Discovering physical concepts with neural networks

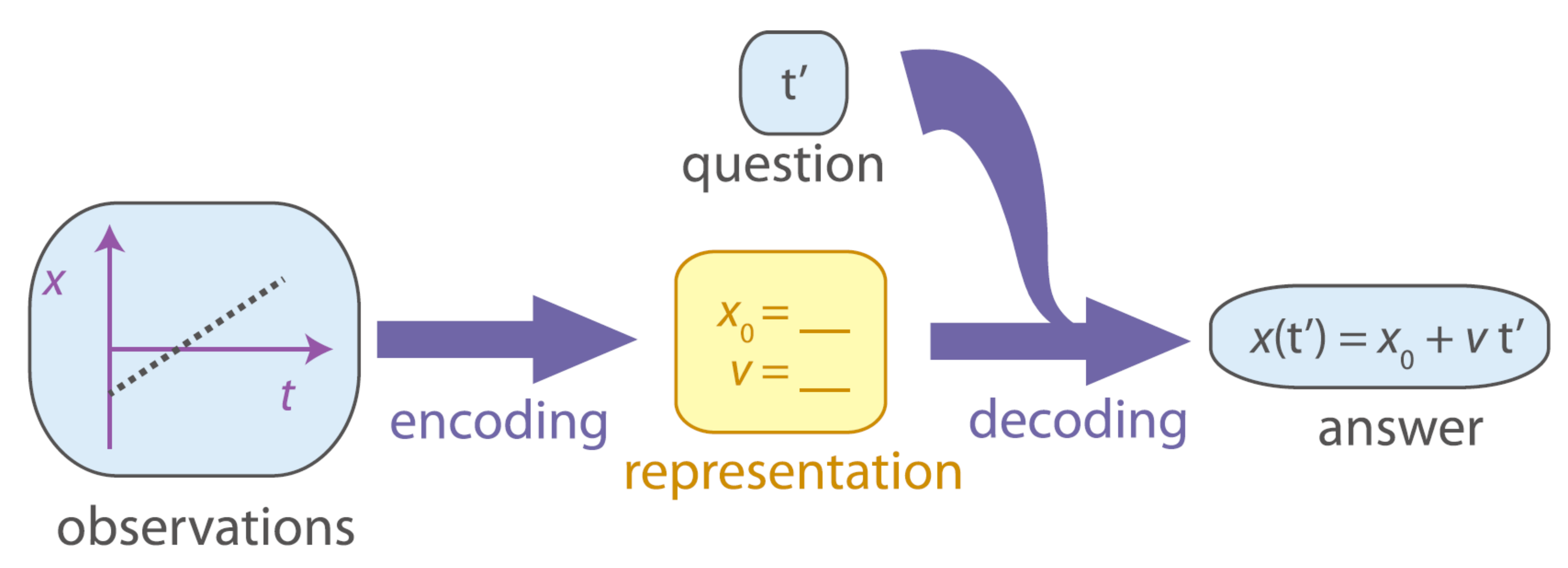
Nils Rörup

May 7th, 2019

Agenda

- Objective: From Observations to Models
- Necessary Network Architecture
- Results on Physical Experiments
- Discussion

Objective

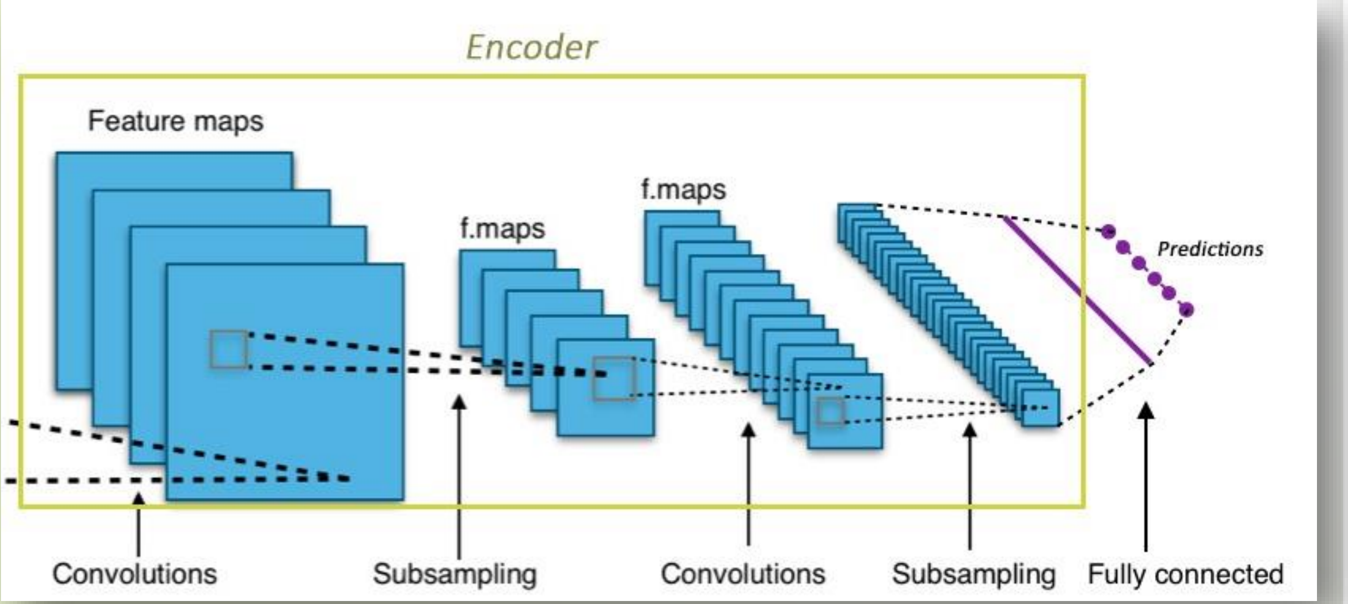


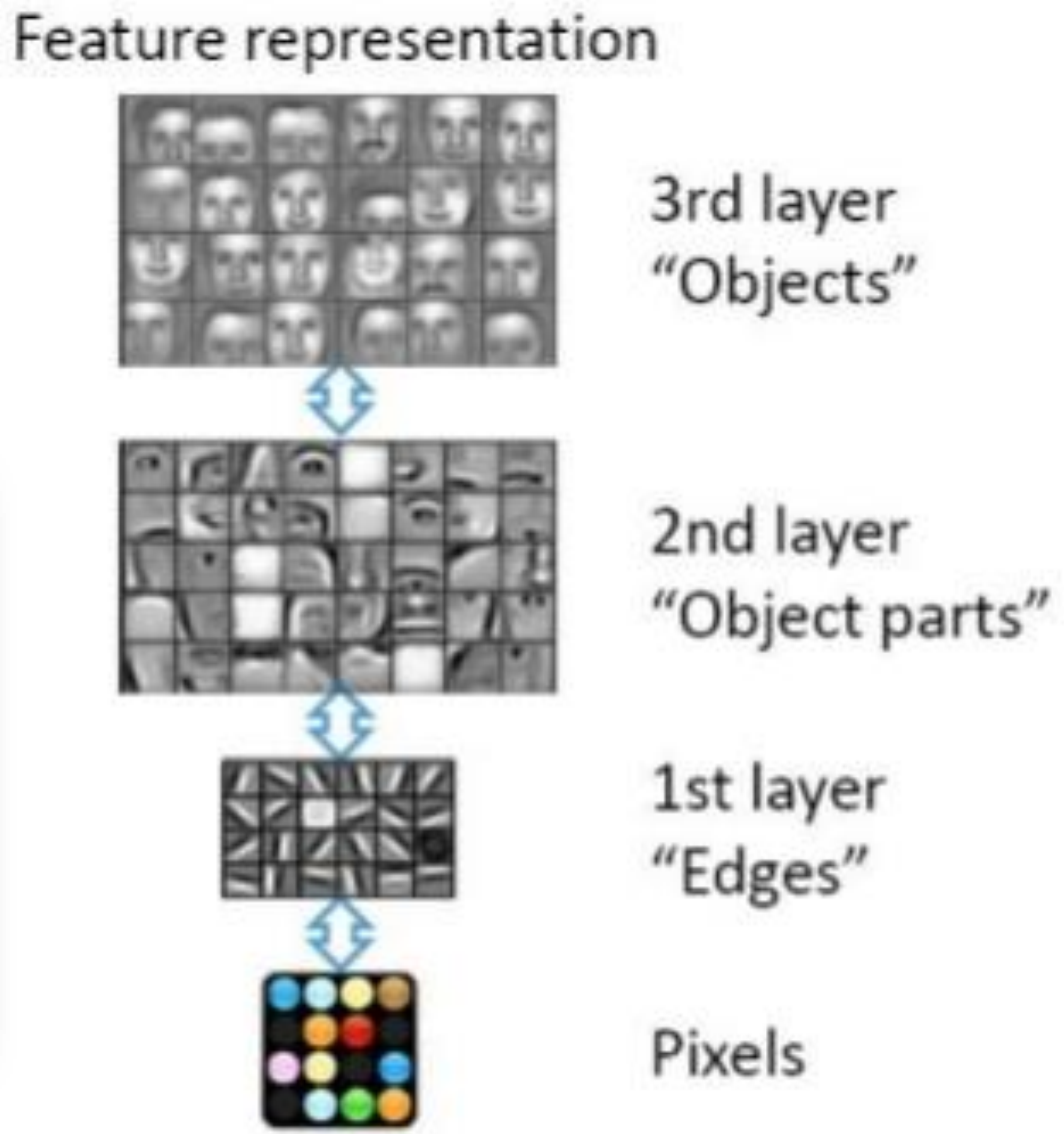
Finding a simple Model

Network Architecture

Representation Learning

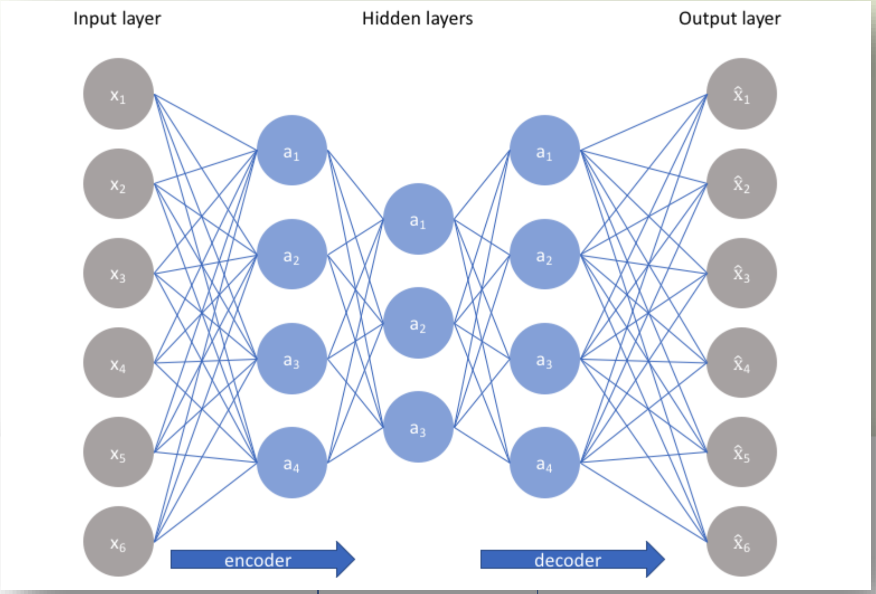
Learning representations:
Levels of Abstraction

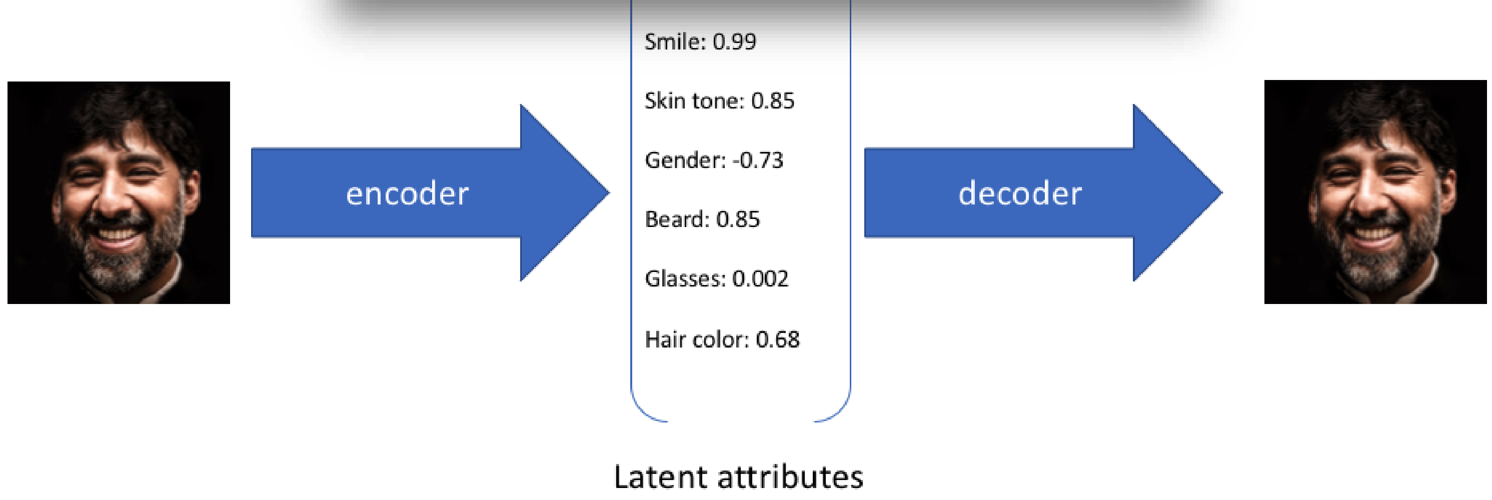




Learning representations: AutoEncoder

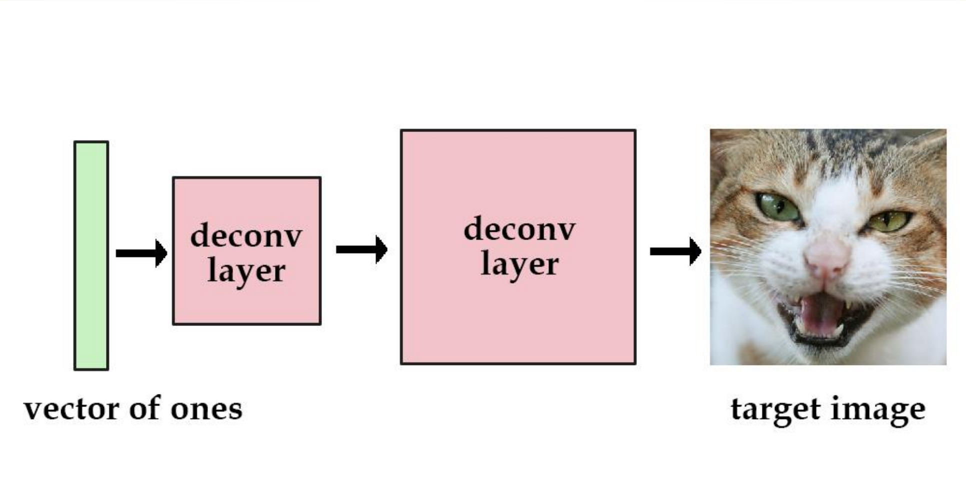
- Bottleneck Architecture





AutoEncoder: A second perspective

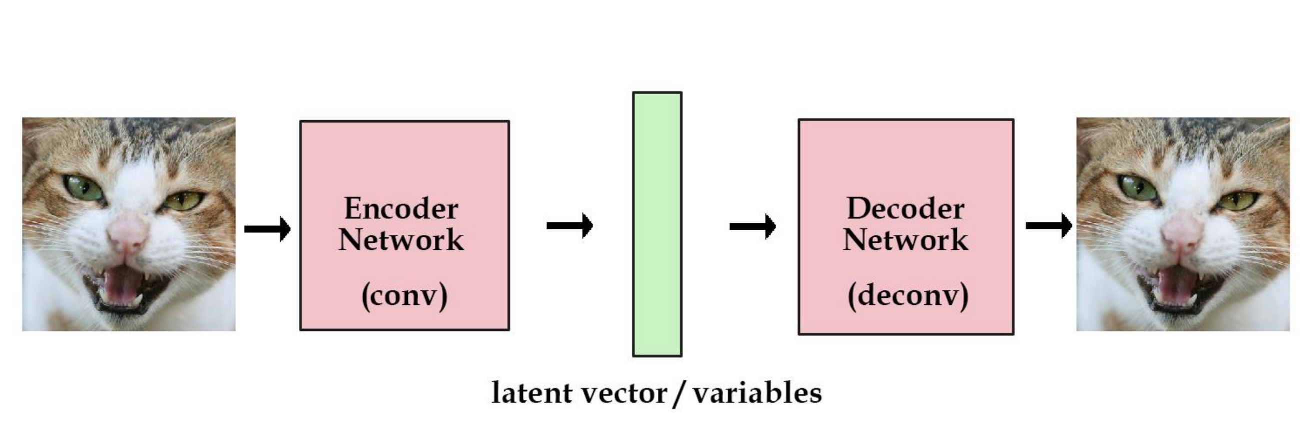
- Use NN as Datastructure:



- Even: Train to output different pictures, based on input vector.

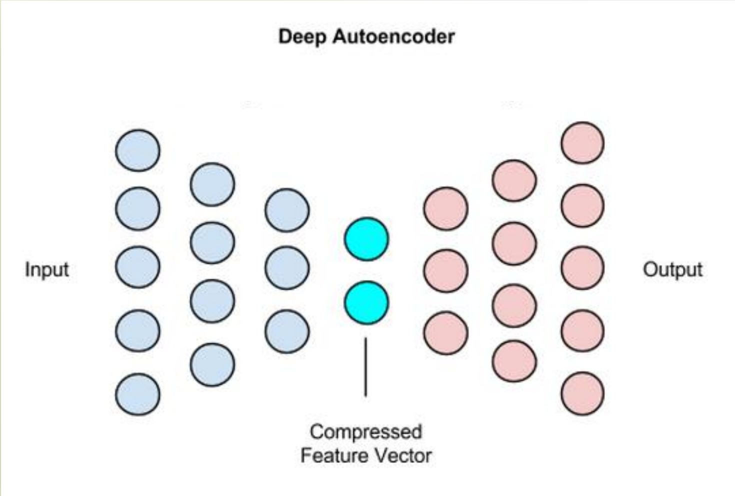
AutoEncoder: A second perspective

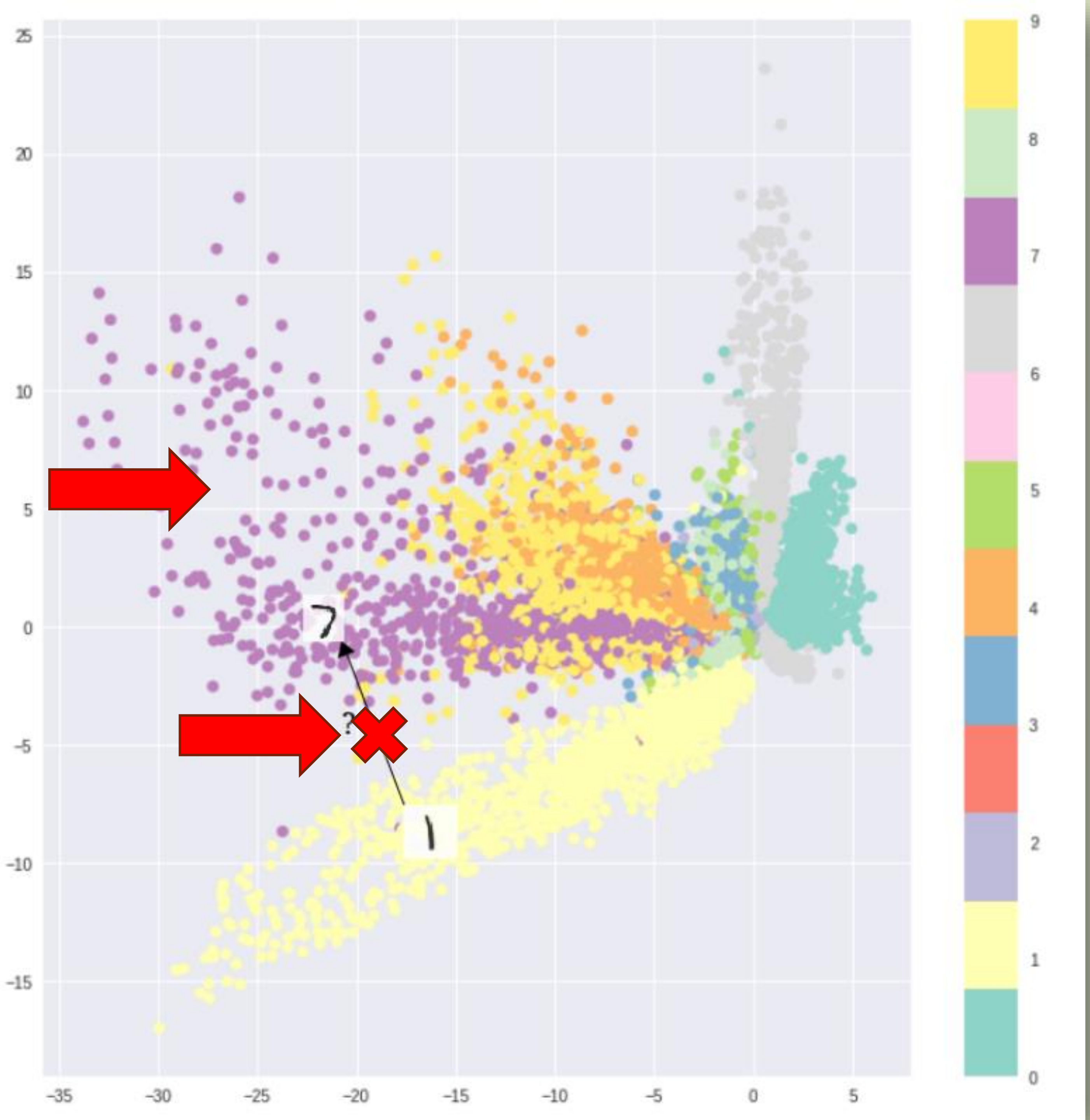
- Train encoder to organise latent space
  - Tada! Autoencoder!



AutoEncoder

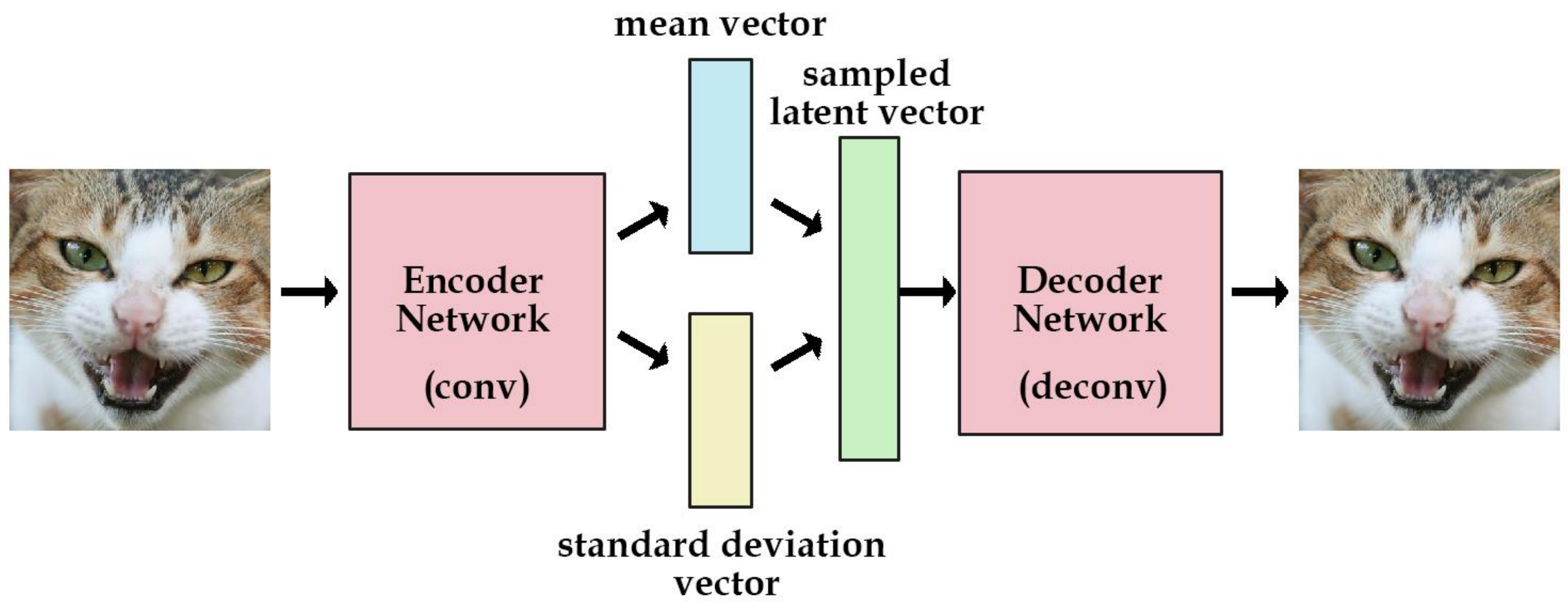
- latent space trained on MNIST.
- 2 latent neurons
  - 2-dim. latent space



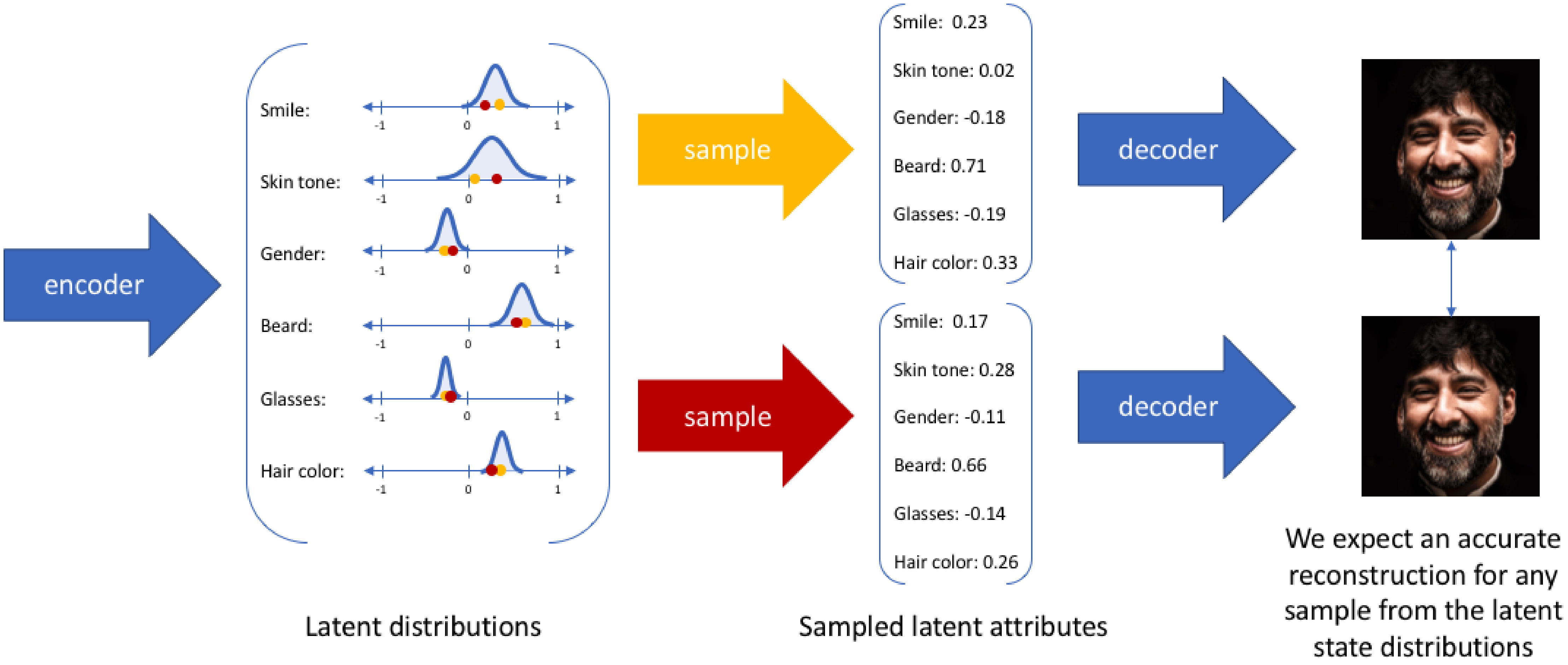


Variational AutoEncoder

- Autoencoder stores points.
- VAE stores Gaussian distributions!



Variational AutoEncoder



Variational AutoEncoder: Loss

- Root Mean Square Error: RSME = $\sum_{i \in test\ set} \left(y_{pred,i} - y_{true,i}\right)^2$
  - Encourages the network to describe the input

Variational AutoEncoder: Loss

- Root Mean Square Error: RSME = $\sum_{i \in test\ set} \left(y_{pred,i} - y_{true,i}\right)^2$
  - Encourages the network to describe the input
- Kullback-Leibler Divergence: $$D_{\mathrm{KL}}(P \parallel Q) = \int_{-\infty}^{\infty} p(x) \log\left(\frac{p(x)}{q(x)}\right) dx$$
  - Distance between distributions

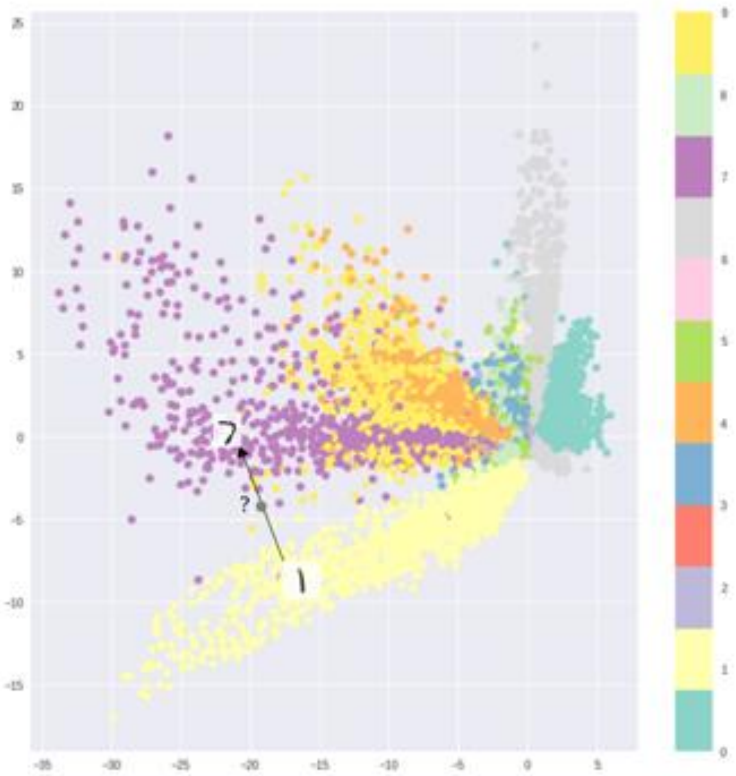
Variational AutoEncoder: Loss

- Root Mean Square Error: RSME = $\sum_{i \in test\ set}\left(y_{pred,i} - y_{true,i}\right)^2$
  - Encourages the network to describe the input
- Kullback-Leibler Divergence: $$D_{\mathrm{KL}}(P \parallel Q) = \int_{-\infty}^{\infty} p(x) \log\left(\frac{p(x)}{q(x)}\right) dx$$
  - Distance between distributions
- $\mathcal{L} = RMSE + \beta \cdot D_{KL}(\ P || N(\mu = 0, \sigma = 1)\ )$

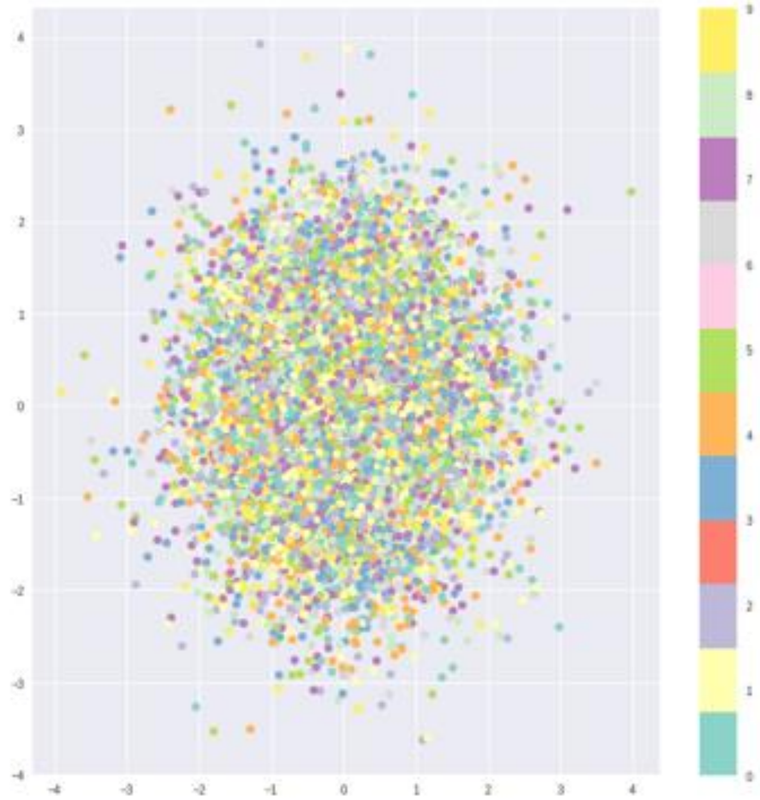
Variational AutoEncoder: latent space

$$\mathcal{L} = RMSE + \beta \cdot D_{KL}(\, P || N(\mu = 0, \sigma = 1)\ )$$

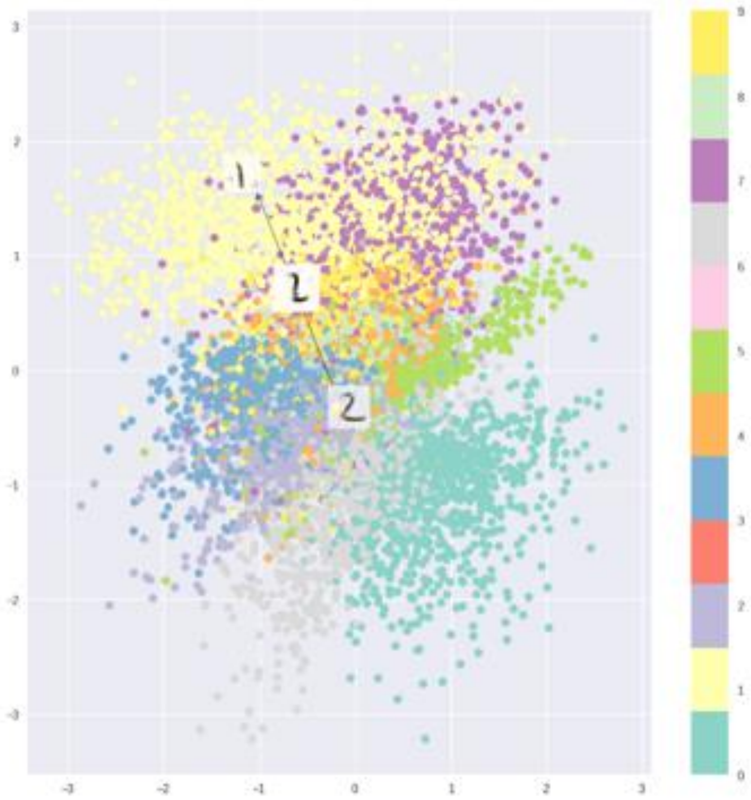
Only reconstruction loss



Only KL divergence

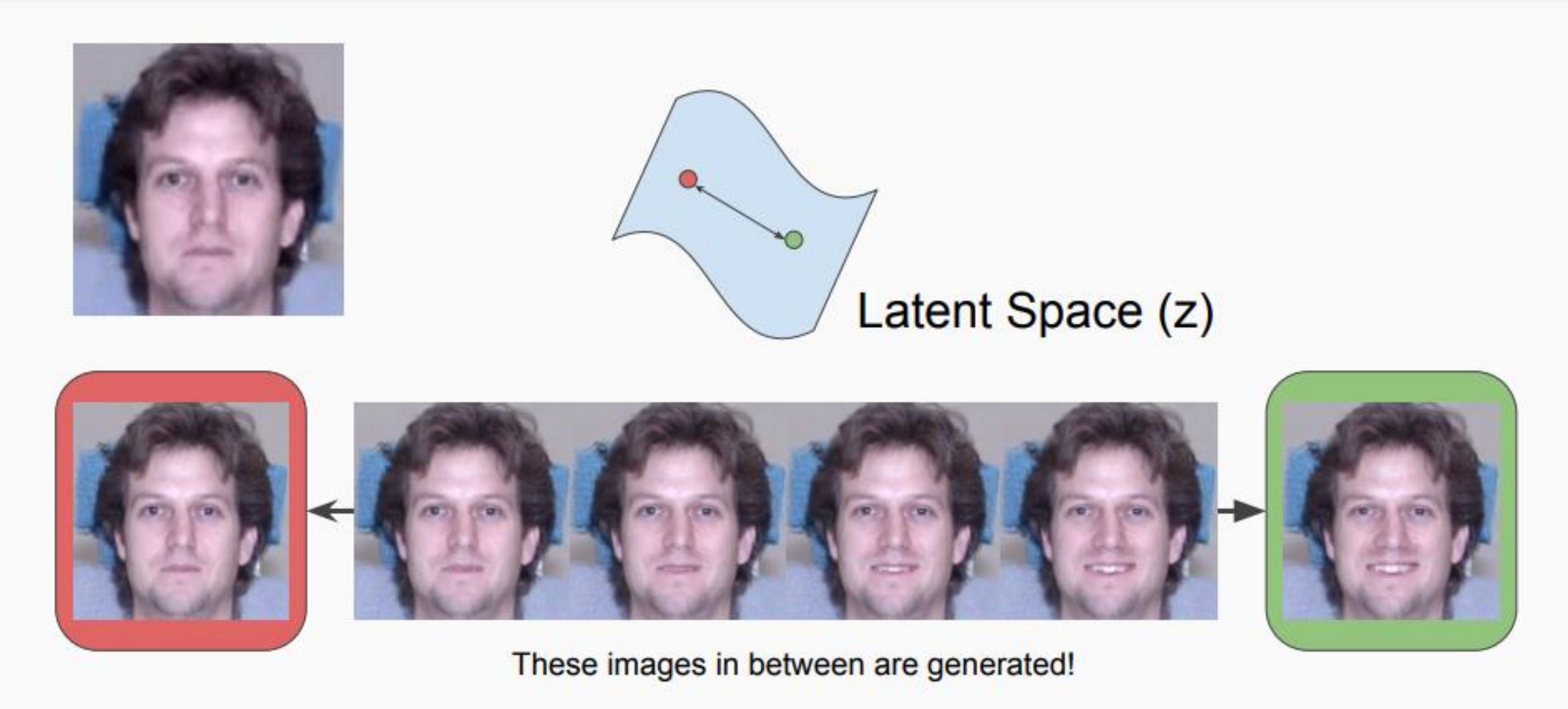


Combination



Variational AutoEncoder

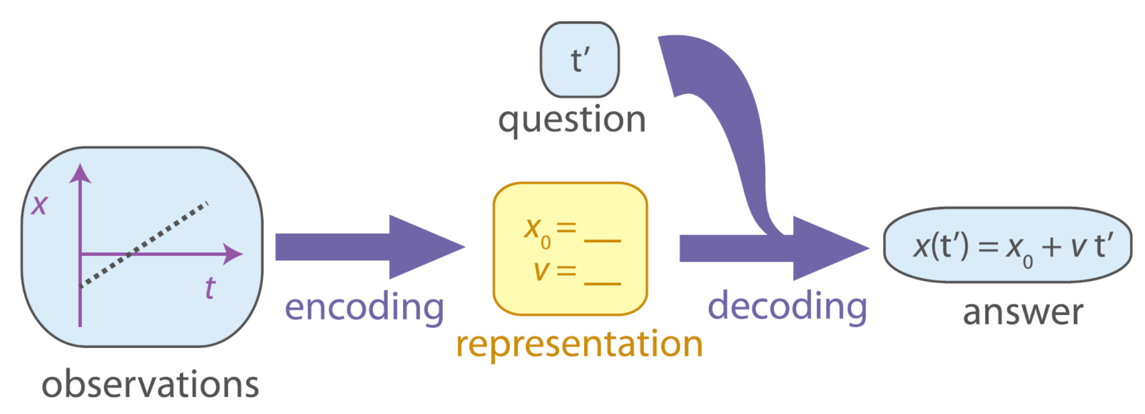
- Result: a dense, continuous latent space.



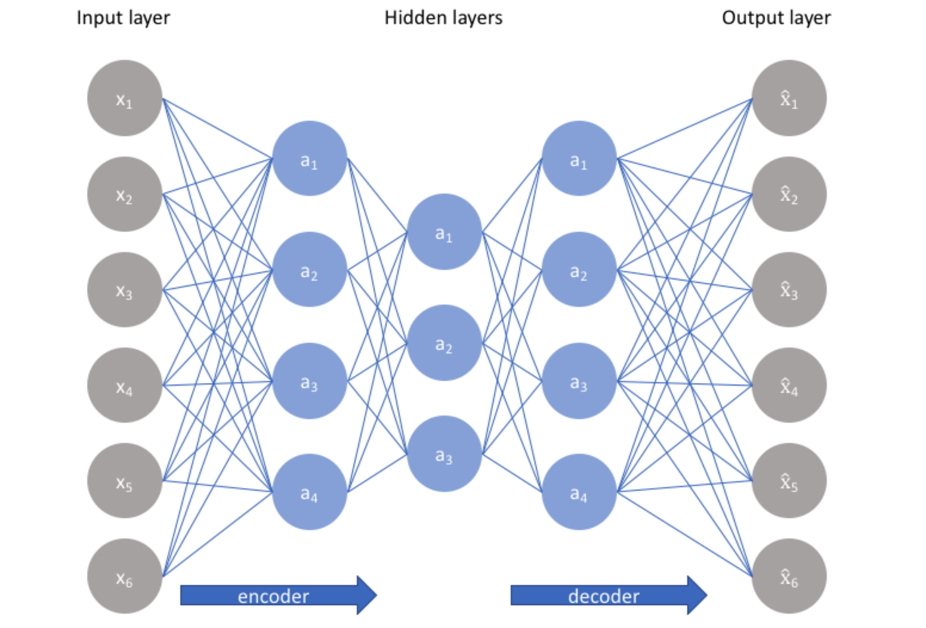
SciNet Architecture

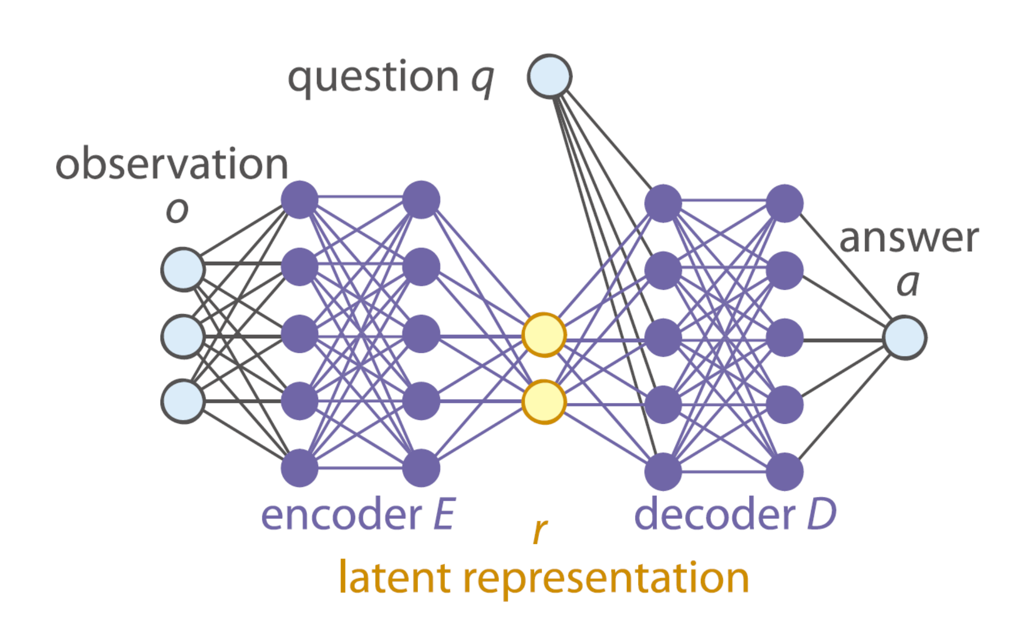
- Train on observations & questions and answers
- Model parameters ~ latent repres.

SciNet!



- Compare to Autoencoder:

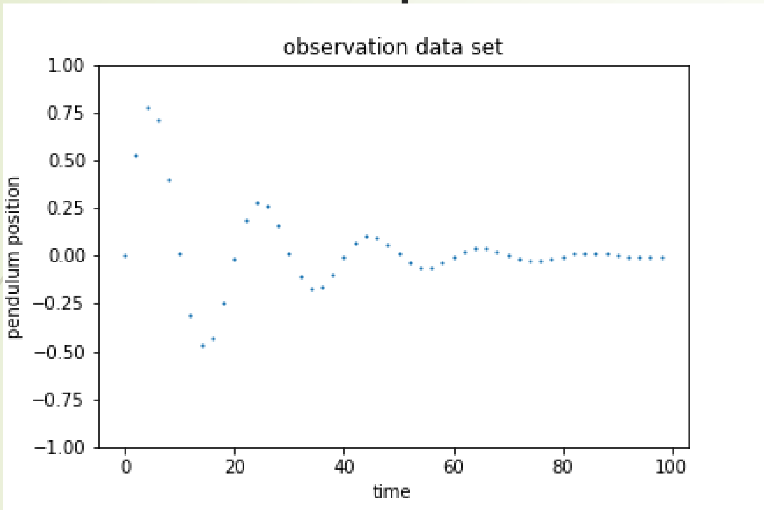


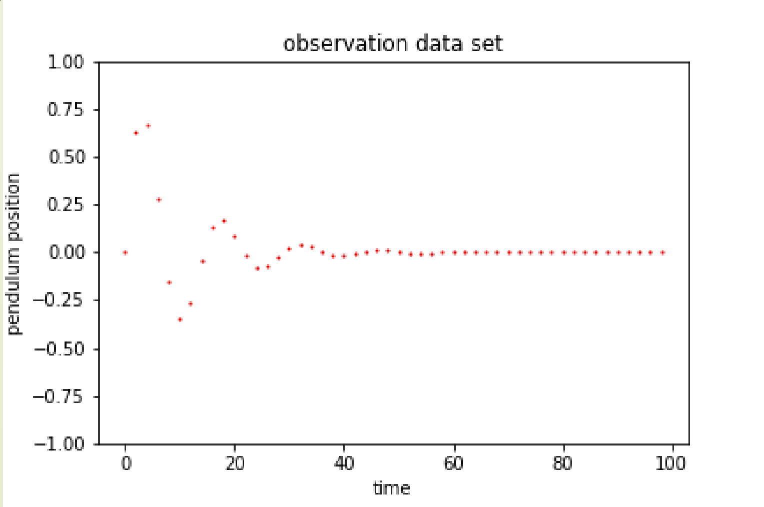


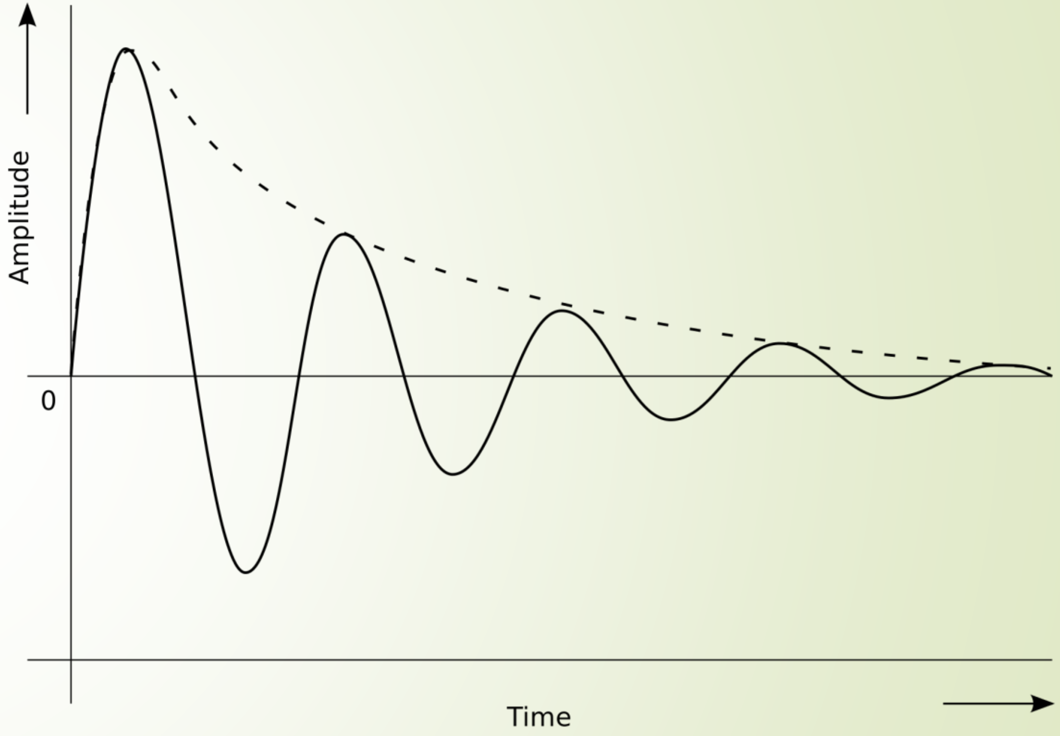
Damped Pendulum

Identifying system parameters

Damped Pendulum

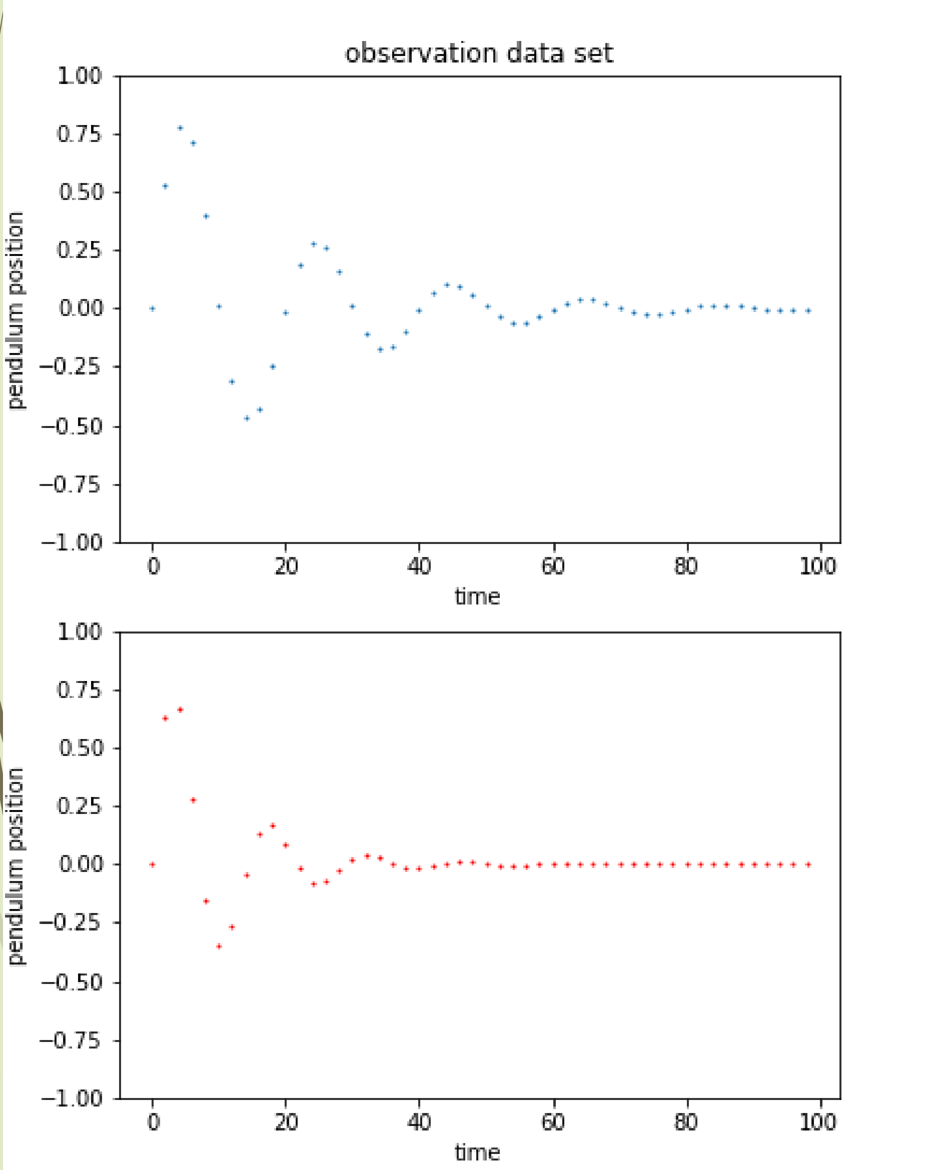


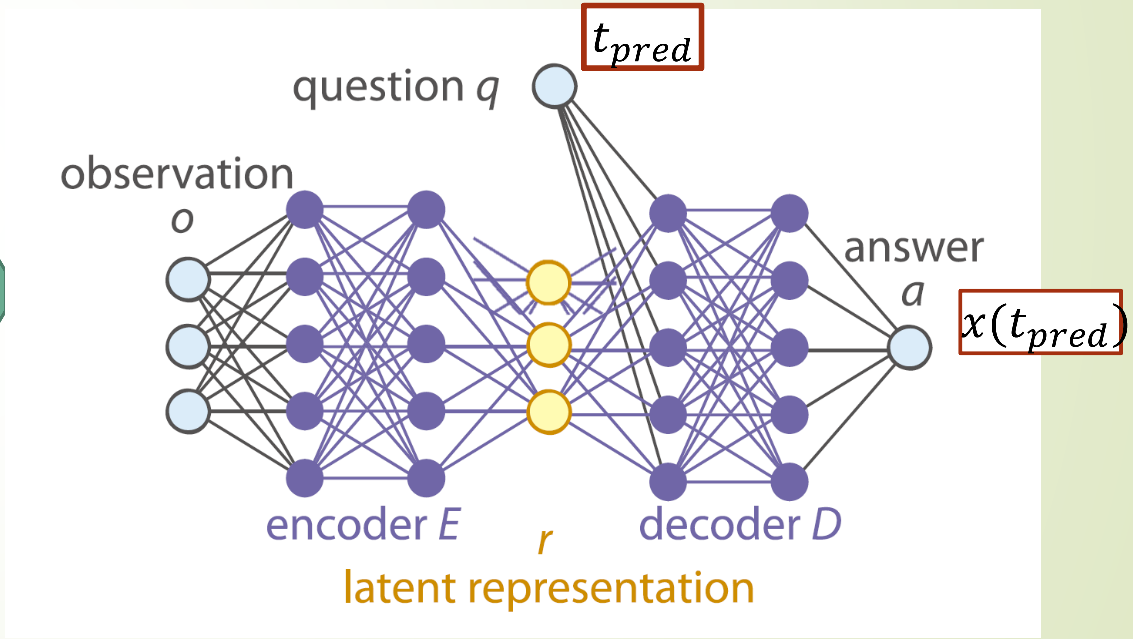




- $x(t) = A_0 e^{-\frac{\kappa t}{2}} \sin(\omega t), \;\; with \;\; \omega = \sqrt{\kappa(1 - \frac{b^2}{4\kappa})}$

Damped Pendulum



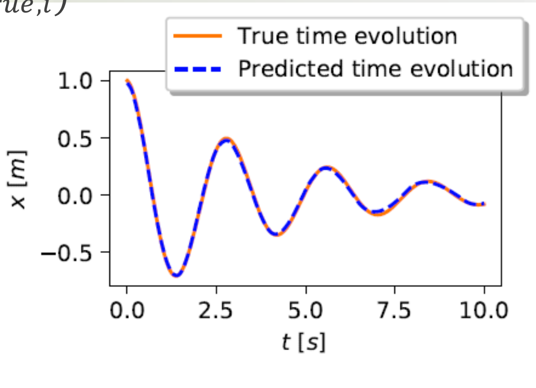


$$\ddot{x} + b\dot{x} + \kappa x = 0$$

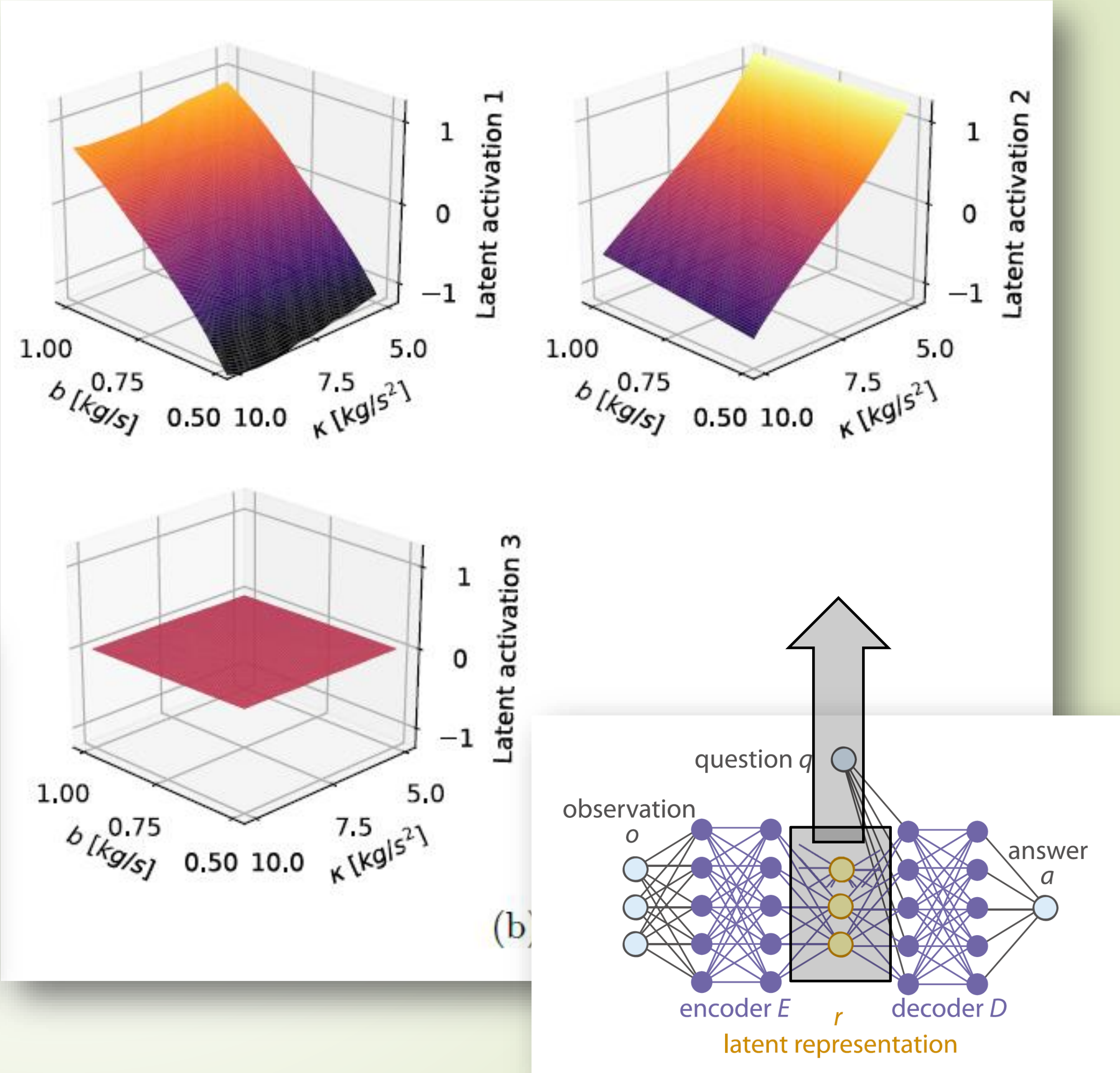
Key Findings

- $\ddot{x} + b\dot{x} + \kappa x = 0$
- κ and b in two latent neurons
- no information in the 3rd neuron.

- RMSE = $\sum_{i \in test\ set} \left(y_{pred,i} - y_{true,i}\right)^2$
- RMSE of 2% of amplitude



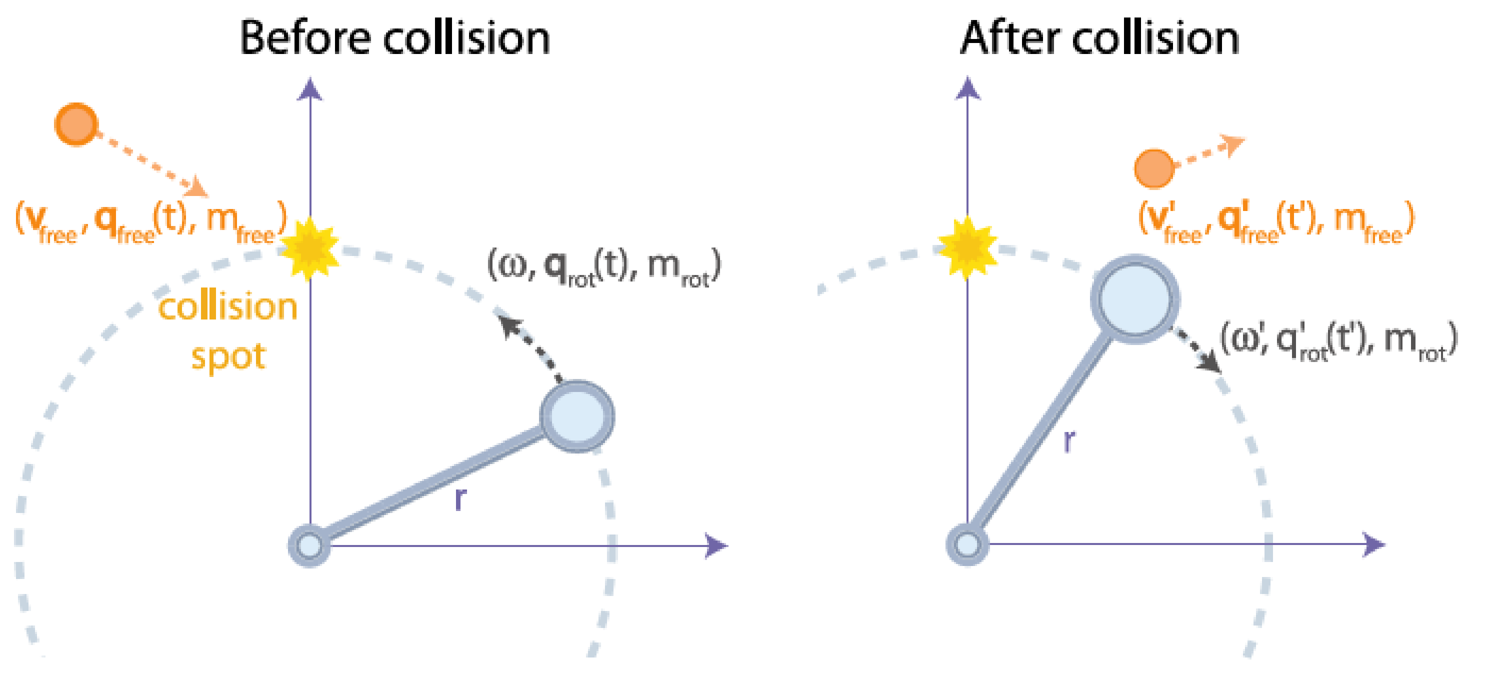
(a)



(b)

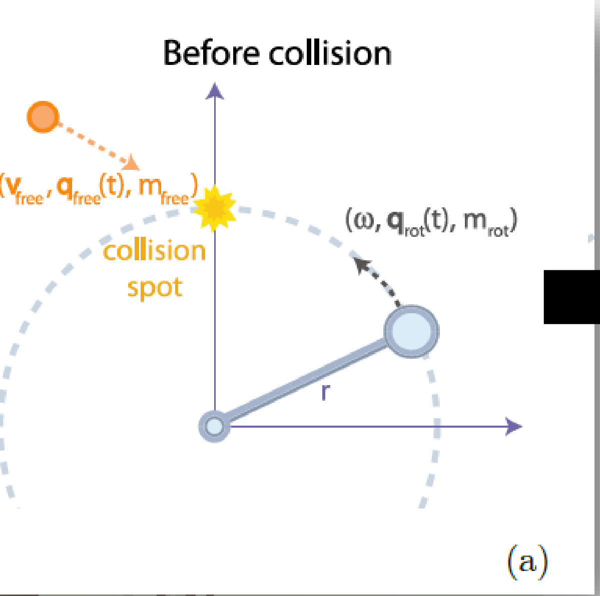
Collision

Using constants of motion

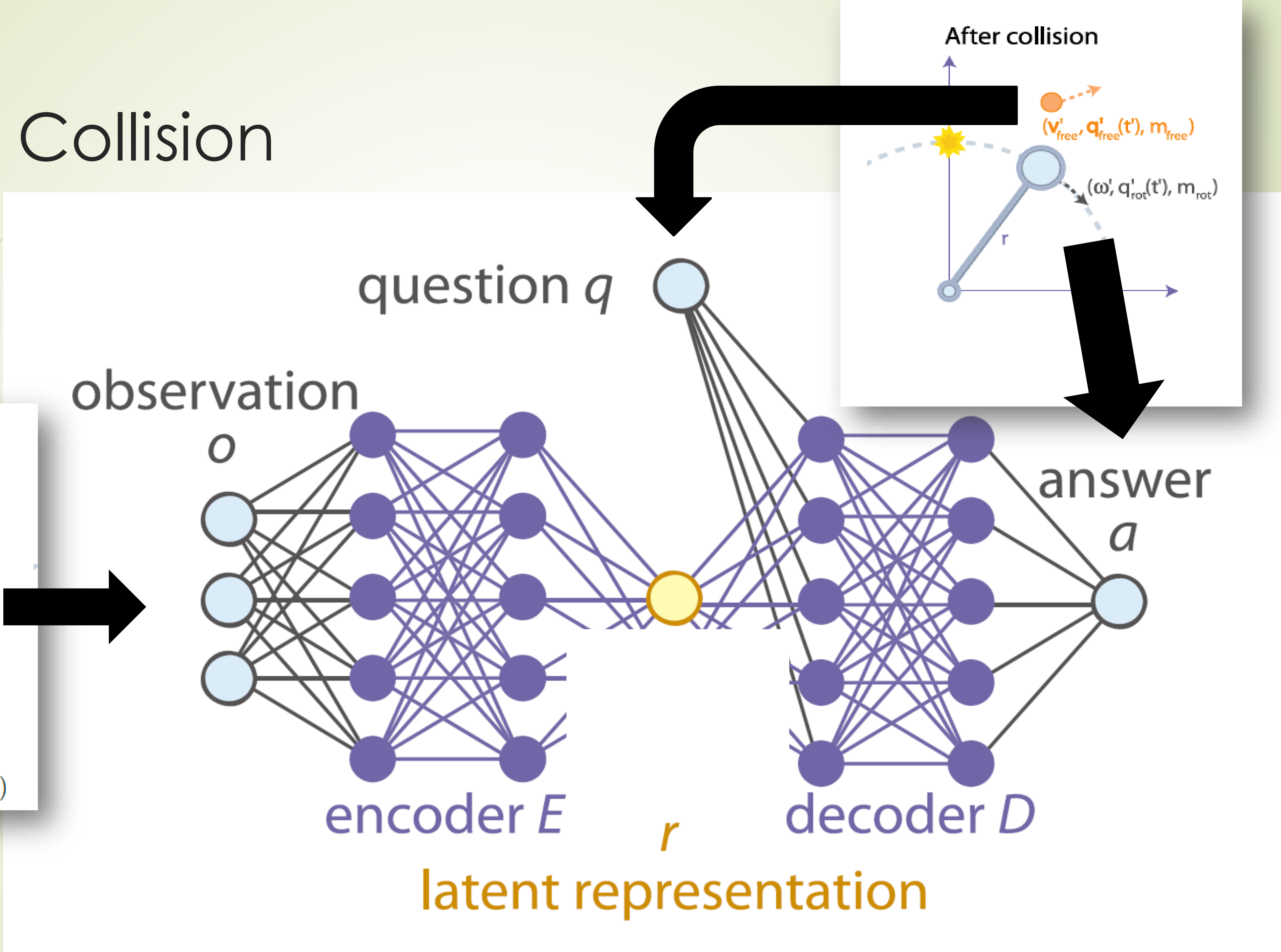

Before collision
After collision
(v_{free}, q_{free}(t), m_{free})
collision spot
(ω, q_{rot}(t), m_{rot})
r
(v'_{free}, q'_{free}(t'), m_{free})
(ω', q'_{rot}(t'), m_{rot})
r
(a)

Collision

Collision

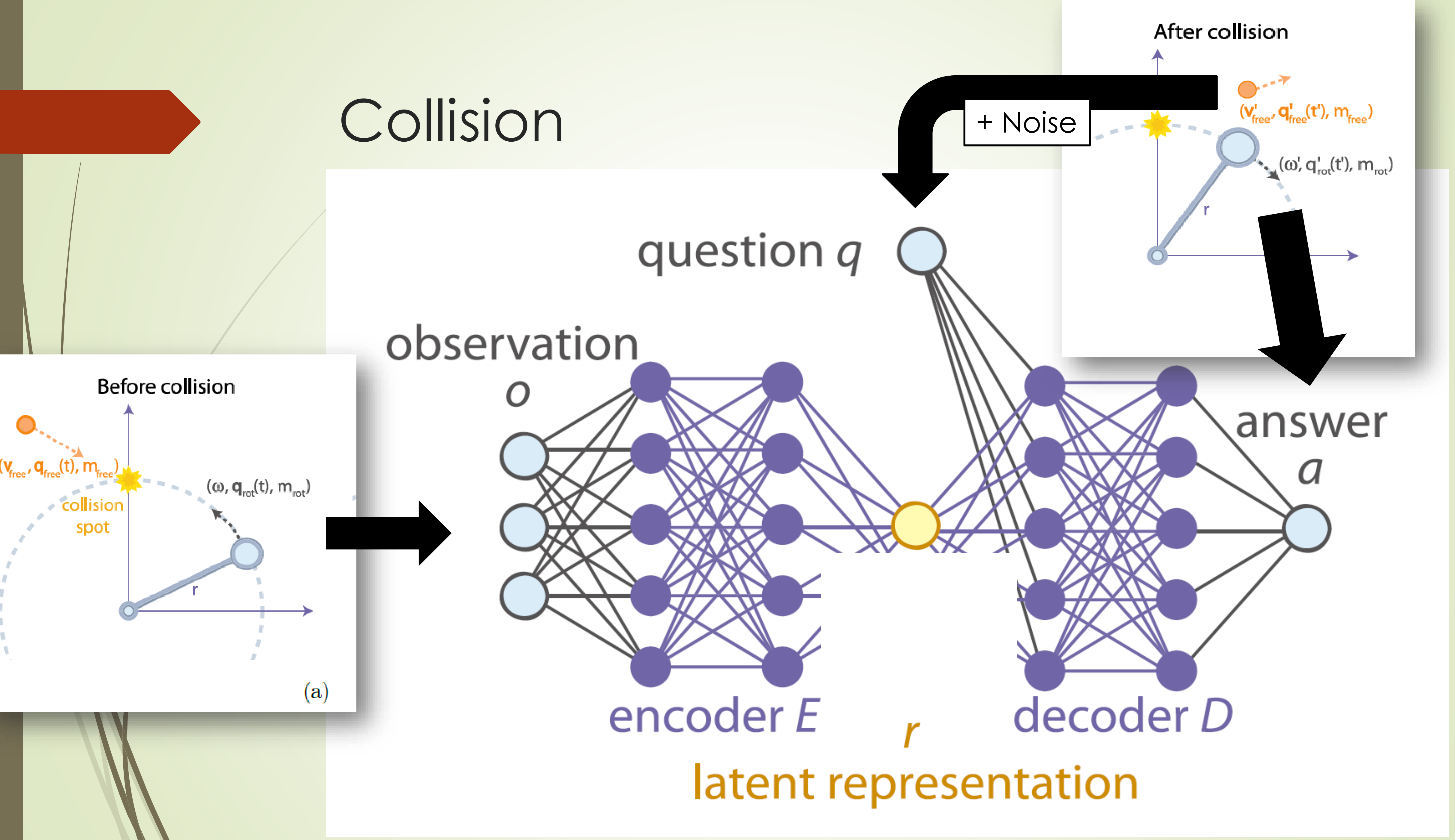


Before collision
(v_free, q_free(t), m_free)
collision spot
(ω, q_rot(t), m_rot)
r
(a)



question q
observation o
answer a
encoder E
r
latent representation
decoder D
After collision
(v'_free, q'_free(t'), m_free)
(ω', q'_rot(t'), m_rot)
r

Collision

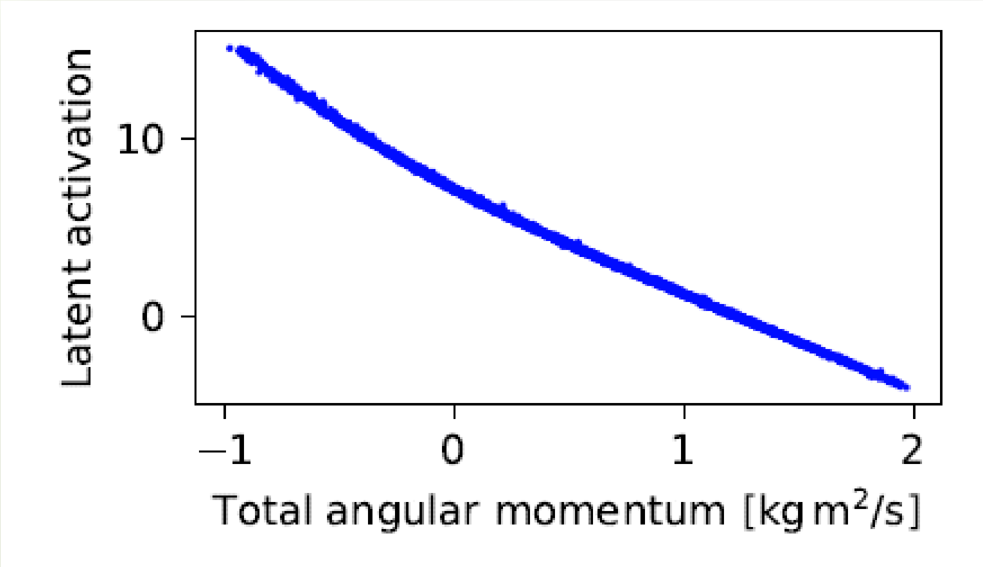


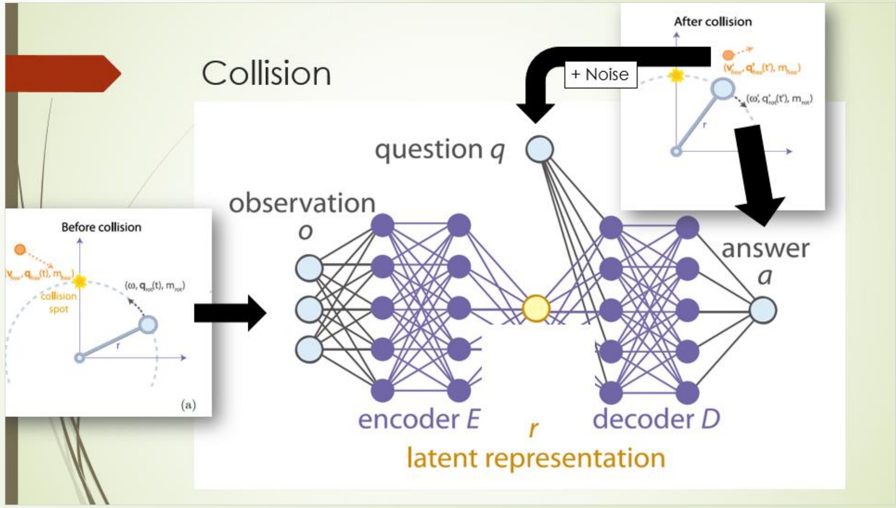
Before collision
$(v_{free}, q_{free}(t), m_{free})$
collision spot
$(\omega, q_{rot}(t), m_{rot})$
r
(a)
+ Noise
After collision
$(v'_{free}, q'_{free}(t'), m_{free})$
$(\omega', q'_{rot}(t'), m_{rot})$
r
question q
observation o
answer a
encoder E
r
decoder D
latent representation

Key Findings: Collision

- stores total angular momentum in the latent neuron!

- RMSE = 4% of radius *r*.
- *SciNet* is resistant to noise.





Qubits

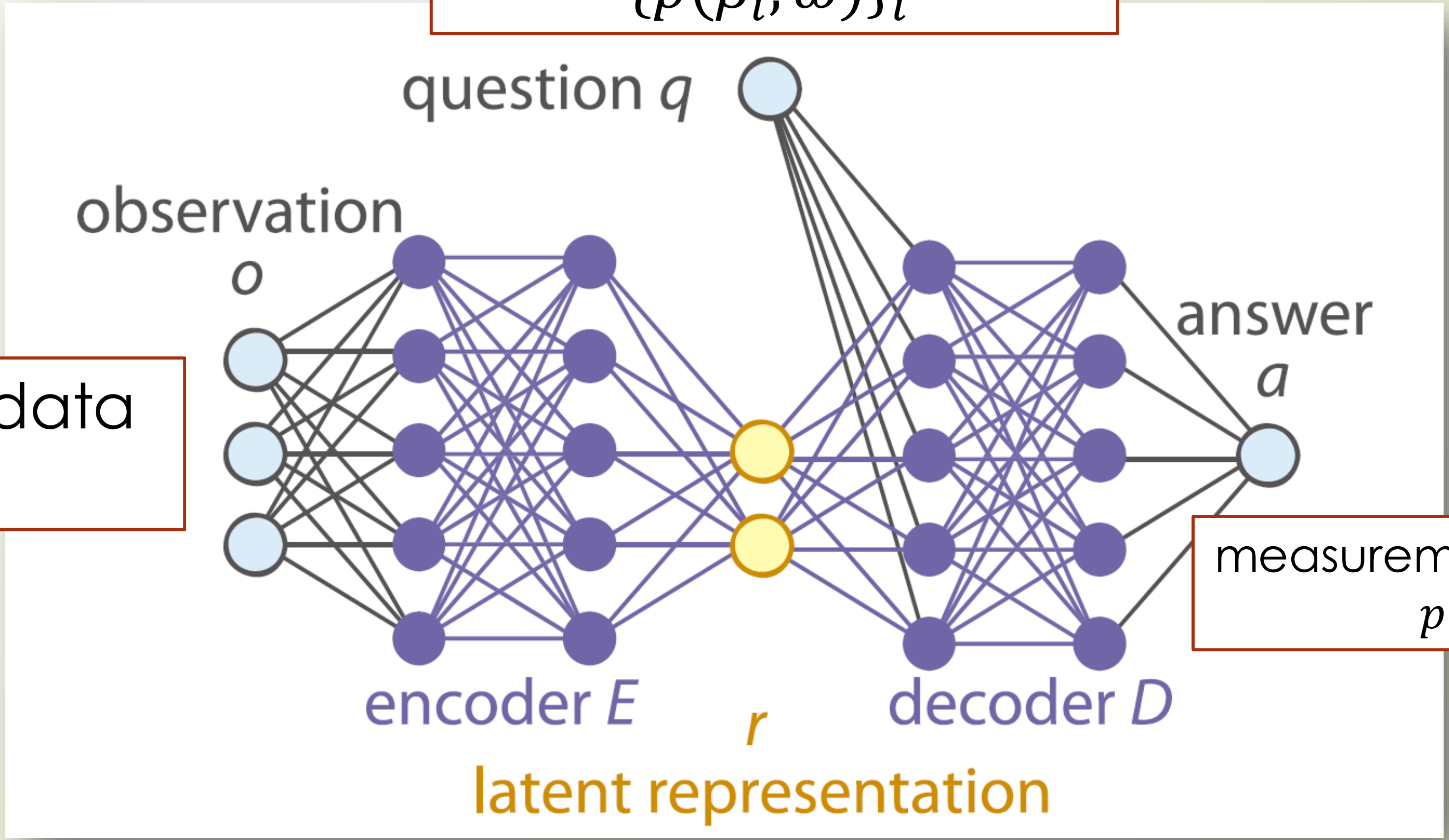
Counting Degrees of Freedom

Qubits: reconstruction of quantum state ψ

Qubits: reconstruction of quantum state ψ

parameterization
of a measurement ω:
$\{p(\beta_i, \omega)\}_i$

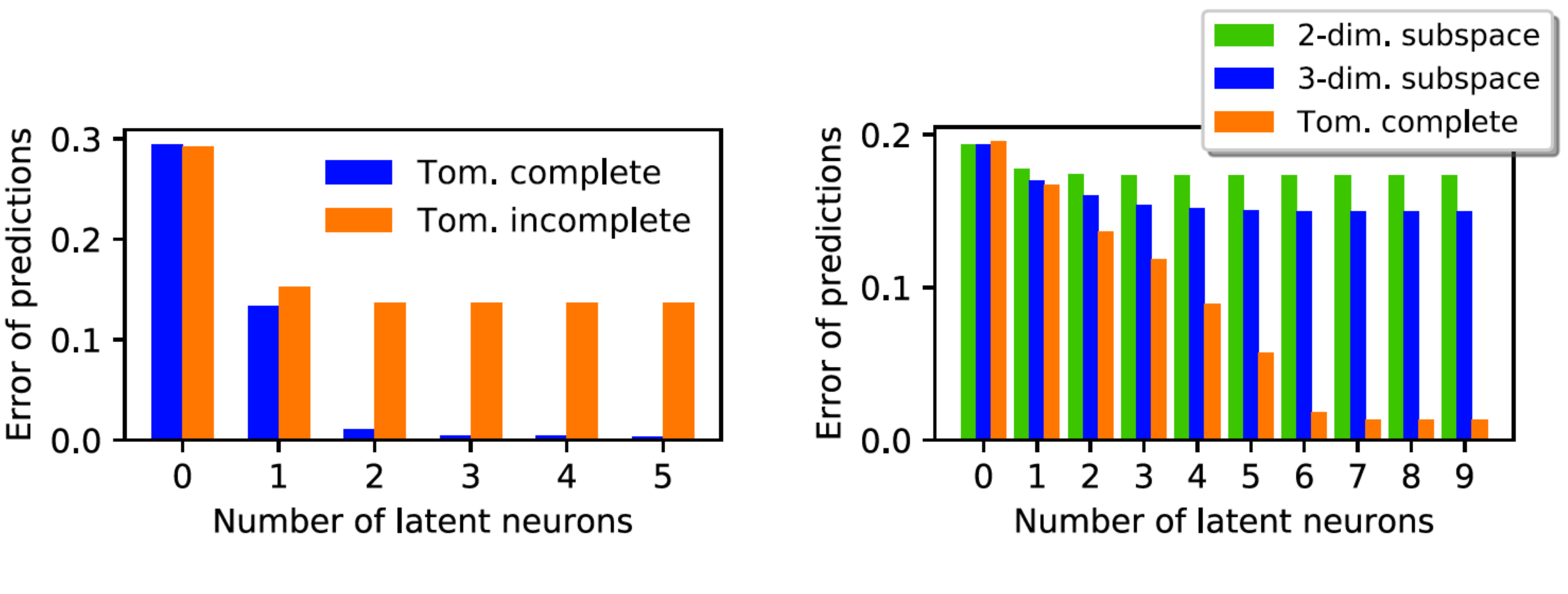
measurement data
$\{p(\alpha_i, \psi)\}_i$



measurement probability
$p(\omega, \psi)$

Key Findings: Qubits

- SciNet determines degrees of freedom in state
- SciNet distinguishes tomographically complete sets



(a) One qubit.

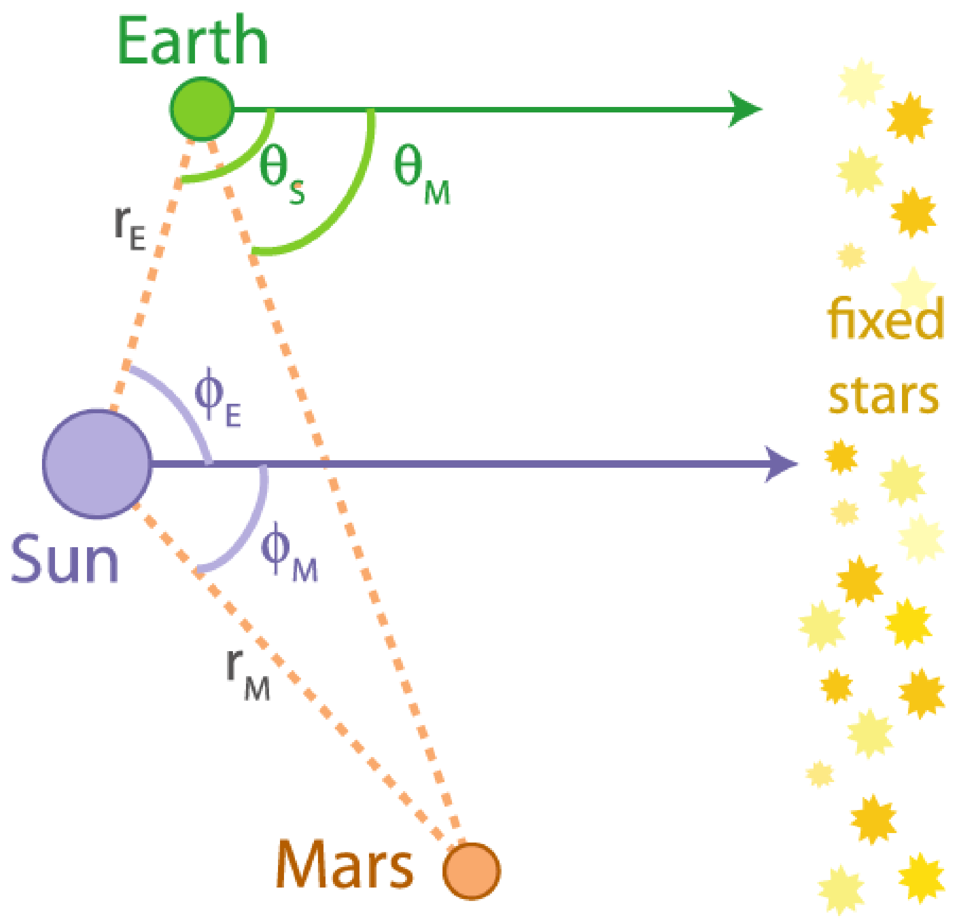
(b) Two qubits.

Heliocentric solar system

Smart choice of coordinates

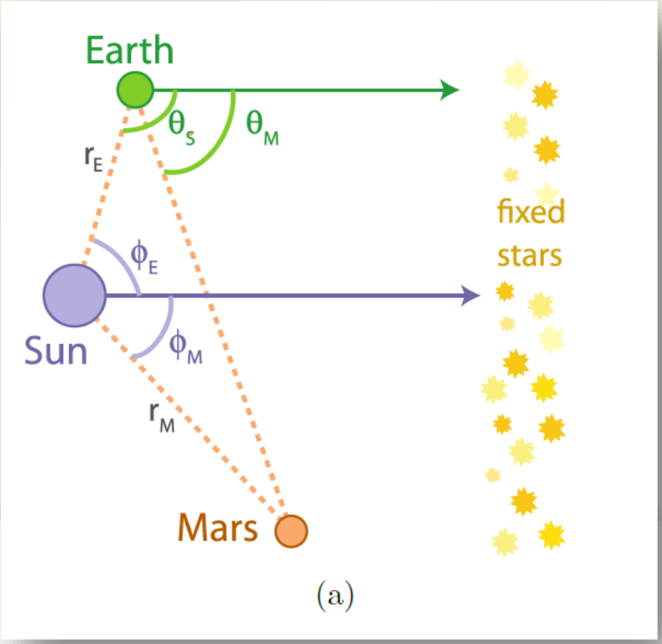
Solar system

- coordinate selection

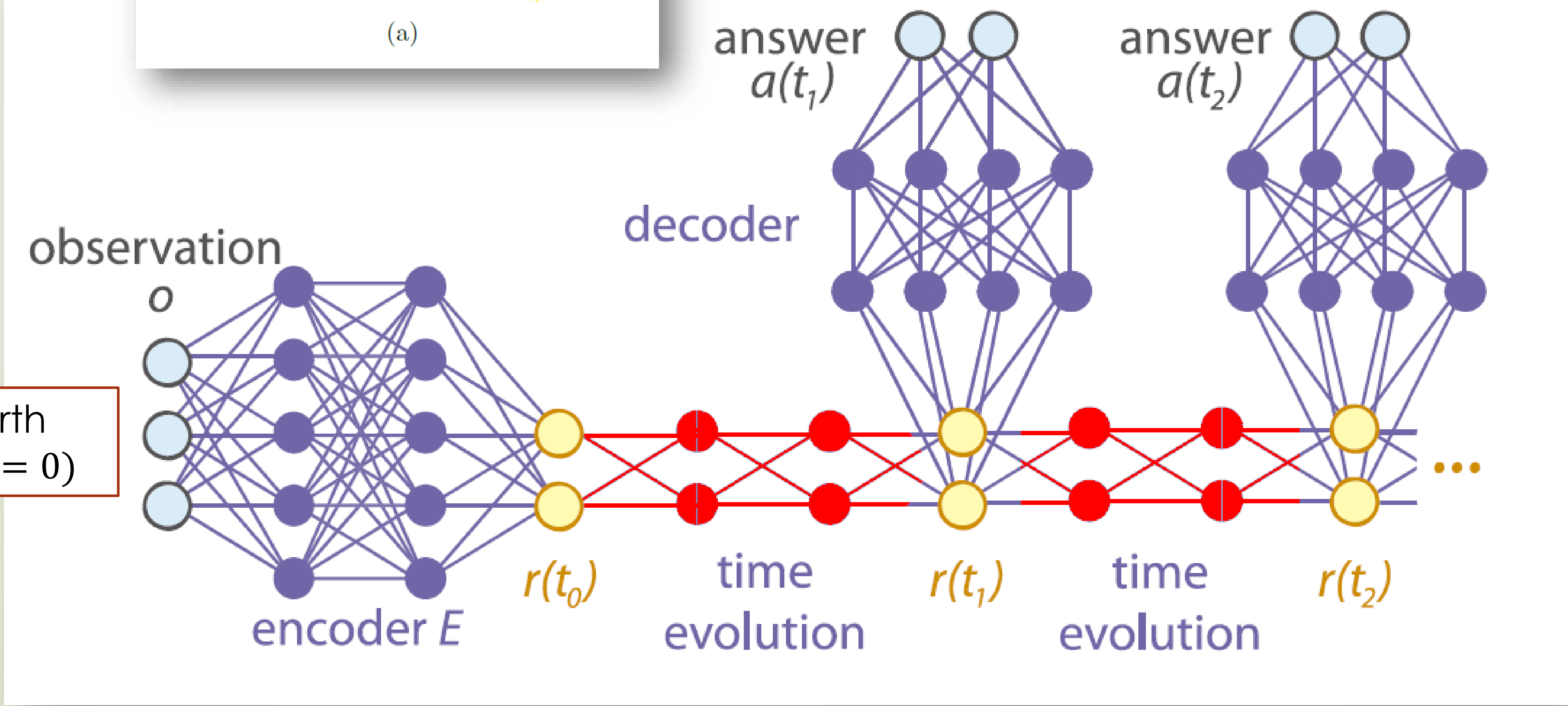


(a)

Solar system



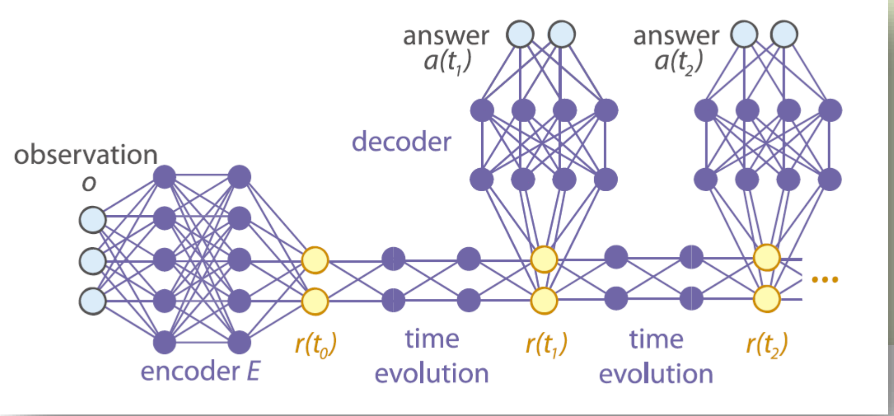
Predicted angles
$\theta_{sun}(t), \theta_{mars}(t)$

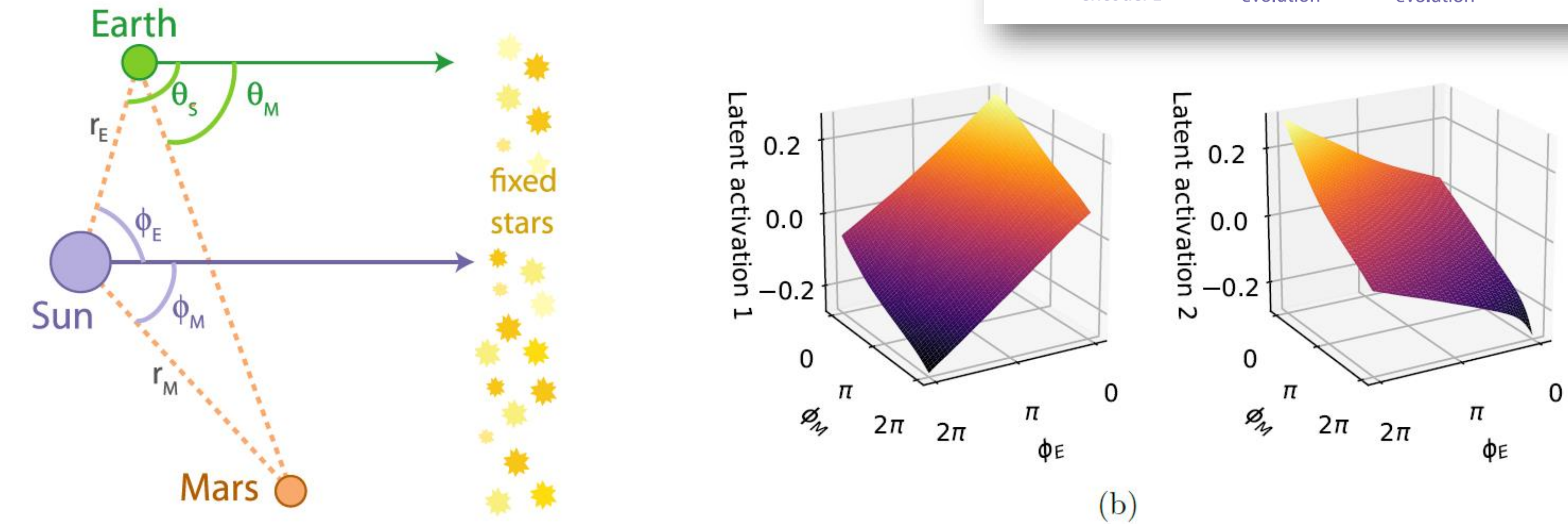


Initial angles from Earth
$\theta_{sun}(t=0), \theta_{mars}(t=0)$

Key Findings: Solar System

- *Storing linear combinations of* $\varphi_{earth}, \varphi_{mars}$
- RMSE = 0.4% (w.r.t. 2π)





(a)

(b)

Discussion

- SciNet:
  - Identifies system parameters in latent neurons
  - Captures constants of motion
  - Counts Degrees of Freedom
  - Finds a preferable coordinate system

Future Work

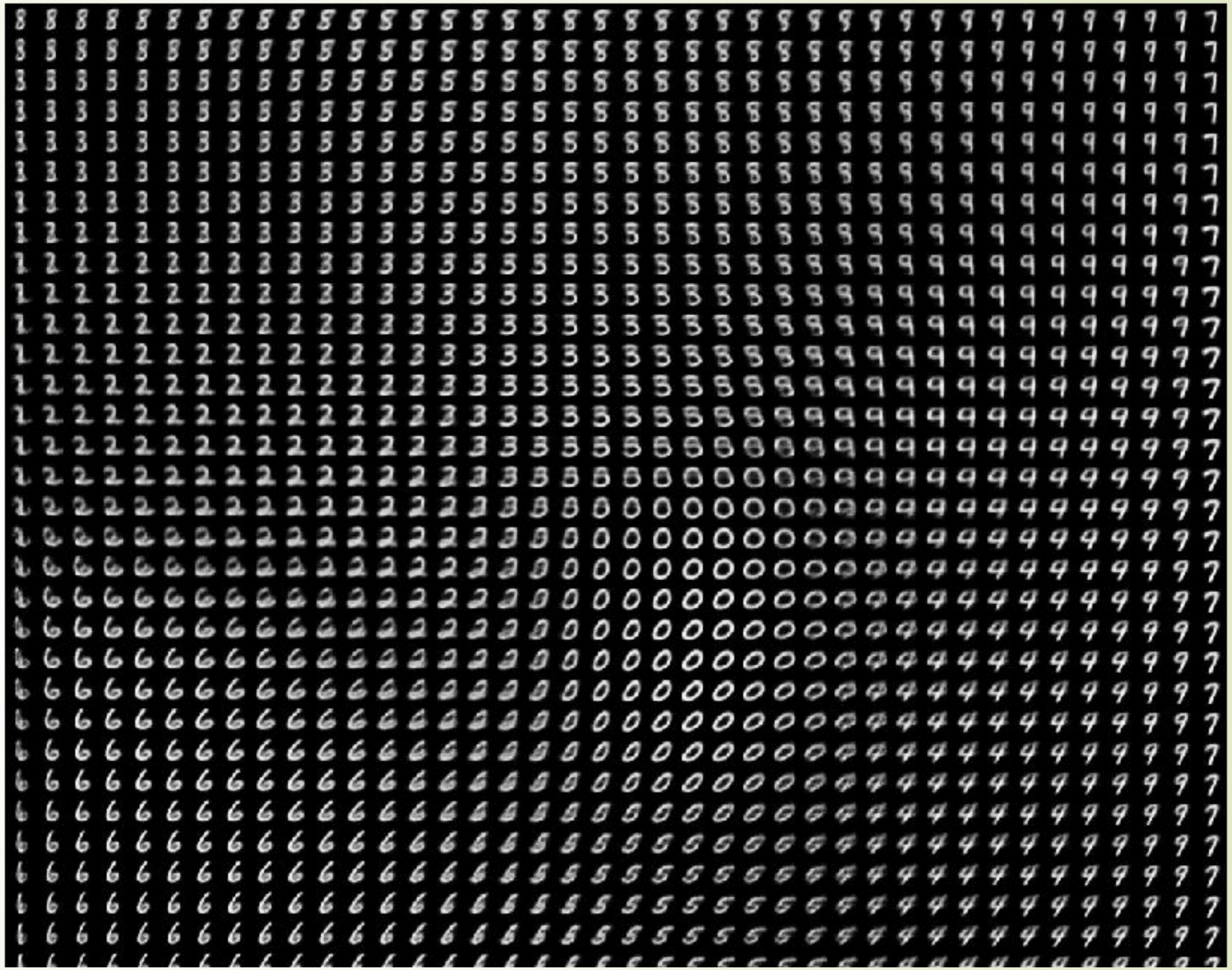
- Strong results on some simple models
  - Validation is simple: Prediction Error
- Determining the d.o.f. needs a cleaner threshold
- Good coordinate systems are valuable but $\varphi(t) = \omega t$ is rare.
  - We have the option to use a more complex time-evolving RNN.
- Interpreting latent variables without comparison model will be hard.

Thank You for the Attention!

Appendix

VAE – latent space

- Result: a dense, continuous latent space.



Variational AutoEncoder: latent space

$$\mathcal{L} = RMSE + \beta \cdot D_{KL}(P || N(\mu = 0, \sigma = 1))$$