& in the ELLIS Unit Sweden



SIAM Program Director for
Algebraic Geometry

Role of AI Foundations

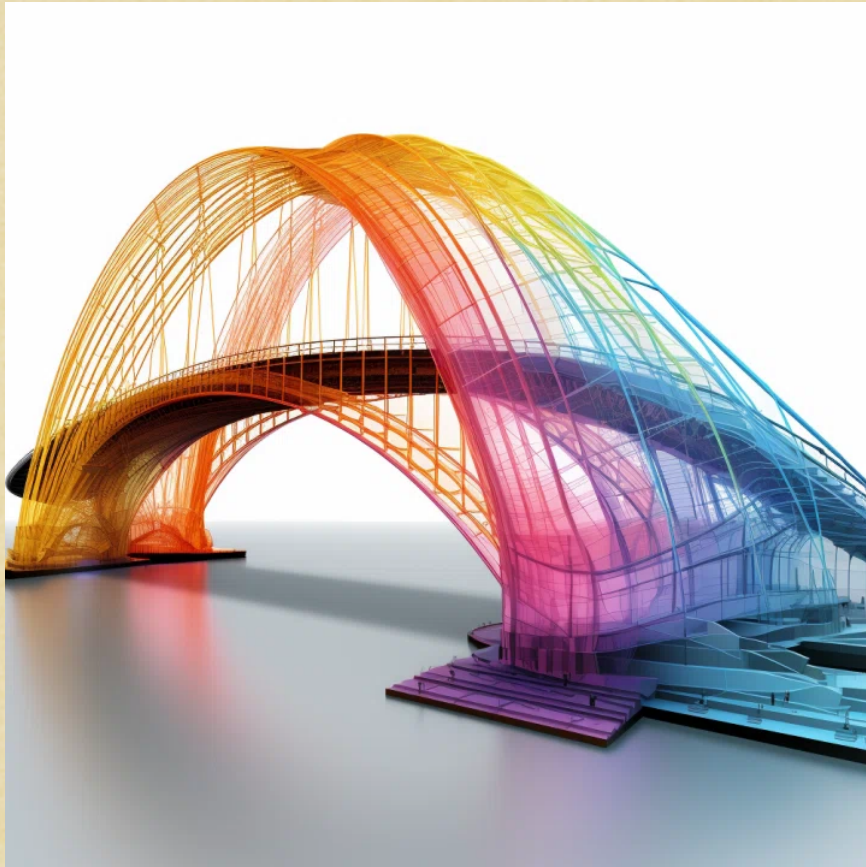
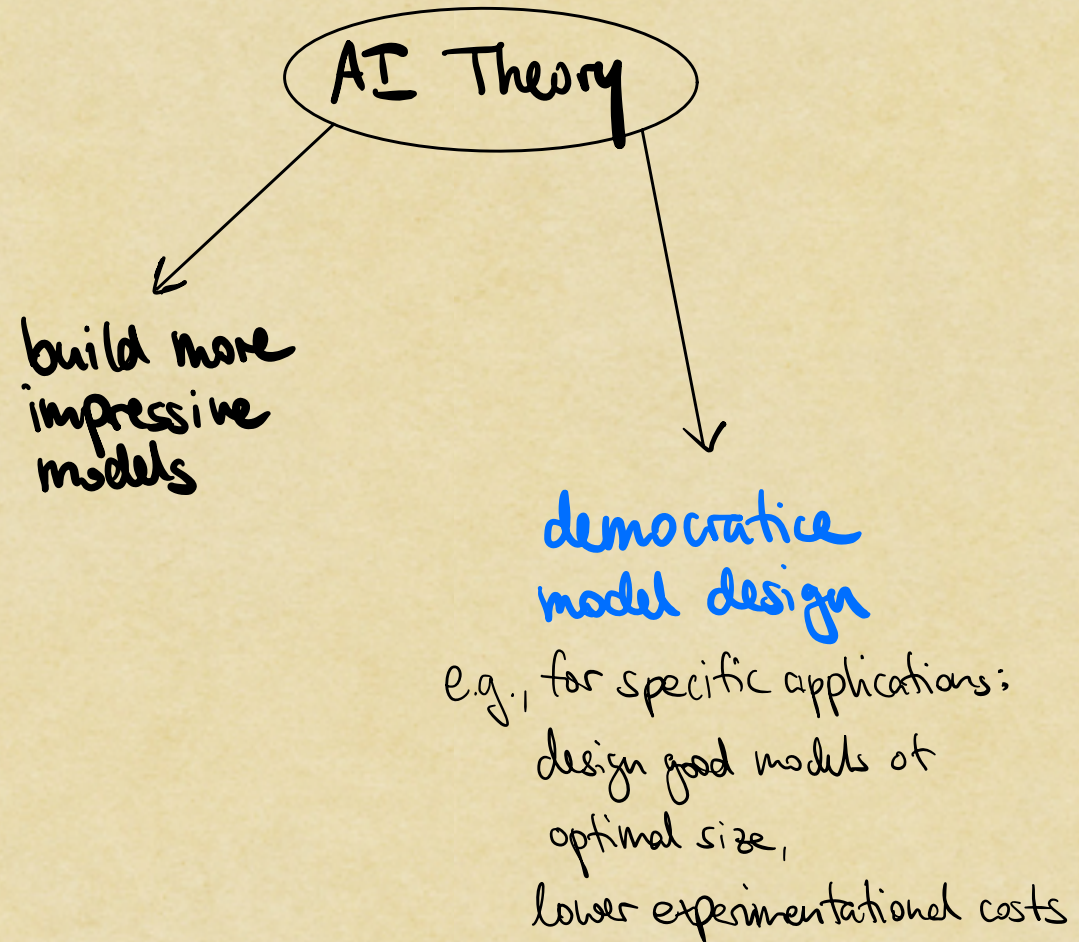
①



Humans have built bridges without mathematical theory for a long time, but theory helps us to compute stability, necessary resources, etc. in advance.

Role of AI Foundations

①



Humans have built bridges without mathematical theory for a long time, but theory helps us to compute stability, necessary resources, etc. in advance.

The Abstract Lens

(2)



Deep learning is like the fashion industry,
with different trends every summer

lottery tickets
implicit bias
highway networks
disentanglement
double descent
neural collapse
skip connections
representation learning
grokking
edge of stability

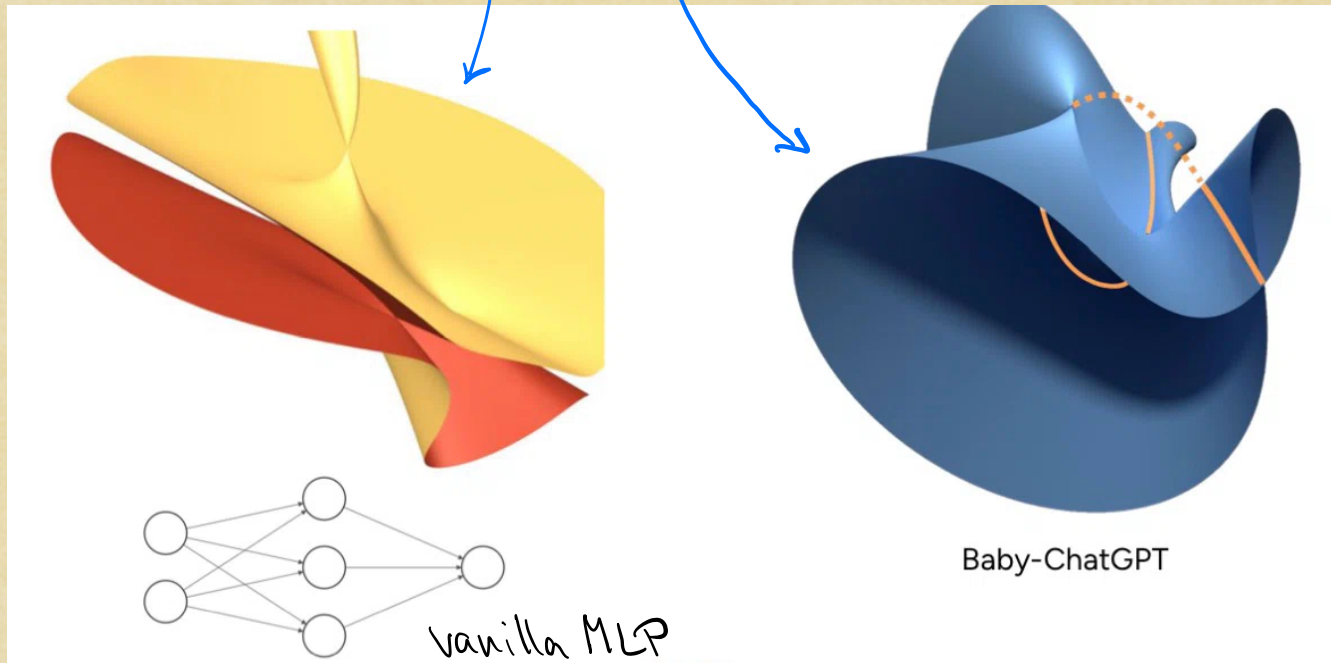
benign overfitting

use the abstract lens to
identify common principles

The Abstract Lens

3

Training a network means to optimize a loss on the **space of functions** parametrized by the network.



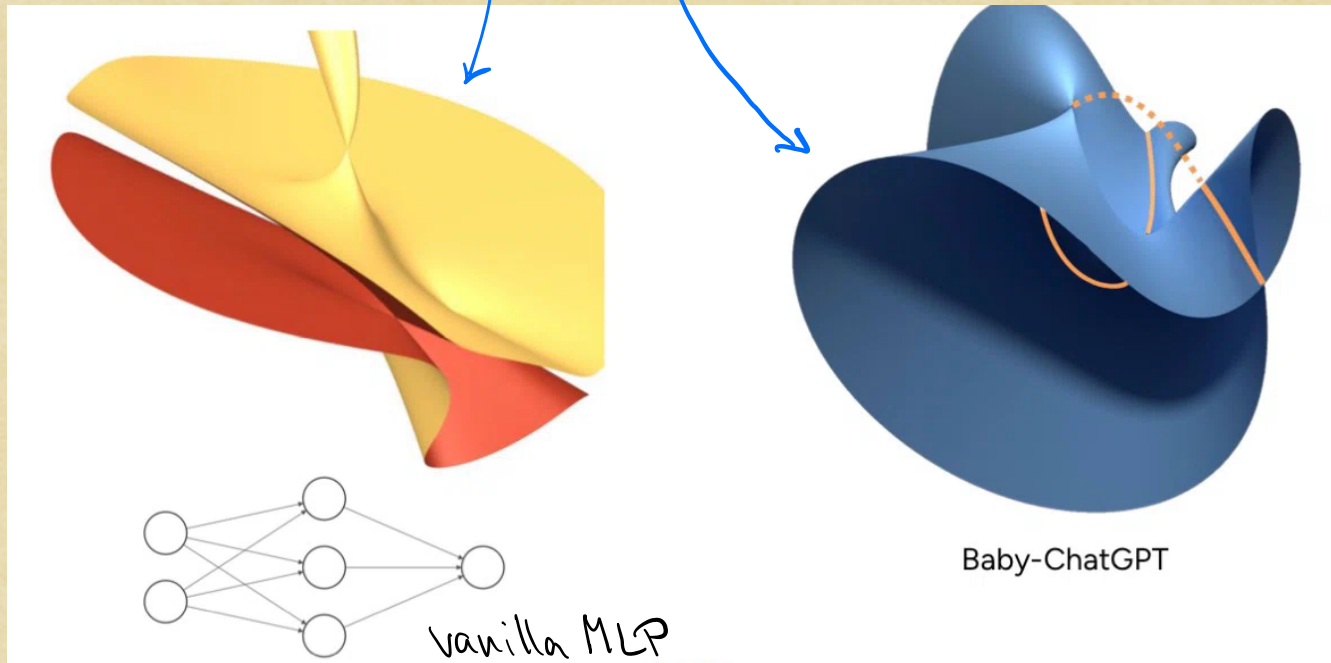
Every design choice
causes a **geometry**
of the function space.

⇒ impacts training &
learning capabilities

The Abstract Lens

3

Training a network means to optimize a loss on the **space of functions** parametrized by the network.



Every design choice causes a **geometry** of the function space.

⇒ impacts training & learning capabilities

Scaling up networks makes the geometry more pronounced!

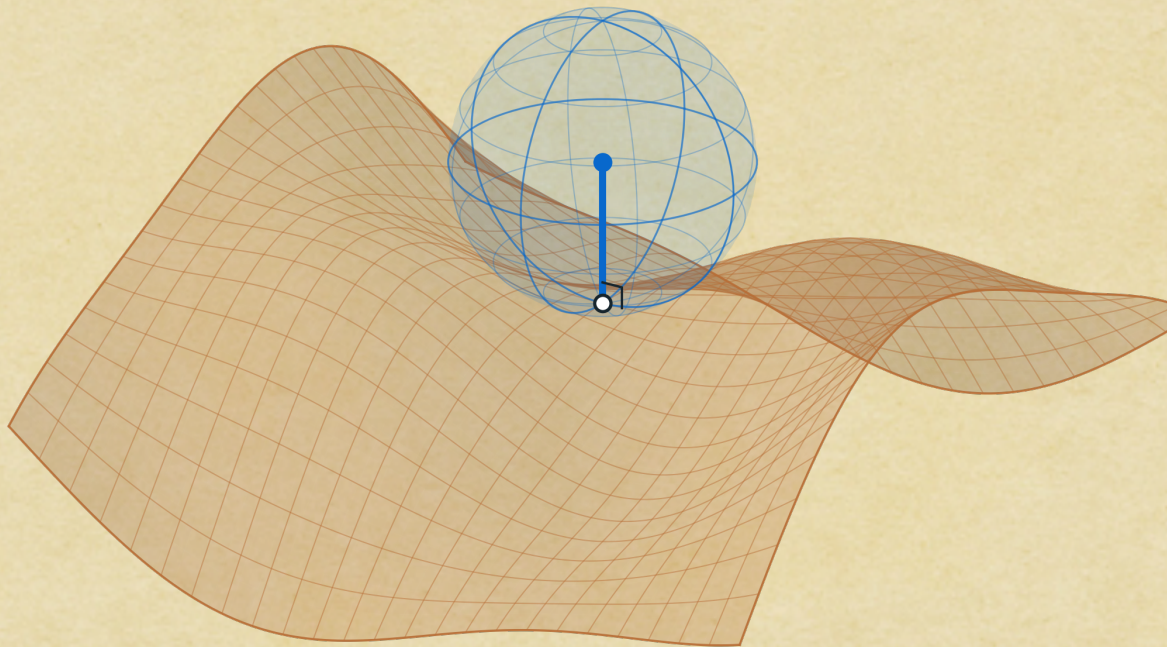
AI Safety

(4)

Potentially, in a few years from now, AI models saturate benchmarks at any task, but what about **reliability** & **trustworthiness**?

Reliable AI systems need to be built on a mix of systematic experiments & **provable guarantees**.

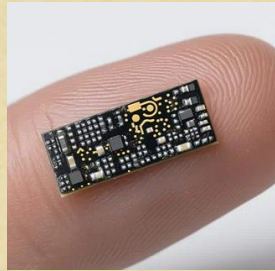
requires mathematical understanding of the inner workings of AI



Examples: Theory underpinning successful AI

⑤

- CNNs based on symmetry
- Diffusion models based on stochastic processes
- Design of TinyML
- Optimal fine tuning

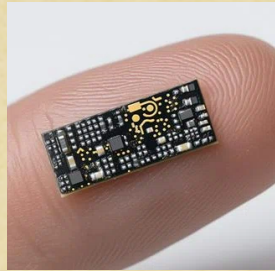


Examples: Theory underpinning successful AI

⑤

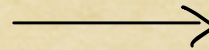
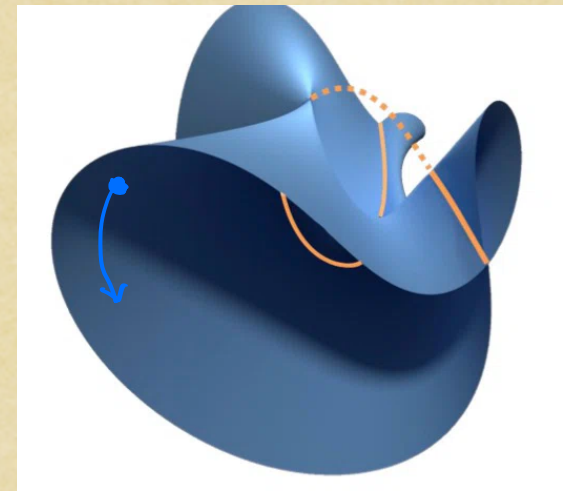
- CNNs based on symmetry
- Diffusion models based on stochastic processes

- Design of TinyML



- Optimal fine tuning:

fine tuning
trains only
along a **small**
portion of
the huge
parameter
space



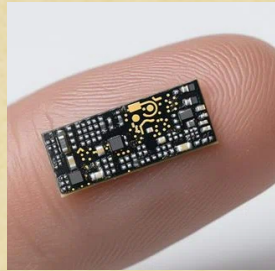
function space

Examples: Theory underpinning successful AI

⑤

- CNNs based on symmetry
- Diffusion models based on stochastic processes

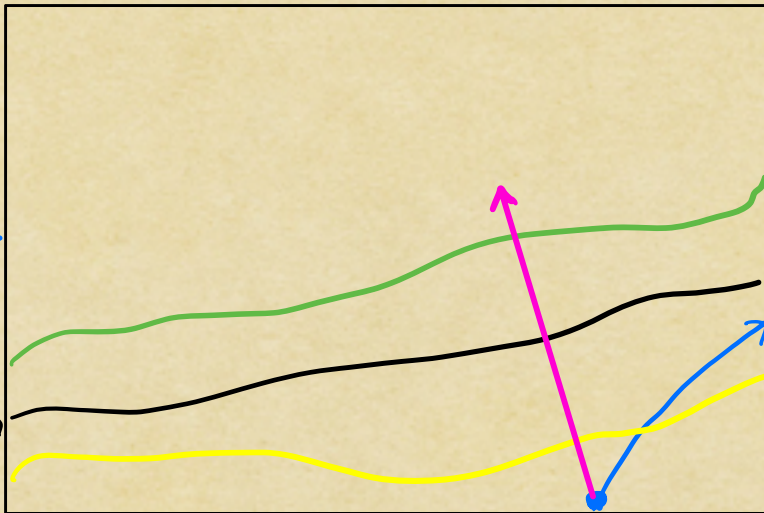
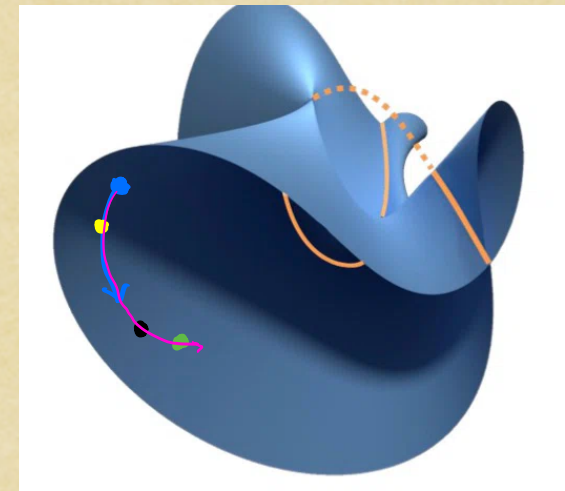
- Design of TinyML



- Optimal fine tuning:

instead of doing standard choices (like LoRA),
fine tune in **directions with least redundancy**
⇒ fastest gain

fine tuning
trains only
along a **small**
portion of
the huge
parameter
space



function space

Some Open Challenges

⑥

How to scale with built in structure / constraints / symmetry?

Why exactly do networks generalize well to new data?

What is an optimal architecture for a given application & hardware?

What is the optimal compute needed for a given task?

⋮