

# 機械学習を援用した 適切なフレームの探索

@DL&P 01/19; arXiv:2211.15269(2022)

Yoshihiro Michishita(Riken, CEMS)

# Outline

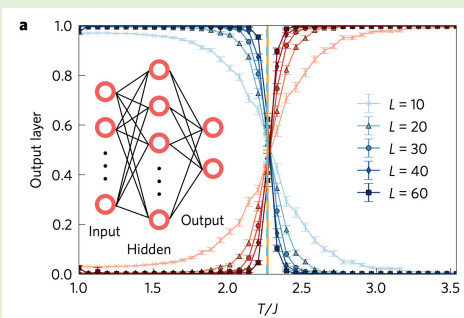
- **Introduction**
  - 機械学習と物理学
  - 物理学における理論解析手法
  - 今回の目標
- **Quick Reivew**
  - 周期駆動系における理論手法(Floquetの手法)
  - 回帰型ニューラルネットワーク(RNN)とその応用
- **Setup & Results**
- **Remarks & Summary**

# Outline

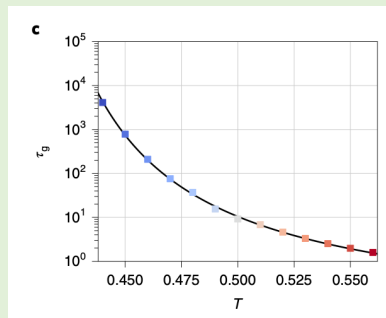
- **Introduction**
  - 機械学習と物理学
  - 物理学における理論解析手法
  - 今回の目標
- **Quick Reivew**
  - 周期駆動系における理論手法(Floquetの手法)
  - 回帰型ニューラルネットワーク(RNN)とその応用
- **Setup & Results**
- **Remarks & Summary**

# Recent application of ML to physics(\*物性屋目線です)

## Detect the phase transition



Ising magnetization  
(Nat. Phys: 13.431(2017))



glass transition  
(Nat. Phys: 10.1038(2020))

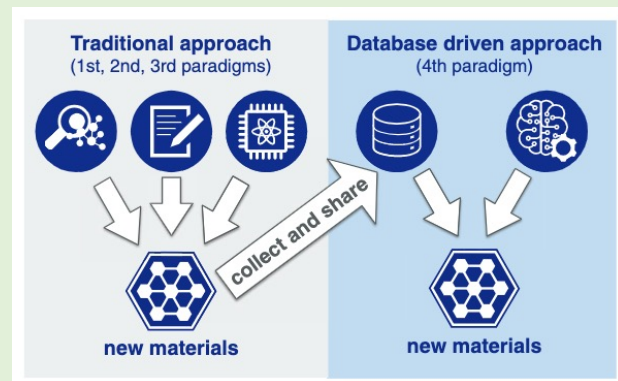
## Calcu. equilibrium or steady state

RBMでスピン系の計算(Science: 335.602-606(2017))  
(PRB: 96.205152(2017))

RBMでSSの計算 (PRB: 99.214306(2019))

RNNを用いた基底状態の表現(PRR: 2.023358(2020))

## Materials Infomatics



(Advanced Science: 6.1900808(2019))

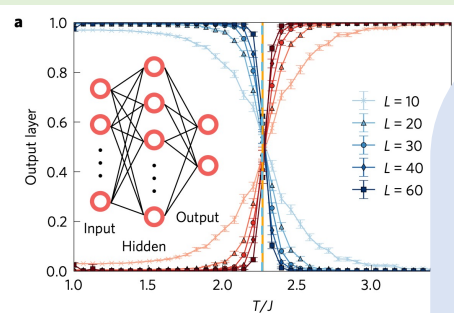
## Remove noise or enhance the accuracy

Gaussian Processを仮定して (IEICE: 10.1587(2010))

NV centerに適用 (Sci. Rep.: 12.13942(2022))

# Recent application of ML to physics (\*物性屋目線です)

## Detect the phase transition



Ising magnetization  
(Nat. Phys: 13.431(2017))

## Calcu. equilibrium or steady state

ce: 335.602-606(2017))

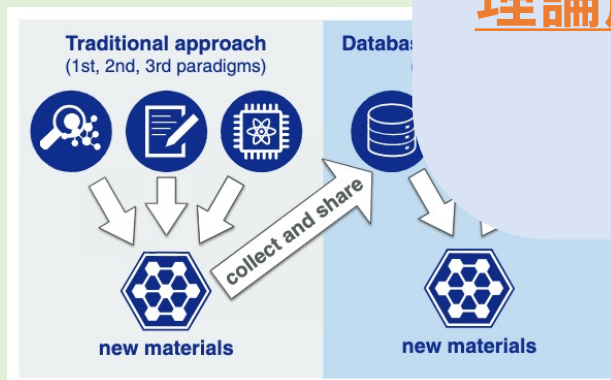
96.205152(2017))

06(2019))

(PRR: 2.023358(2020))

多くの場合、実験の補助や数値計算の手法  
に使用されている。。

## Materials Inform



(Advanced Science: 6.1900808(2019))

理論解析手法の開拓に機械学習を使えないか？

e the accuracy

EICE: 10.1587(2010))

NV centerに適用 (Sci. Rep.: 12.13942(2022))

# How did we physicists develop analytical methods?

## ● Scale separation & Reduction

- Nonlinear system  $\Rightarrow$  reduction
- The Hubbard model (Lattice model)  $\Rightarrow$  Heisenberg model
- Open quantum system  $\Rightarrow$  Markov app., GKSL equation
- Periodic driving system  $\Rightarrow$  high-frequency expansion
- (Renormalization Group  $\Rightarrow$  cutoff scale)
- (DMRG, Tensor network  $\Rightarrow$  SVD & reduction)

# How did we physicists develop analytical methods?

## ● Example 1: Nonlinear system

簡単な例として力学系(多変数の非線形方程式)を考える

$$\frac{dx}{dt} = f(x, y), \quad \frac{dy}{dt} = g(x, y)$$

一般に多変数の非線形方程式は解く事が難しいので、有効的に変数を減らしたい。

上手く減らせる事が出来る成功例として、時間スケールが分離している系がある ( $\|f\| \ll \|g\|$ )

変数  $y$  の緩和のスピードが  $x$  に比べ十分早い時、 $y$  は一瞬で  $x$  を固定した時の定常状態におちる

$$\frac{dy}{dt} = g(x, y) = 0 \quad \Rightarrow \quad y^* = G(x) \quad (g(x, G(x)) = 0)$$

この時、 $x$  の緩和の時間スケールで見れば、 $y$  は  $x$  の関数になっているので

$$\frac{dx}{dt} = f(x, G(x)) = f^*(x) \quad \text{となり、上式と合わせ1変数でダイナミクスを記述出来る}$$

(\* 繰り込みや開放量子系における量子マスター方程式・GKSL方程式もほぼ同じことをする)

# How did we physicists develop analytical methods?

## ● Example 2: Hubbard model (Large-U, Half-filling)

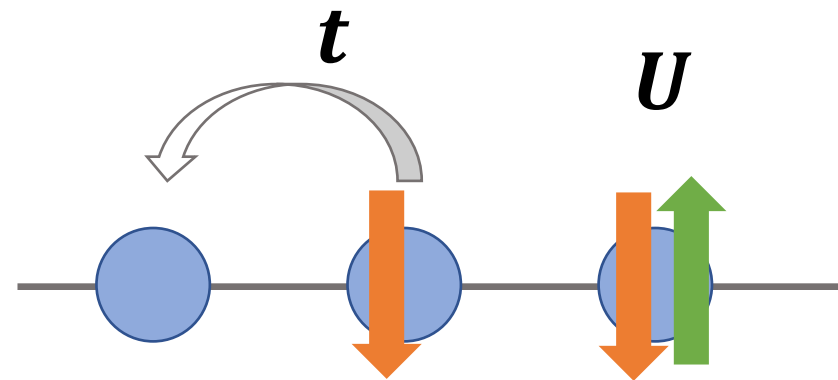
$$H = \sum_{i,\sigma} \{t(c_{i+1,\sigma}^\dagger c_{i,\sigma} + c_{i,\sigma}^\dagger c_{i+1,\sigma}) + \mu n_{i,\sigma} + \frac{U}{2} n_{i,\sigma} n_{i,\bar{\sigma}}\}, \quad n_{i,\sigma} = c_{i,\sigma}^\dagger c_{i,\sigma}$$

上式の  $U \gg t, \mu = -U/2$  の場合を考える。非線形項(相互作用項)を小さくして摂動論で解析したい

$t = 0$  の基底状態を考え、そこから  $t$  についての摂動を行うと、低エネルギー有効ハミルトニアンは

$$H_{eff} = P_0 e^S H e^{-S} P_0 = \sum_{i,\sigma} J (s_i s_{i+1} - \frac{1}{4}) + O\left(\left(\frac{t}{U}\right)^4\right), \quad J = \frac{2t^2}{U}, S = -\frac{t}{U} \sum_i (c_{i+1,\sigma}^\dagger c_{i,\sigma} + c_{i,\sigma}^\dagger c_{i+1,\sigma})$$

$J$  はいま小さいので摂動論で解析が出来る





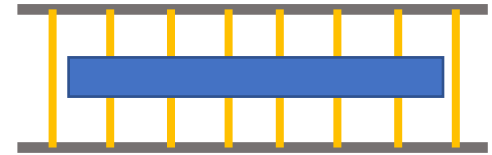
# How did we physicists develop analytical methods?

## ● Example 3: Materials in Cavity (Strong Coupling)

$$H_C = \int dx \psi_x^\dagger \left[ \frac{(-i\partial_x - qA)^2}{2m} + V(x) \right] \psi_x + \omega_c a^\dagger a, \quad A = A_0(a^\dagger + a)$$

$$= \int dx \psi_x^\dagger \left[ \frac{(-i\partial_x)^2}{2m} + V(x) \right] \psi_x + g \int dx \psi_x^\dagger \left[ \frac{i\partial_x}{m} \right] \psi_x * B + \Omega b^\dagger b,$$

$$\Omega = \omega_c \sqrt{1 + \frac{2Nq^2 A_0^2}{m\omega_c}}, \quad g = \frac{qA_0}{m} \sqrt{\frac{\omega_c}{\Omega}}, \quad b^\dagger + b = \sqrt{\frac{\Omega}{\omega_c}} (a^\dagger + a)$$



$g$ の大きな極限を考える。以下のユニタリ変換を用いる

$$U = \exp \left[ -i\xi_g \pi \left( \int dx \psi_x^\dagger \left[ \frac{-i\partial_x}{m} \right] \psi_x \right) \right], \quad \pi = i(b^\dagger - b), \quad \xi_g \rightarrow \propto g^{-\frac{1}{2}} (g \rightarrow \infty)$$

すると

$$H_U = \int dx \psi_x^\dagger \left[ \frac{(-i\partial_x)^2}{2m} + V(x + \xi_g \pi) \right] \psi_x + \Omega b^\dagger b$$

と出来、強結合領域でも解析が出来る。

# How did we physicists develop analytical methods?

## ● Example 4: Floquet System (周期駆動系, High-frequency reg.)

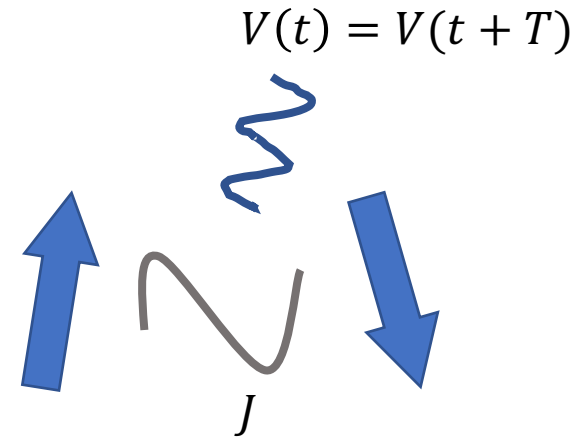
$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t), \quad i \frac{d}{dt} |\psi(t)\rangle = \hat{H}(t) |\psi(t)\rangle, \quad V(t) = V(t+T), \quad \Omega = \frac{2\pi}{T}$$

$$\begin{aligned} i \frac{d}{dt} |\tilde{\psi}(t)\rangle &= i \frac{d}{dt} \hat{U}(t) |\psi(t)\rangle \\ &= \hat{U}(t) (\hat{H}(t) - i \frac{d}{dt}) \hat{U}^\dagger(t) |\tilde{\psi}(t)\rangle = \hat{H}_r(t) |\tilde{\psi}(t)\rangle \end{aligned}$$

これに対しては、良いUの与え方が知られていて、それを用いれば

$$H_r(t) = H_r^{(n)} + O(1/\Omega^{n+1}, t)$$

なる回転座標系での有効ハミルトニアンが与えられ、しかも実際の時間発展との差は十分小さい事が知られている。



# How did we physicists develop analytical methods?

- ここまでの例で言いたかったこと

- 例1: スケールの分離がある場合、変数の消去(reduction)が出来る場合がある

# How did we physicists develop analytical methods?

## ●ここまでの例で言いたかったこと

- 例1: スケールの分離がある場合、変数の消去(reduction)が出来る場合がある
- 例2-4: スケールの分離がある時に、reductionや摂動論を用いたい場合、  
一般にはそれが出来るframeにユニタリ変換や射影を行う必要がある

# How did we physicists develop analytical methods?

## ●ここまでの例で言いたかったこと

- 例1:スケールの分離がある場合、変数の消去(reduction)が出来る場合がある
- 例2-4:スケールの分離がある時に、reductionや摂動論を用いたい場合、  
一般にはそれが出来るframeにユニタリ変換や射影を行う必要がある

なので、物理の(手でできる範囲の)理論解析とは、

1. (scaleの分離がある場合に)摂動論やreductionが出来る良いframeを見つける
2. (妥当な近似として)摂動論やreductionを実行し有効モデルを作る
3. 個々の物性を得られたモデルで解析

# How did we physicists develop analytical methods?

## ●ここまでの例で言いたかったこと

- 例1:スケールの分離がある場合、変数の消去(reduction)が出来る場合がある
- 例2-4:スケールの分離がある時に、reductionや摂動論を用いたい場合、  
一般にはそれが出来るframeにユニタリ変換や射影を行う必要がある

なので、物理の(手でできる範囲の)理論解析とは、

ここを機械学習にやらせたい

1. (scaleの分離がある場合に)摂動論やreductionが出来る良いframeを見つける
2. (妥当な近似として)摂動論やreductionを実行し有効モデルを作る
3. 個々の物性を得られたモデルで解析

# How did we physicists develop analytical methods?

- 今回の目的

機械学習を援用して理論解析手法を開拓する

# How did we physicists develop analytical methods?

- 今回の目的

機械学習を援用して理論解析手法を開拓する



ある一般的なモデルのあるパラメータ領域において  
適切なframeを機械学習を援用して見つける



# How did we physicists develop analytical methods?

## ●今回の目的

機械学習を援用して理論解析手法を開拓する



ある一般的なモデルのあるパラメータ領域において  
適切なframeを機械学習を援用して見つける



そもそもスケールの分離が存在するのか一見分からない場合に  
(implicitなscale分離がある場合)適切なFrameを機械学習で見つける

# How did we physicists develop analytical methods?

## ●今回の目的

機械学習を援用して理論解析手法を開拓する



ある一般的なモデルのあるパラメータ領域において  
適切なframeを機械学習を援用して見つける



そもそもスケールの分離が存在するのか一見分からない場合に  
(implicitなscale分離がある場合)適切なFrameを機械学習で見つける

# Outline

- **Introduction**
  - 機械学習と物理学
  - 物理学における理論解析手法
  - 今回の目標
- **Quick Review**
  - 周期駆動系における理論手法(Floquetの手法)
  - 回帰型ニューラルネットワーク(RNN)とその応用
- **Setup & Results**
- **Remarks & Summary**

# Quick review of Floquet theory & RF

## ➤ Formalism

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t) \quad i \frac{d}{dt} |\psi(t)\rangle = \hat{H}(t) |\psi(t)\rangle \quad \hat{U}(t) = \exp[i\hat{K}(t)]$$

$$i \frac{d}{dt} |\tilde{\psi}(t)\rangle = i \frac{d}{dt} \hat{U}(t) |\psi(t)\rangle = \hat{H}_r(t) |\tilde{\psi}(t)\rangle \quad \hat{H}_r(t) = \hat{U}(t) (\hat{H}(t) - i\partial_t) \hat{U}^\dagger(t)$$

$\hat{H}(t) = \hat{H}(t+T)$  のとき、 $\hat{U}_F(t) = \hat{U}_F(t+T)$ ,  $\hat{H}_r(t) = \hat{H}_F$  を満たす  $U_F(t)$  が存在

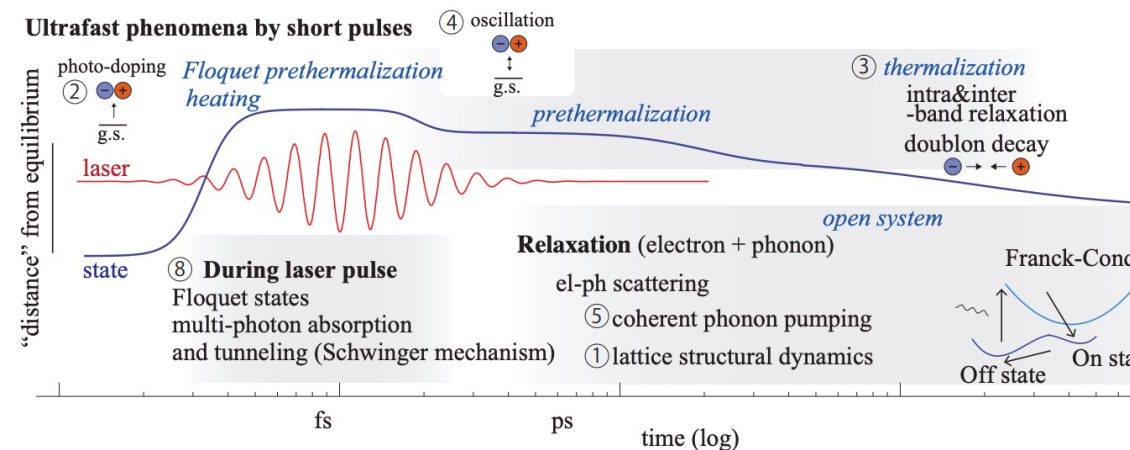
さらに高周波領域  $\|H_0\| \ll \Omega$  の時、 $U_F(t), H_F$  を  $1/\Omega$  展開する方法がある (van-Vleck, Floquet-Magnus)

$$H_r(t) = H_F^{(n)} + O\left(\frac{1}{\Omega^{n+1}}, t\right)$$

## ➤ Floquet pre-thermalization

Hamiltonian が local かつ、 $t = mT$  の時  
(arXiv:1509.03968(2015))

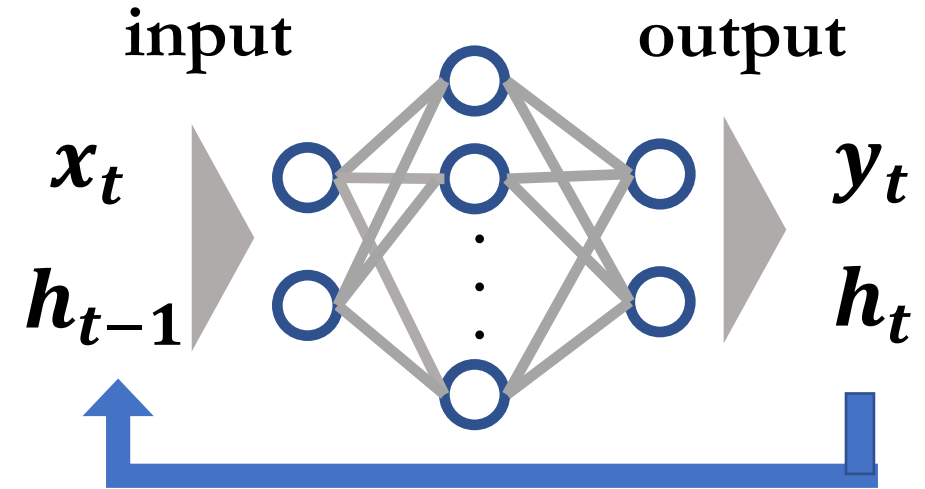
$$\|\mathcal{T} \exp[-i \int ds H(s)] - \exp[-i H_F^{(n)} t]\| \leq \exp[-O(\Omega)] t$$



# Quick review of RNN

## ● RNN

時系列データを学習する際に、あるニューラルネットワークを繰り返し使用する事で、少ないパラメータ数で表現できる。(Google翻訳などで使われている(\*DeepLはCNNらしい))



## ➤ 勾配爆発・勾配消失

$$h_{t+1} = W_h \sigma(h_t) + W_{in} \sigma(x_t) + b$$

$$\frac{\partial l(y_t)}{\partial \theta} = \sum_k \frac{\partial l(y_t)}{\partial h_t} \frac{\partial h_t}{\partial h_k} \frac{\partial h_k}{\partial \theta}$$

\*安直には力学系のアトラクターと  
ベイスンの振る舞いと  
同じ

$$\frac{dh_{t+1}}{dt} = F(h_t) + h_t \approx F(h_t)$$

(arXiv:1211.5063(2013))



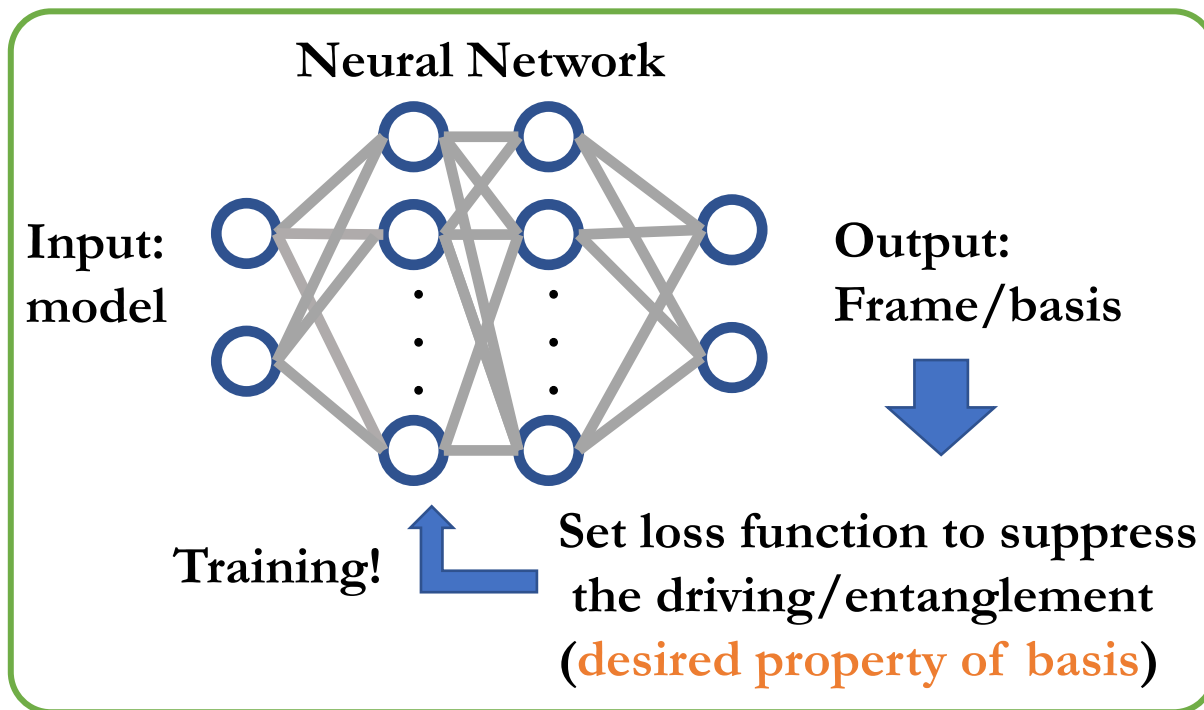
# Outline

- **Introduction**
  - 機械学習と物理学
  - 物理学における理論解析手法
  - 今回の目標
- **Quick Review**
  - 周期駆動系における理論手法(Floquetの手法)
  - 回帰型ニューラルネットワーク(RNN)とその応用
- **Setup & Results**
- **Remarks & Summary**

# What I want to do with machine learning

- 今回提案するフレームワーク

## Step1: Numerically derive the unitary tr.



## Step2: Convert the results into operators

We get (Rotating) frame numerically  
 $\{U(t) = \exp[iK(t)]\}$

We can also calculate the parameter dependence

Physicist

$\hat{U} = \exp[-i\hat{K}(\psi^\dagger, \psi)]?$

Construct the framework of theoretical analysis

If it is not valid, we get the frame numerically in another model and construct another candidate

Check the validity by hand

# Demonstration

## ● First step

- Interacting quantum Two-Spin model under driving

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t)$$

$$\hat{H}_0 = - \sum_{\alpha} (J_{\alpha} \hat{S}_1^{\alpha} \otimes \hat{S}_2^{\alpha} + h_{\alpha} \sum_i \hat{S}_i^{\alpha})$$

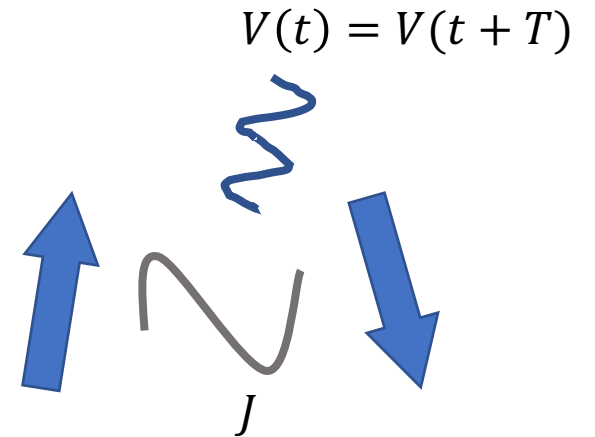
$$\hat{V}(t) = - \sum_{\alpha} \xi_{\alpha} \sin(\Omega t) \sum_i \hat{S}_i^{\alpha}$$

$$\vec{J} = (J_x, J_y = 0, J_z), \quad \vec{h} = (h_x = 0, h_y = 0, h_z), \quad \vec{\xi} = (\xi, 0, 0)$$

$$H_0 = \begin{pmatrix} -J_z - 2h_z & 0 & 0 & -J_x \\ 0 & J_z & -J_x & 0 \\ 0 & -J_x & J_z & 0 \\ -J_x & 0 & 0 & -J_z + 2h_z \end{pmatrix}$$

$$V = \begin{pmatrix} 0 & -\xi & -\xi & 0 \\ -\xi & 0 & 0 & -\xi \\ -\xi & 0 & 0 & -\xi \\ 0 & -\xi & -\xi & 0 \end{pmatrix}$$

$$\begin{pmatrix} |\uparrow\uparrow\rangle \\ |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \\ |\downarrow\downarrow\rangle \end{pmatrix}$$





# Setup

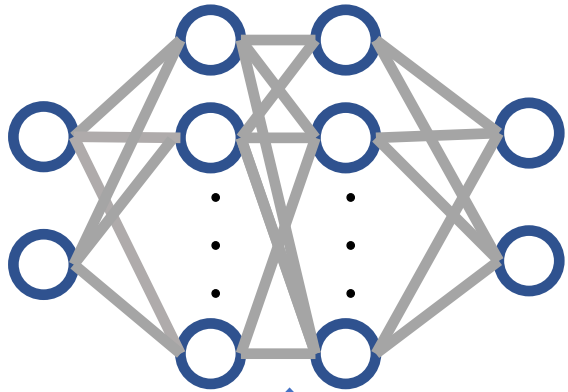
## Model of Machine Learning

Neural Network

$\Omega$

$$H(t) = H(\alpha + \tau)$$

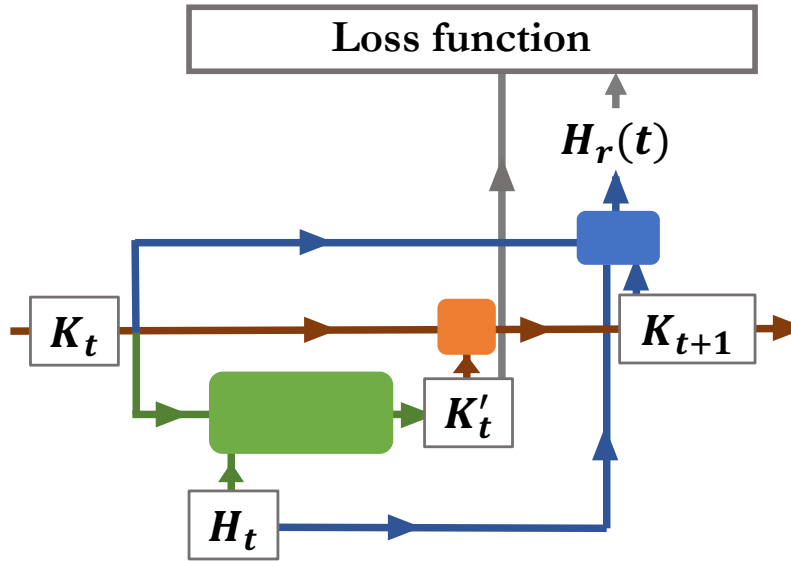
Input:  
 $H(t)$   
 $K(t)$



Output:  
 $\frac{dK(t)}{dt}$



$K(t + \delta t)$   
 $H_r(t)$



■ NN  
■  $K_{t+1} = K_t + K'_t dt$   
■  $H_r(t) = \exp[iK_t] \left( H_t - i \frac{\partial}{\partial t} \right) \exp[-iK_t]$

### Remarks:

Model:

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t)$$

$$\hat{H}_0 = \sum_{\alpha} (J_{\alpha} \hat{S}_1^{\alpha} \otimes \hat{S}_2^{\alpha} + h_{\alpha} \sum_i \hat{S}_1^{\alpha})$$

$$\hat{V}(t) = - \sum_{\alpha} \xi_{\alpha} \sin(\Omega t) \sum_i \hat{S}_i^{\alpha}$$

Formalism:

$$\hat{H}_r(t) = \hat{U}(t) (\hat{H}(t) - i \partial_t) \hat{U}^{\dagger}(t)$$

$$\hat{U}(t) = \exp[i\hat{K}(t)]$$

$$\hat{K}'(t) = \frac{d}{dt} \hat{K}(t)$$

$$l(t) = \sum_n \gamma^n |H_r(t) - H_r(t - n\delta t)| / N_T + \sum_t |K'(t)| / N_T$$

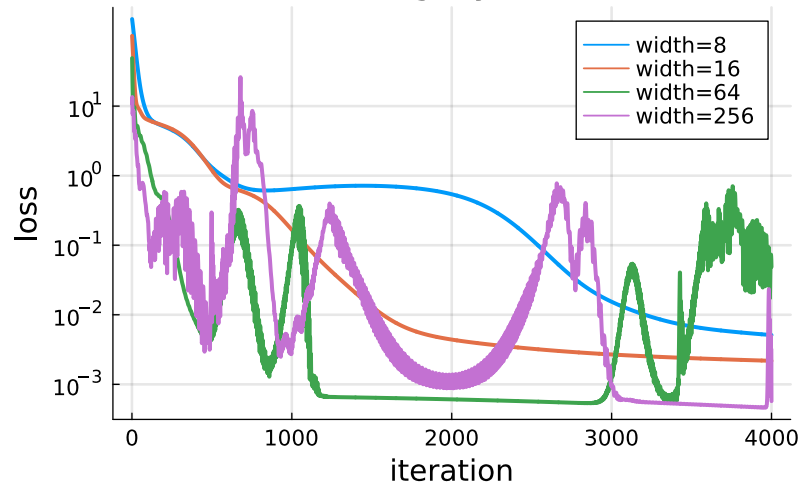
$|M| = \sum_{i,j} |M_{ij}|^2 \rightarrow \text{tr} D^2$   
 $\frac{\text{KL dev. } \mathcal{L} = \mathcal{L}(\delta t)}{N_T} = \frac{\mathcal{I}}{\delta t}$

Transformer

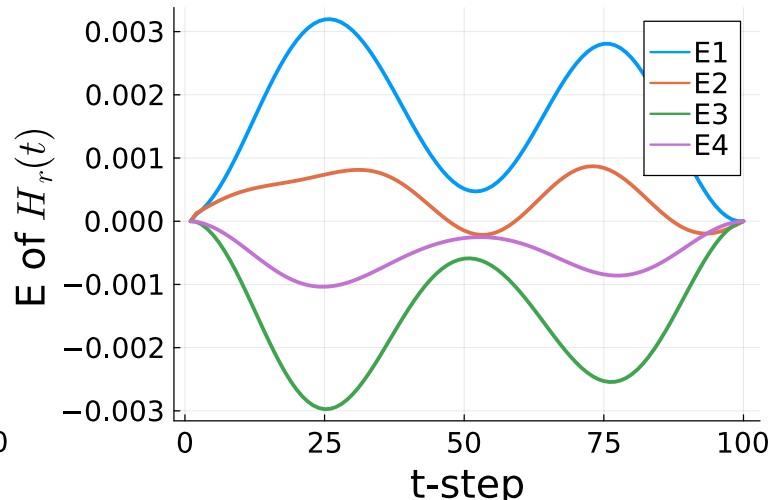
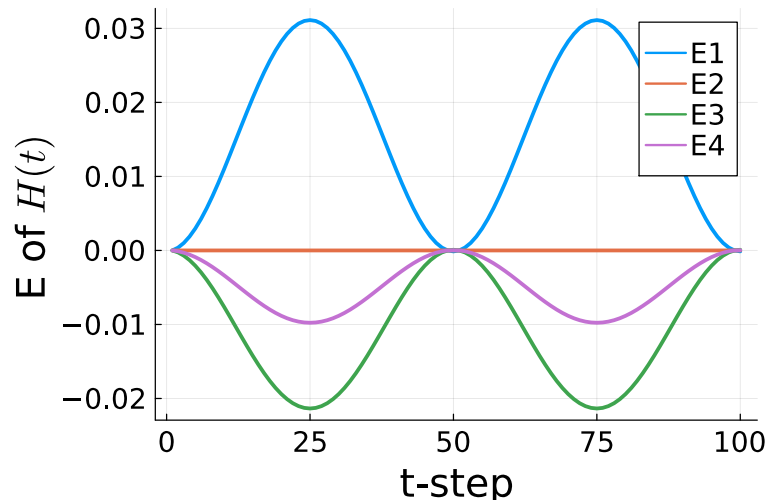
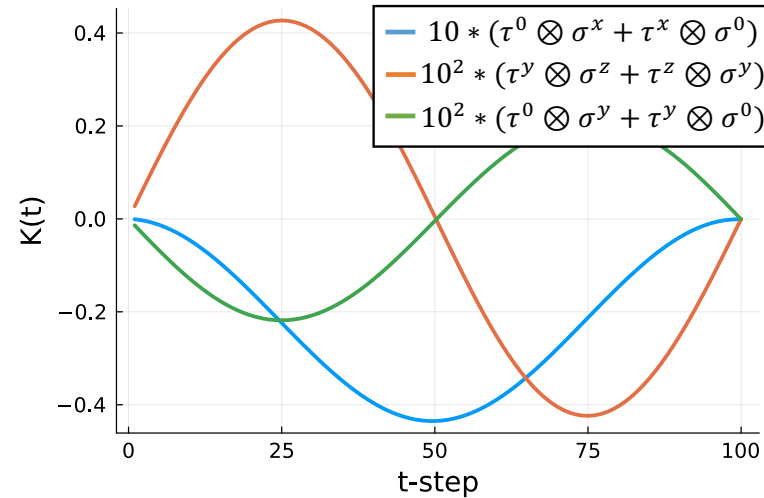
# Step 1: Numerically derive the RF with ML

## High-Frequency regime

Learning dynamics



(a)  $(\xi, J_z, J_x, h_z, \Omega) = (0.2, 1.0, 0.7, 0.5, 10)$



Remarks:

**Model:**

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t)$$

$$\hat{H}_0 = \sum_{\alpha} (J_{\alpha} \hat{S}_1^{\alpha} \otimes \hat{S}_2^{\alpha} + h_{\alpha} \sum_i \hat{S}_1^{\alpha})$$

$$\hat{V}(t) = - \sum_{\alpha} \xi_{\alpha} \sin(\Omega t) \sum_i \hat{S}_i^{\alpha}$$

**Formalism:**

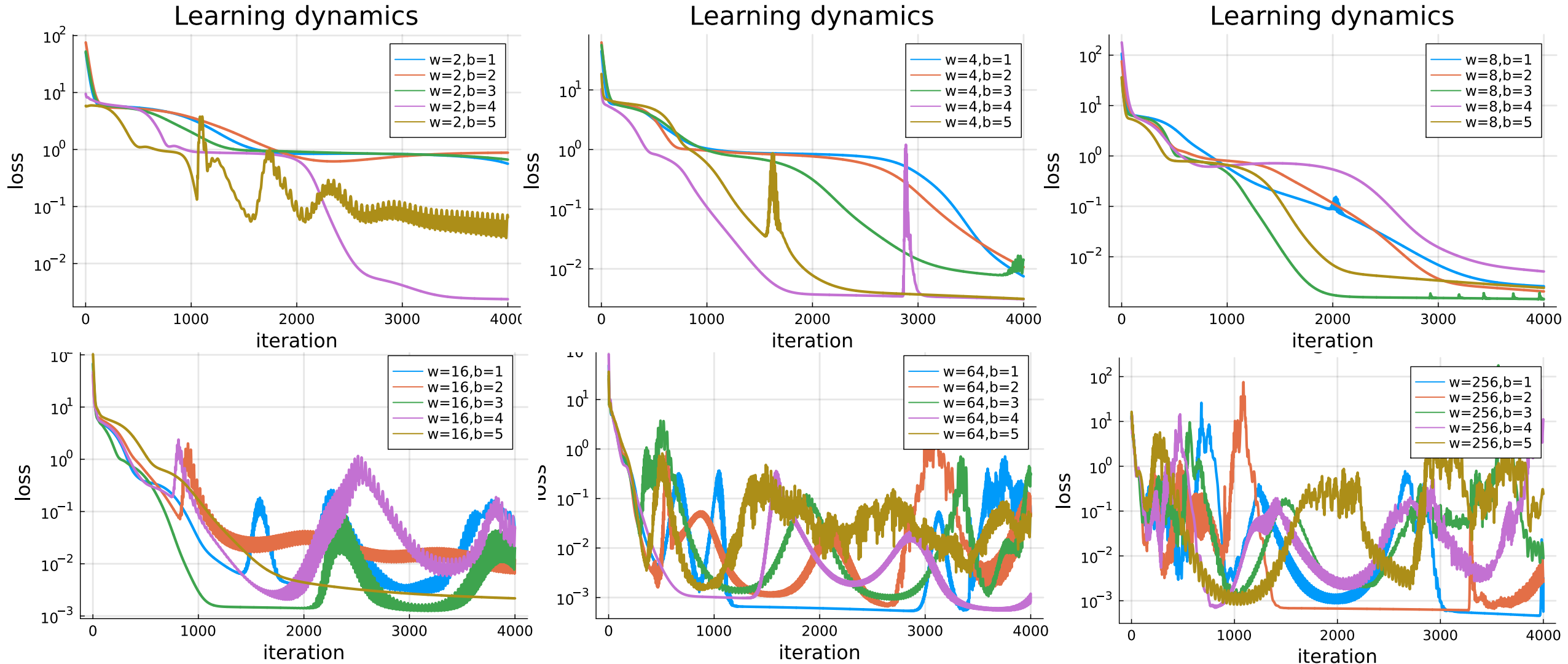
$$\hat{H}_r(t) = \hat{U}(t)(\hat{H}(t) - i\partial_t)\hat{U}^+(t)$$

$$\hat{U}(t) = \exp[i\hat{K}(t)]$$

$$\hat{K}'(t) = \frac{d}{dt}\hat{K}(t)$$

# Step 1: Numerically derive the RF with ML

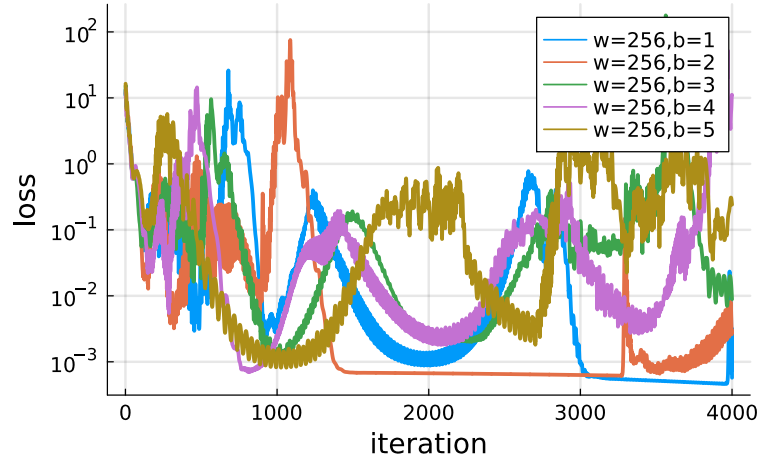
- hidden layer width dependence of learning dynamics *Adabelief(?)*



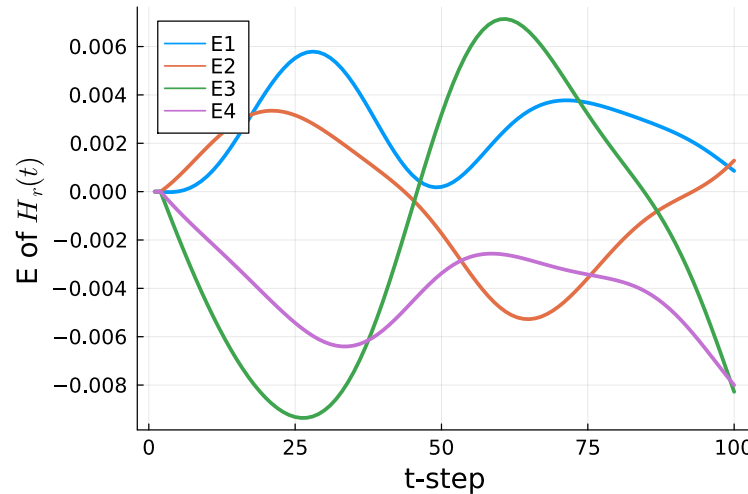
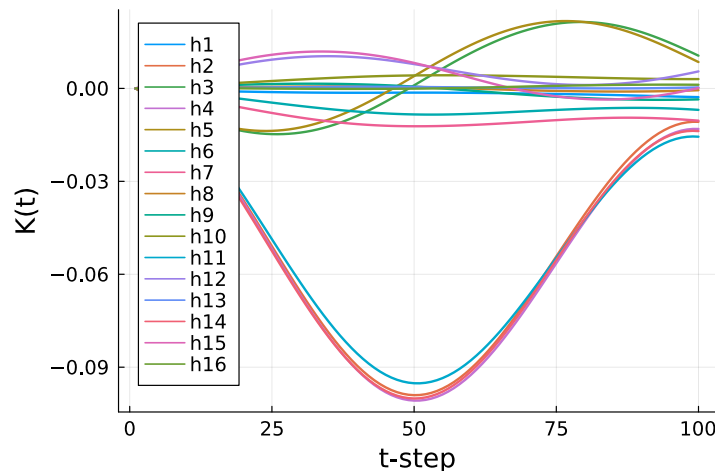
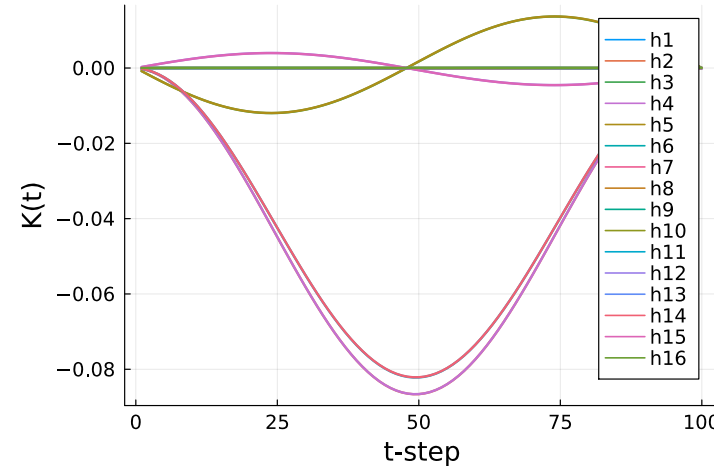
# Step 1: Numerically derive the RF with ML

- If learning doesn't work well,

Learning dynamics



Success case



Remarks:

Model:

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t)$$

$$\hat{H}_0 = \sum_{\alpha} (J_{\alpha} \hat{S}_1^{\alpha} \otimes \hat{S}_2^{\alpha} + h_{\alpha} \sum_i \hat{S}_1^{\alpha})$$

$$\hat{V}(t) = - \sum_{\alpha} \xi_{\alpha} \sin(\Omega t) \sum_i \hat{S}_i^{\alpha}$$

Formalism:

$$\hat{H}_r(t) = \hat{U}(t)(\hat{H}(t) - i\partial_t)\hat{U}^+(t)$$

$$\hat{U}(t) = \exp[i\hat{K}(t)]$$

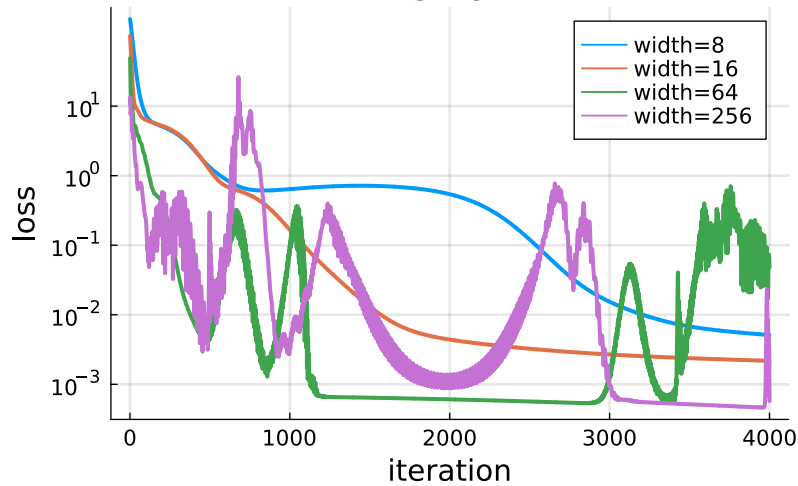
ここまでの主張

Hidden layerの幅を増やし過ぎると学習が上手くいかない。  
減らし過ぎると初期状態によっては上手く行くが安定しない。  
今回の場合はw=8くらいが一番安定

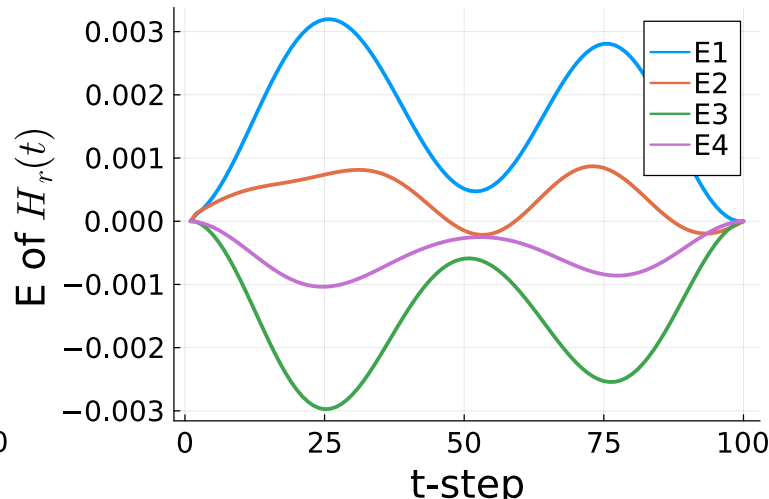
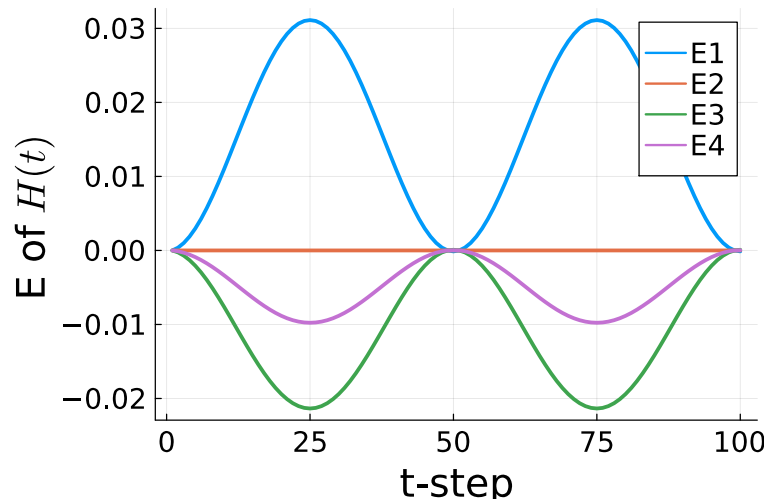
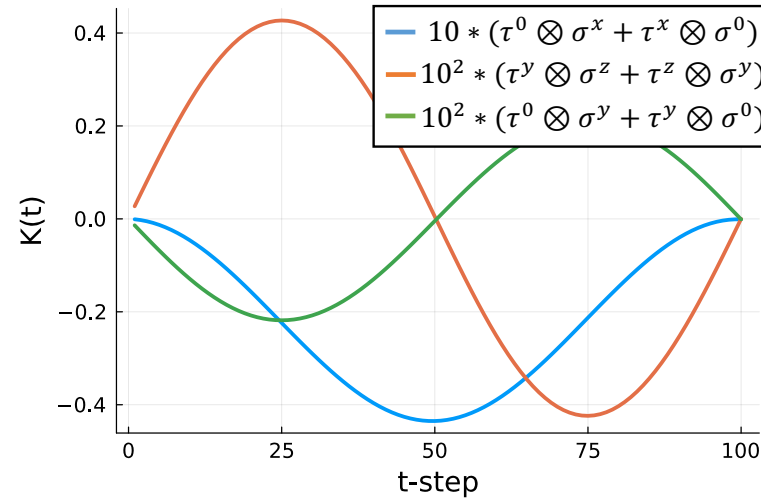
# Step 1: Numerically derive the RF with ML

## High-Frequency regime

Learning dynamics



(a)  $(\xi, J_z, J_x, h_z, \Omega) = (0.2, 1.0, 0.7, 0.5, 10)$



### Remarks:

#### Model:

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t)$$

$$\hat{H}_0 = \sum_{\alpha} (J_{\alpha} \hat{S}_1^{\alpha} \otimes \hat{S}_2^{\alpha} + h_{\alpha} \sum_i \hat{S}_1^{\alpha})$$

$$\hat{V}(t) = - \sum_{\alpha} \xi_{\alpha} \sin(\Omega t) \sum_i \hat{S}_i^{\alpha}$$

#### Formalism:

$$\hat{H}_r(t) = \hat{U}(t)(\hat{H}(t) - i\partial_t)\hat{U}^+(t)$$

$$\hat{U}(t) = \exp[i\hat{K}(t)]$$

#### ここまでの主張

Hidden layerの幅を増やし過ぎると学習が上手くいかない。  
減らし過ぎると初期状態によっては上手く行くが安定しない。  
今回の場合はw=8くらいが一番安定

# Step 2: Convert the result to operator form

## Parameter dependence & translate it to operator form

### Remarks:

#### Model:

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t)$$

$$\hat{H}_0 = \sum_{\alpha} (J_{\alpha} \hat{S}_1^{\alpha} \otimes \hat{S}_2^{\alpha} + h_{\alpha} \sum_i \hat{S}_i^{\alpha})$$

$$\hat{V}(t) = - \sum_{\alpha} \xi_{\alpha} \sin(\Omega t) \sum_i \hat{S}_i^{\alpha}$$

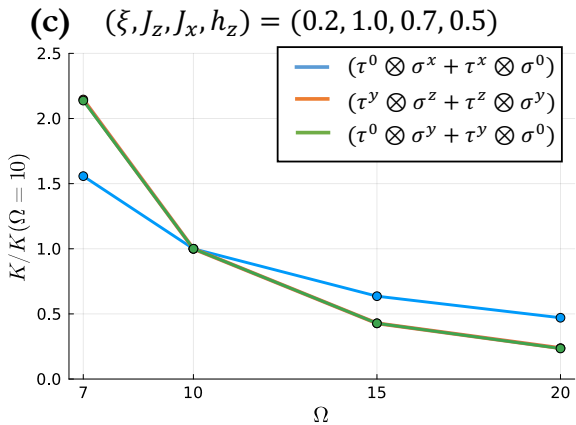
#### Formalism:

$$\hat{H}_r(t) = \hat{U}(t)(\hat{H}(t) - i\partial_t)\hat{U}^+(t)$$

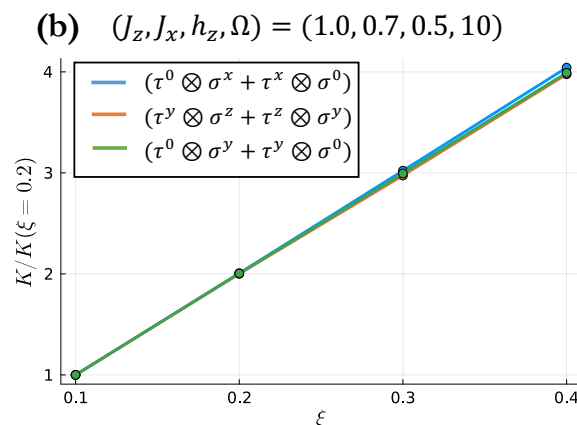
$$\hat{U}(t) = \exp[i\hat{K}(t)]$$

#### ここまでの主張

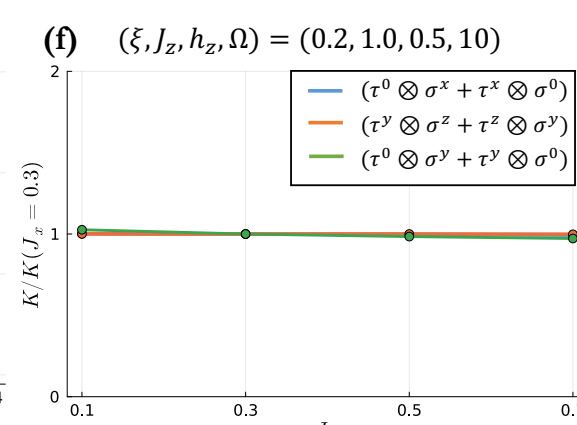
Hidden layerの幅を増やし過ぎると学習が上手くいかない。  
減らし過ぎると初期状態によっては上手く行くが安定しない。  
今回の場合はw=8くらいが一番安定



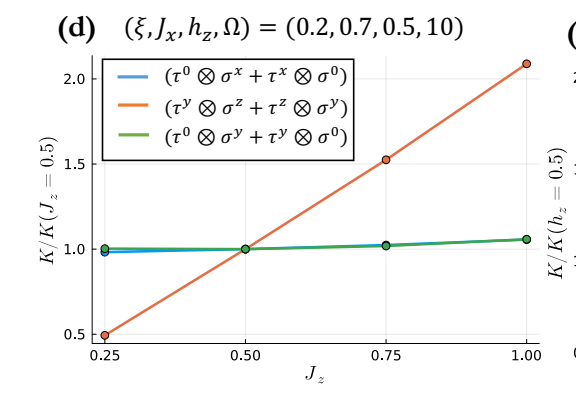
→ bl  $\propto \Omega^{-1}$ , or, gr  $\propto \Omega^{-2}$



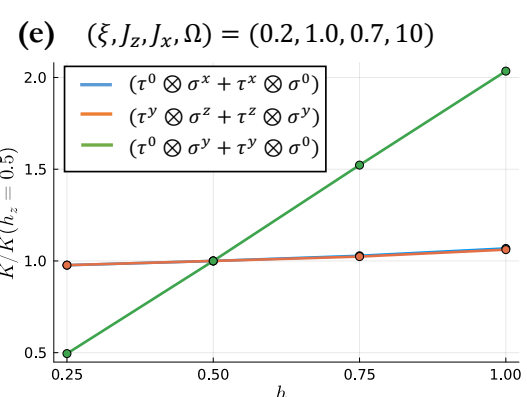
→ bl, or, gr  $\propto \xi^1$



→ bl, or, gr  $\propto J_x^0$



bl, gr  $\propto J_z^0$ , or  $\propto J_z^1$



bl, or  $\propto h_z^0$ , gr  $\propto h_z^1$

$$\hat{K}(t) = \xi \left[ \frac{1}{\Omega} (1 - \frac{1}{\Omega} \frac{dV}{dt}) + \frac{1}{\Omega} (1 - \cos(\Omega t)) (\tau^0 \otimes \sigma^x + \tau^x \otimes \sigma^0) + \frac{2J_z}{\Omega^2} \sin(\Omega t) (\tau^y \otimes \sigma^z + \tau^z \otimes \sigma^y) + \frac{2h_z}{\Omega^2} \sin(\Omega t) (\tau^0 \otimes \sigma^y + \tau^y \otimes \sigma^0) \right]$$

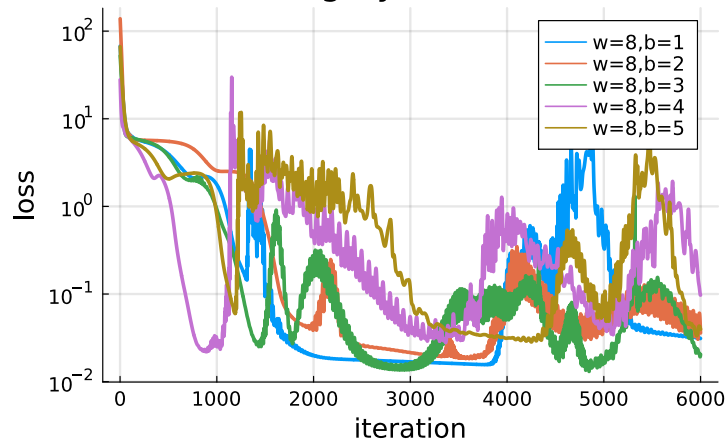
$$\hat{K}(t) = \int dt' \{ \hat{V}(t') - i \int dt'' [\hat{V}(t''), \hat{H}_0] \}$$

# Results

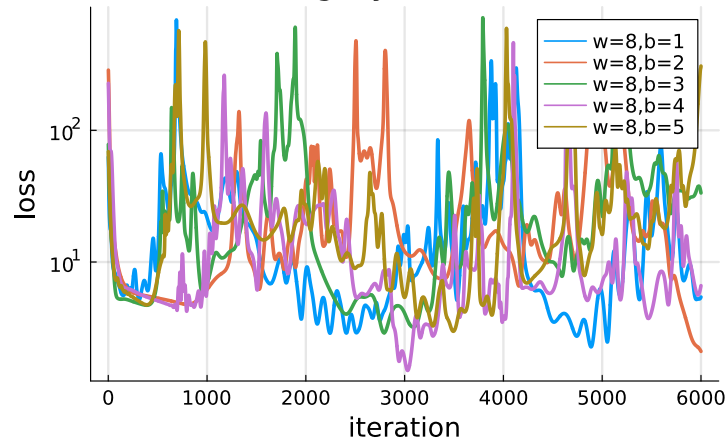
## ● Resonant Drive

$$\Omega = 2, \xi = 0.2, J_z = 1.0, h_z = 0.5$$

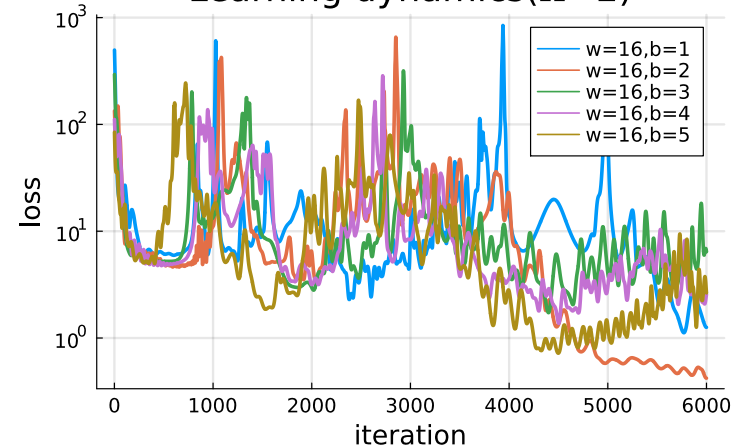
Learning dynamics( $\Omega=5$ )



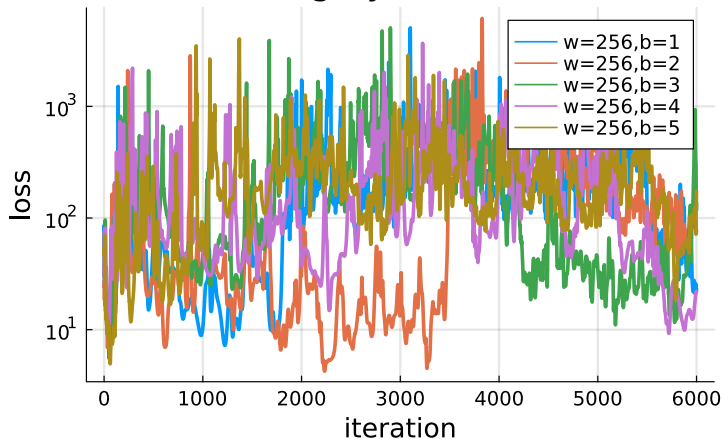
Learning dynamics( $\Omega=2$ )



Learning dynamics( $\Omega=2$ )



Learning dynamics( $\Omega=2$ )



### Remarks:

#### Model:

$$\widehat{H}_0 = \sum_{\alpha} (J_{\alpha} \widehat{S}_1^{\alpha} \otimes \widehat{S}_2^{\alpha} + h_{\alpha} \sum_i \widehat{S}_i^{\alpha})$$

$$\widehat{V}(t) = - \sum_{\alpha} \xi_{\alpha} \sin(\Omega t) \sum_i \widehat{S}_i^{\alpha}$$

#### Formalism:

$$\widehat{U}(t) = \exp[i\widehat{K}(t)]$$

#### ここまでの主張

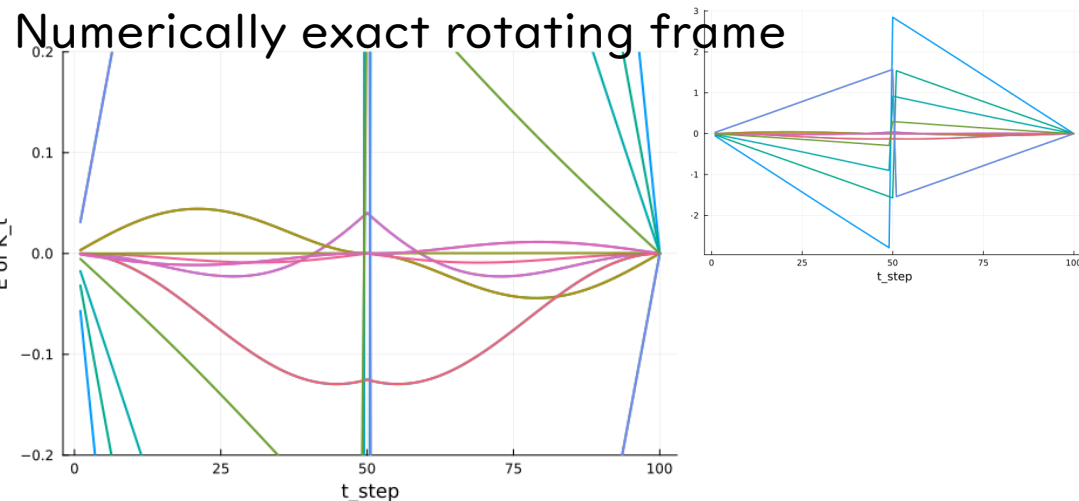
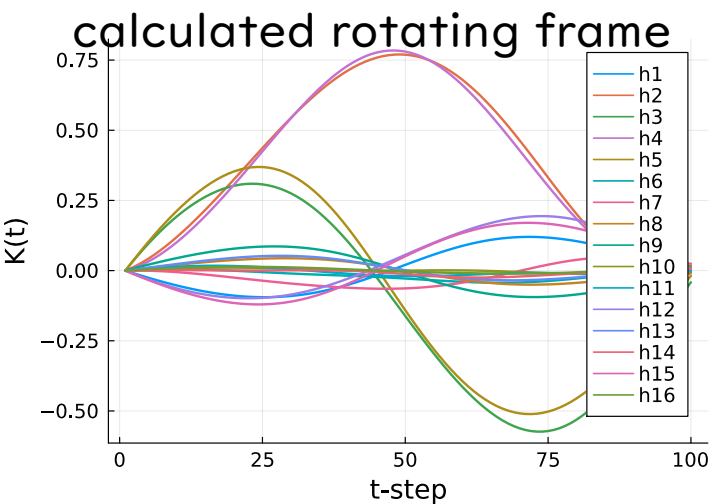
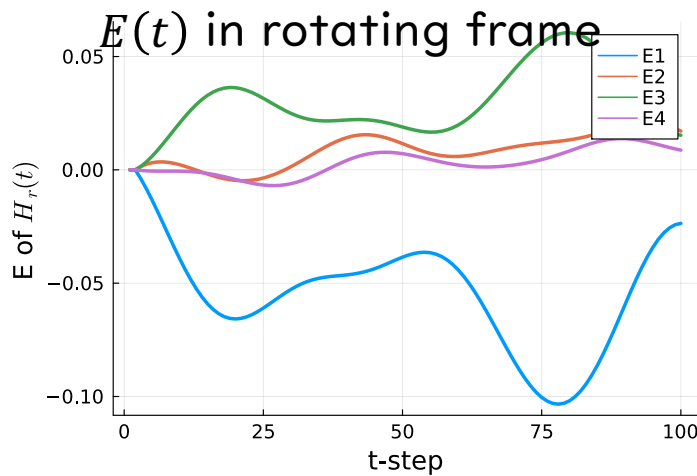
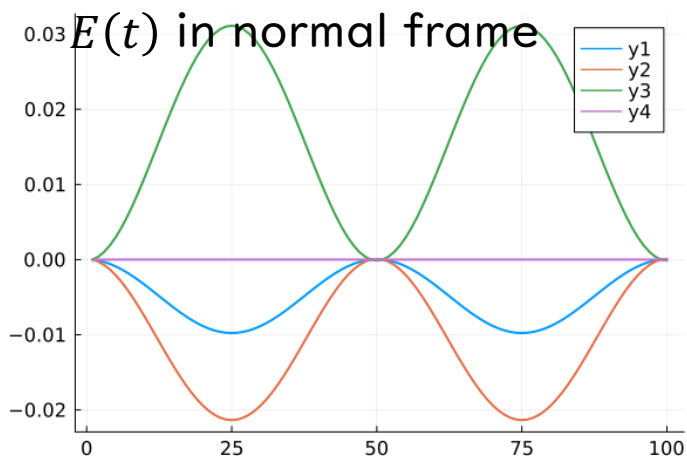
Hidden layerの幅を増やし過ぎると学習が上手くいかない。  
減らし過ぎると初期状態によっては上手く行くが安定しない。  
今回の場合は $w=8$ くらいが一番安定

# Results

## ● Resonant Drive

$$\Omega = 2, \xi = 0.2, J_z = 1.0, h_z = 0.5$$

(w=16, b=2の結果)



### Remarks:

#### Model:

$$\widehat{H}_0 = \sum_{\alpha} (J_{\alpha} \widehat{S}_1^{\alpha} \otimes \widehat{S}_2^{\alpha} + h_{\alpha} \sum_i \widehat{S}_1^{\alpha})$$

$$\widehat{V}(t) = - \sum_{\alpha} \xi_{\alpha} \sin(\Omega t) \sum_i \widehat{S}_i^{\alpha}$$

#### Formalism:

$$\widehat{U}(t) = \exp[i\widehat{K}(t)]$$

#### ここまでの主張

Hidden layerの幅を増やし過ぎると学習が上手くいかない。  
減らし過ぎると初期状態によっては上手く行くが安定しない。  
今回の場合はw=8くらいが一番安定



# Remarks & Outlook

## ● ツッコミどころ

- Fourier変換した周波数の空間でやれば普通のNNで出来る
  - ・普通のNNの方が収束性も高いし自然に時間周期性が担保される
  - ・最終的な(僕個人の)目標として、開放系や格子模型に適用したいので、RNNが使えるか試したかった
- 人間が関数形を推察するパートがあるのはどうなの？
  - ・先に人間が関数系をそれなりに指定して、そこに最適化していく方法もあるがそっちの方が微妙な気がした
  - ・次元解析であったり、物理的なpictureを要求すれば、推察することはある程度可能であるはず。(それで推察できないような関数は、そもそも手の解析に向いてない場合が多いはず。)

# Remarks & Outlook

## ● 上手くいかなかった所

### ➤ Resonantの解析で上手く収束しなかった

- ・有限の表現能力で表現できるような良いFrameがない(implicitなscale分離も無い)ということなのか、それともRNNの最適化の仕方をもう少し工夫する必要があるのか現状わからない

### ➤ (今回の場合)汎化性能はない

- ・今回は実はパラメータごとに学習(最適化)を行っている。なのでNNの役割は「有限の表現能力で良く表現できる」という条件を課す事に留まっている。
- ・本当にNNを使う必要があるのか?という疑問も湧いてくるけど、関数形はわかっていないがある程度簡単な形のものを探す為には、これしかない気もする。

# Remarks & Outlook

## ●まとめ

$$H(t) \xrightarrow{V(t)} H_F, \quad H_F \xrightarrow{U(t)} H(t)$$

Floquet engineering.  $H_F \Rightarrow H(t)$

➤ 適切なFrameの探索に知られている場合については成功した

- ・ 自明なスケール分離が存在する場合は成功
- ・ Floquetでたまたま上手くいっただけなのでは？他にもできるのか？

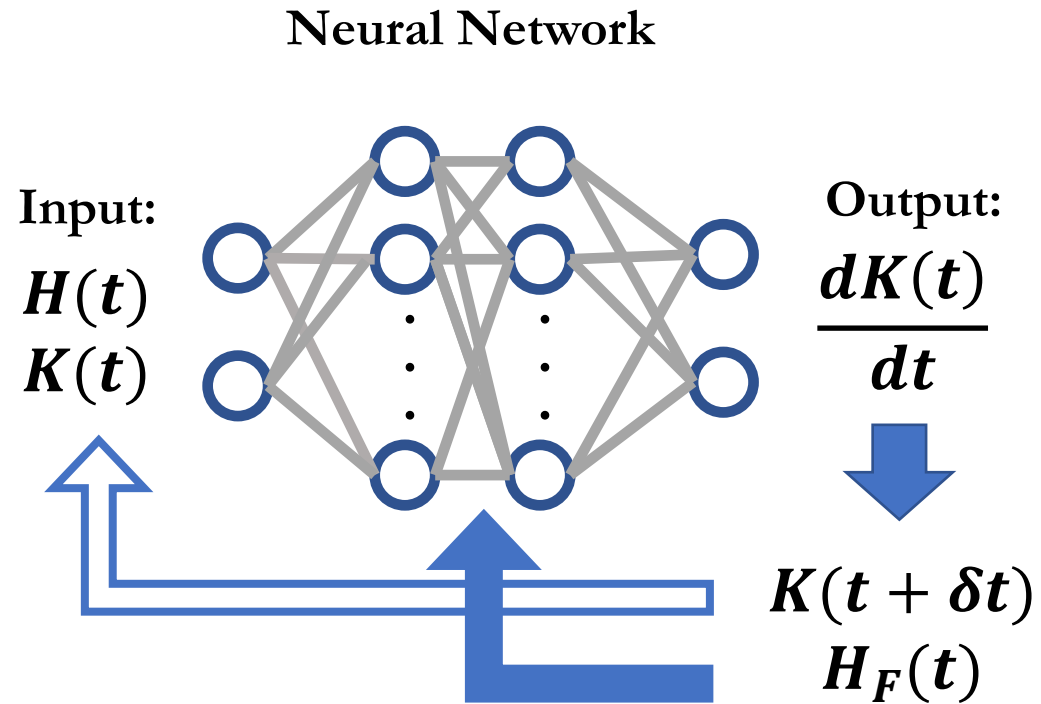
➤ とはいえまだ大分課題も多い

- ・ 上手くいかなかった場合にそれはスケールの分離が存在しないのか、それとも単にRNNの学習が難しいだけなのかがわからない(loss関数の性質が変化するときはあるのかも)
- ・ 人間が推察する部分がある。(ノーヒントで解くよりは簡単だし、次元解析であったり、物理的なpictureを要求すれば、推察することはある程度可能であるはず。(それで推察できないような関数は、そもそも手の解析に向いてない場合が多いはず。))

# Seup

## ● Model of Machine Learning

- Reinforcement Learning(?) / Recurrent Neural Network



$$l(t) = \sum_n \gamma^n |H_F(t) - H_F(t - n\delta t)| / N_T + \delta_{t,T} |H_F(t) - H_F(0)|$$

### Remarks:

Motivation:

Construct a new framework of theoretical analysis

Model:

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t)$$

$$\hat{H}_0 = -\sum_{\alpha} J_{\alpha} \hat{S}_1^{\alpha} \otimes \hat{S}_2^{\alpha} - \sum_i h \cdot \hat{S}_i$$

$$\hat{V}(t) = -\sum_i \tilde{F}(t) \cdot \hat{S}_i$$

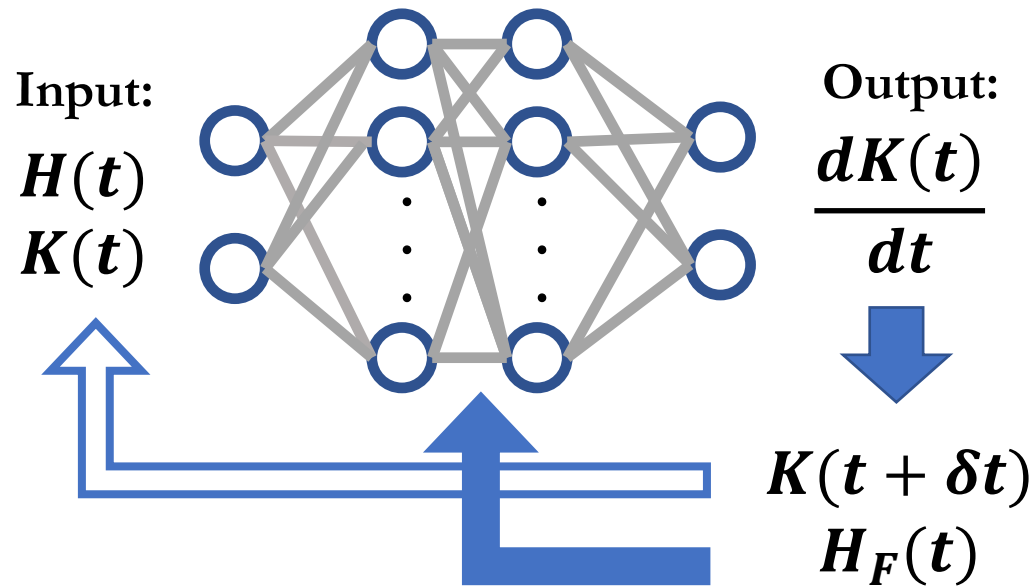
Formalism:

$$e^{-iK^{(n)}(t)} \left[ H(t) - i\hbar \frac{d}{dt} \right] e^{iK^{(n)}t} \simeq \tilde{H}^{(n)}(t)$$

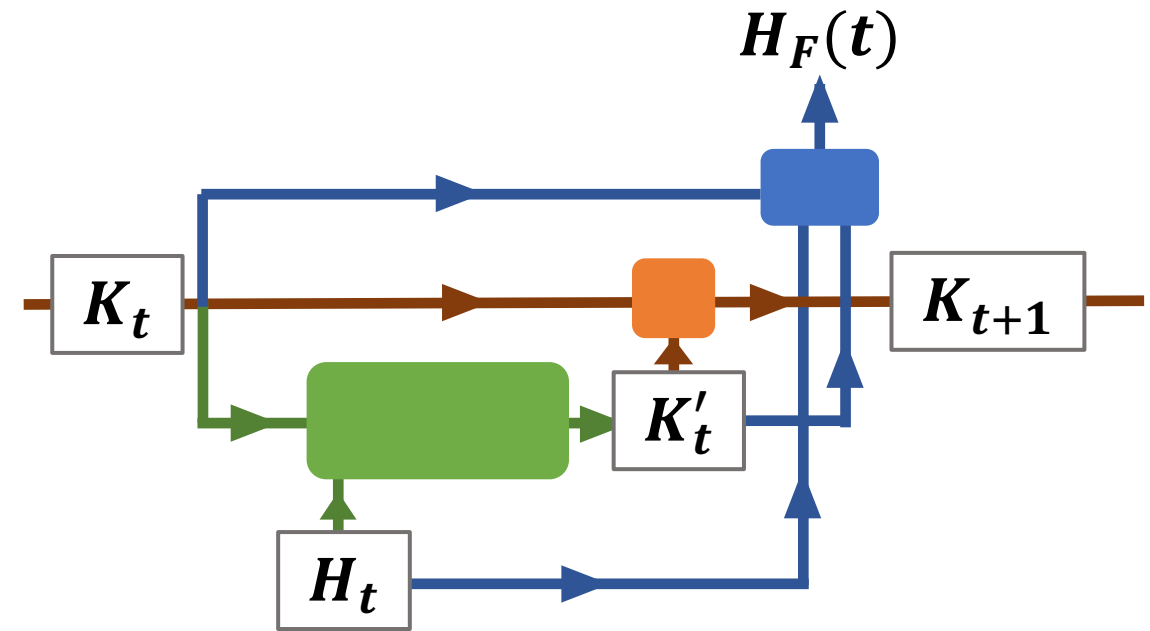
# Setup

- Other Application of RNN to Physics

Neural Network



$$l(t) = \sum_n \gamma^n |H_F(t) - H_F(t - n\delta t)| / N_T + \delta_{t,T} |H_F(t) - H_F(0)|$$



■ NN

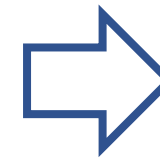
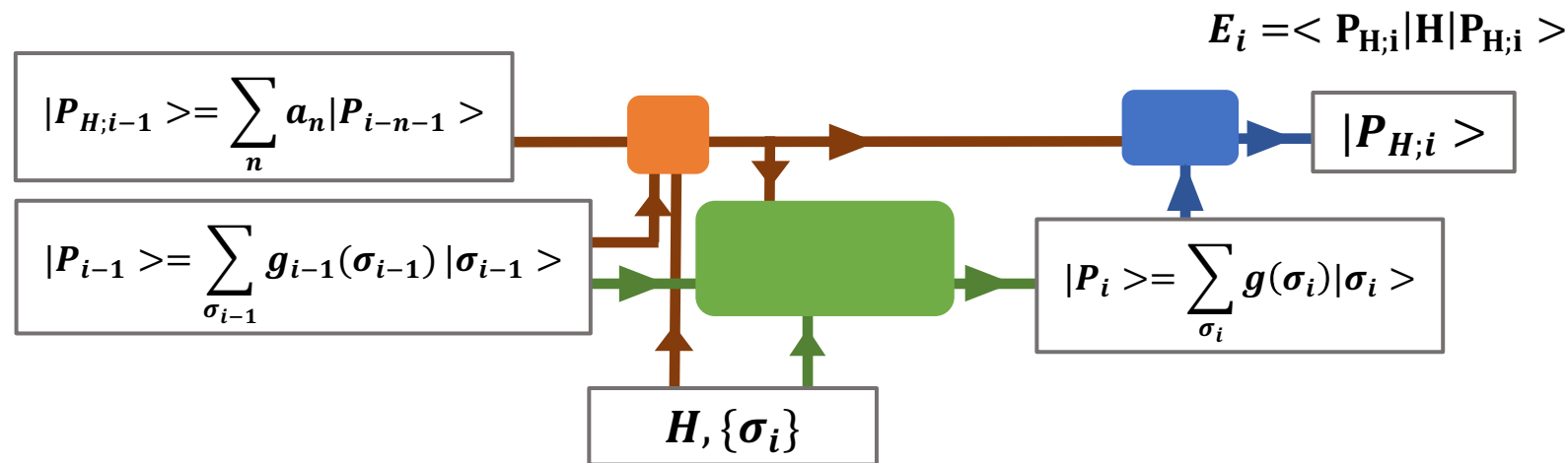
■  $K_{t+1} = K_t + K'_t dt$

■  $H_F(t) = \exp[iK_t] (H_t - K'_t) \exp[-iK_t]$

# Setup

## ● Application of RNN to Physics

- Solve the ground state of 1D(2D) spin system (by Gated Recurrent Unit(GRU))



$E_i$ の代わりに  $U_i = \langle P_{H;i} | U | P_{H;i} \rangle$  をloss関数に選べば、相互作用の効果を摂動的に解析できる基底を選んでくれる？

■ NN      ■  $a_n$  を計算 (forget gate)

■ 状態の履歴に  $|P_i \rangle$  を追加



MPSによる計算と対応  
(PRR:2,023358(2020))