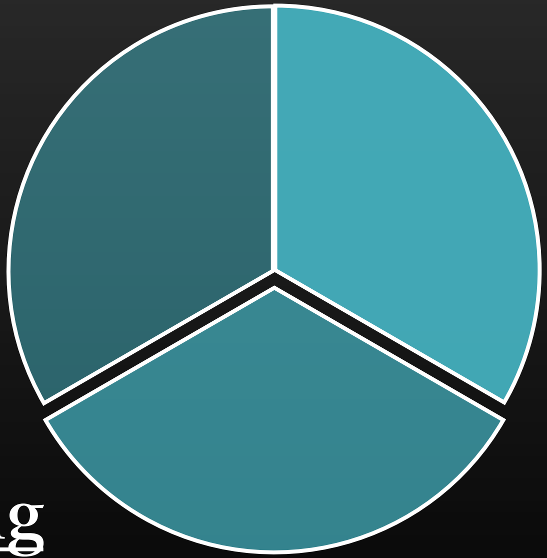
Analysis of Flavor Models by Diffusion Model

Satsuki Nishimura (Kyushu U.), Hajime Otsuka, Haruki Uchiyama

Soon to be submitted [arXiv:2503.xxxxx]

Machine Learning ($\subset$ AI)

- A technique in which a computer extracts hidden rules or patterns as it iteratively learns data.



Supervised Learning

Reinforcement Learning

S. Nishimura, C. Miyao, H. Otsuka,
・JHEP12(2023)021 (2304.14176 [hep-ph])
・2409.10023 [hep-ph]

<u>Unsupervised Learning</u>

K. Ishiguro, S. Nishimura, H. Otsuka,
・JHEP08(2024)133 (2312.07181 [hep-th])

Diffusion Model（DM）

- A type of generative AI, which is used in image generation.

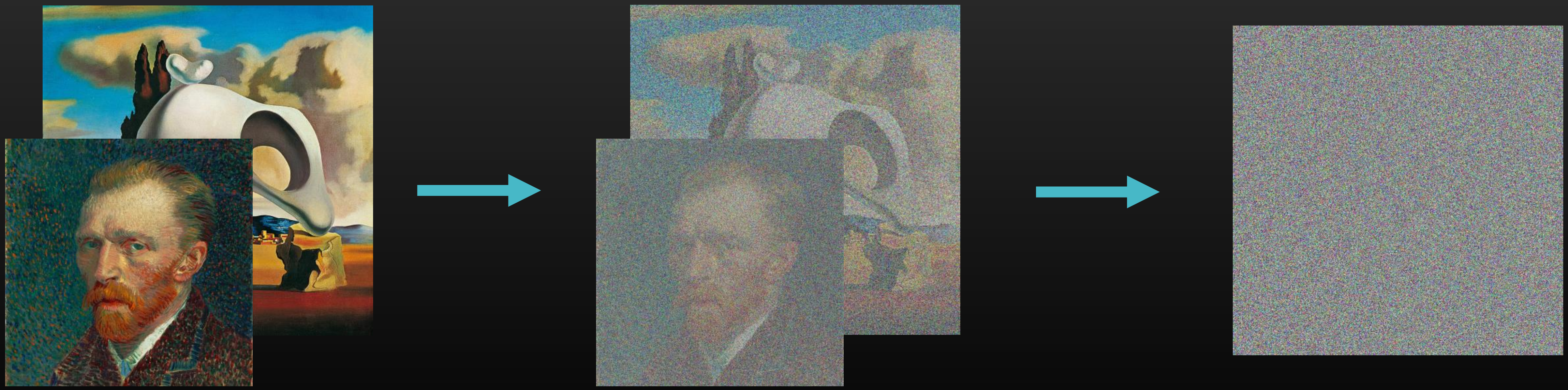
  Ex) AI services “Stable Diffusion” & “DALL-E in ChatGPT”

- Diffusion process　：Adding noises to data

  Reverse process　：Remove noises from Gaussian noise

Diffusion Model（DM）

- Diffusion process：Initial pictures are added noises until they become pure Gaussian noises.



Diffusion Model（DM）

- Reverse process : Pictures are generated by denoising.

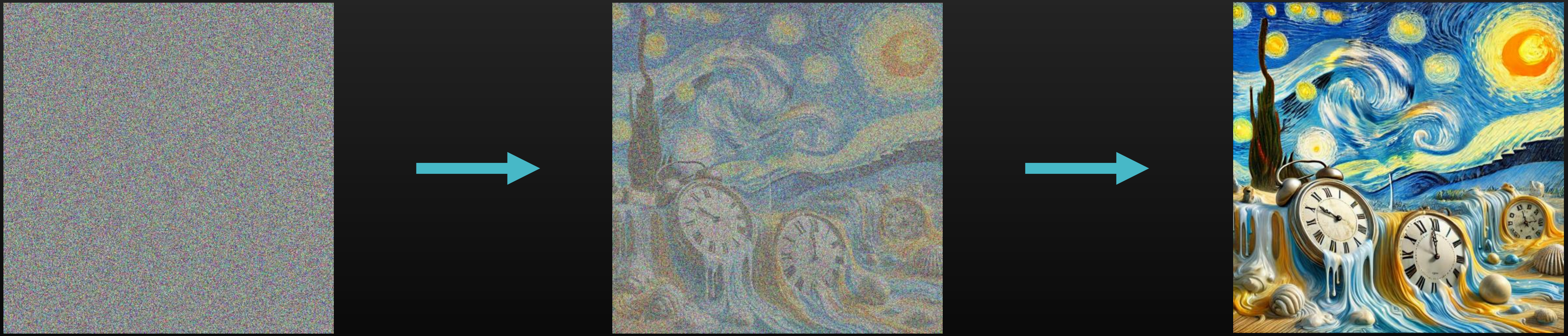
Ex) Make an image with a touch of Salvador Dalí.

Diffusion Model（DM）

- Conditional DM : Pictures are generated with conditional labels.

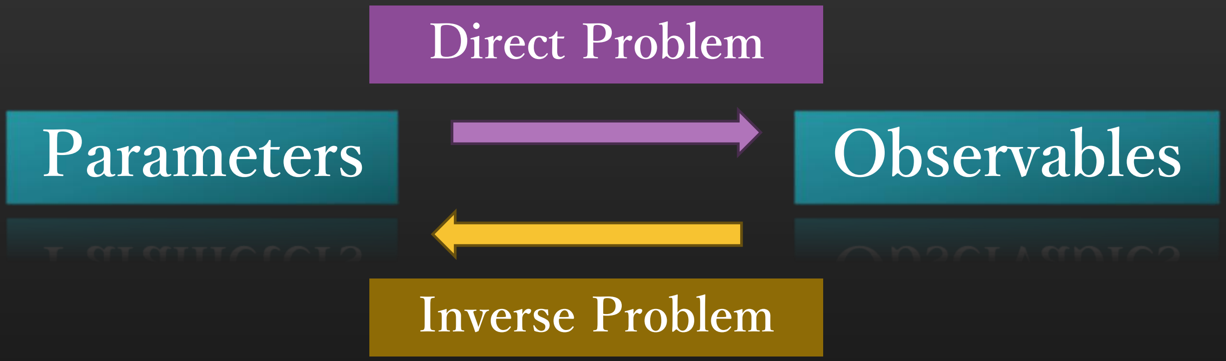
Ex) Make an image of "The Persistence of Memory" by Salvador Dalí,

with a touch of Van Gogh.



Key Point

- Using the diffusion model, we analyze S'_4 modular flavor model.

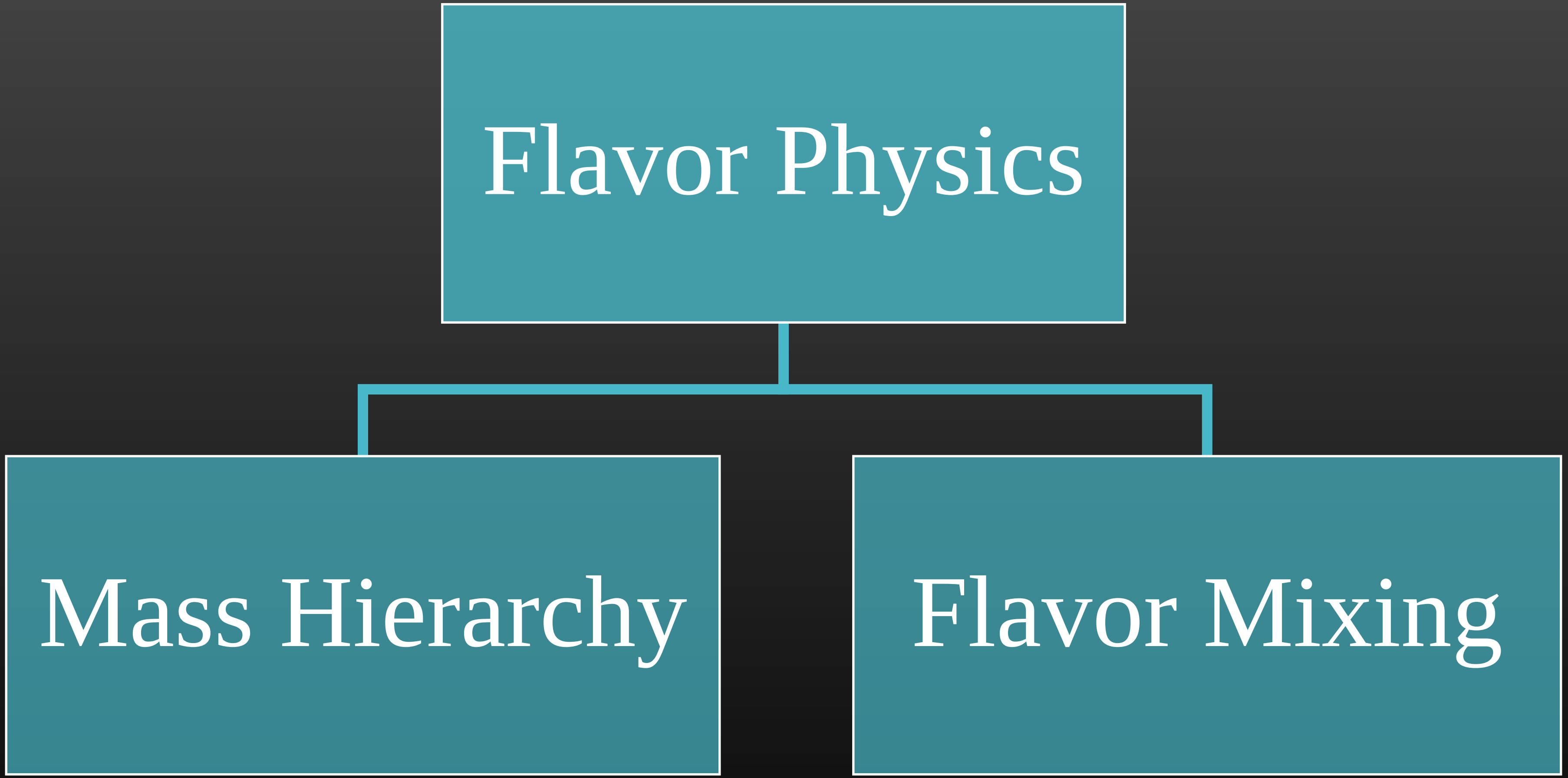


- We found various parameters as an inverse problem, and show the S'_4 model occurs spontaneous CP violation.

The Standard Model (SM)

- SM describes the behavior of elementary particles with a high degree of accurately. It is valid for $\sim 10^{-18}$ m. However, there are many problems. (neutrino masses, generation, ……)

- The search for new physics beyond the Standard Model (BSM) is the challenge in particle physics.


Flavor Physics
Mass Hierarchy
Flavor Mixing

Modular Flavor Model（1）

- Modular flavor symmetry can explain the hierarchical structure of the quarks & the leptons.

  Ex) $SL(2,Z)$, $\Gamma_2 \simeq S_3$（symmetric group of degree 3）,

  S'_4（double covering of S_4）

- Yukawa couplings are representation of modular symmetry. These depend on moduli τ remaining as a degree of freedom.

Modular Flavor Model（2）

- Superpotential of S'_4 modular flavor model

Y.Abe, T.Higaki, J.Kawamura, T.Kobayashi
Phys.Lett.B 842 (2023) 137977

$$W = H_u \left\{ \sum_{a=1}^{2} \alpha_a \left(Q Y_3^{(k_{ua}+k_Q)} \right)_1 u_a^c + \alpha_3 \left(Q Y_{\hat{3}}^{(k_{u3}+k_Q)} u_3^c \right)_1 \right\}$$
$$+ H_d \sum_{i=1}^{3} \left\{ \beta_i \left(Q Y_3^{(k_{di}+k_Q)} d_i^c \right)_1 \right\}$$

- Under modular transformation, modular form with weight k is defined as a function such that $f_i(\gamma\tau) = (c\tau + d)^k \rho_{ij}(\gamma) f_j(\tau)$. Y_3^k means 3 representation with weight k.

Modular Flavor Model（3）

- When representations and weights are fixed,
  Yukawa couplings are calculated using modular forms.

$$Y_u = \begin{pmatrix} \alpha_1 \left[Y_{3}^{k_{u1}+k_Q}(\tau)\right]_1 & \alpha_2 \left[Y_{3}^{k_{u2}+k_Q}(\tau)\right]_1 & \alpha_3 \left[Y_{\hat{3}}^{k_{u3}+k_Q}(\tau)\right]_1 \\ \alpha_1 \left[Y_{3}^{k_{u1}+k_Q}(\tau)\right]_3 & \alpha_2 \left[Y_{3}^{k_{u2}+k_Q}(\tau)\right]_3 & \alpha_3 \left[Y_{\hat{3}}^{k_{u3}+k_Q}(\tau)\right]_3 \\ \alpha_1 \left[Y_{3}^{k_{u1}+k_Q}(\tau)\right]_2 & \alpha_2 \left[Y_{3}^{k_{u2}+k_Q}(\tau)\right]_2 & \alpha_3 \left[Y_{\hat{3}}^{k_{u3}+k_Q}(\tau)\right]_2 \end{pmatrix}$$

- To reproduce masses and mixings of the quarks,
  $O(1)$ constants α, β and moduli τ should be determined properly.

Motivation

- Diffusion model is known as "Generative AI"
  (generate new data and information based on learning)

- It may be possible to
  - generate unknown parameters that reproduce
    the flavor structure in an observable-driven way (Inverse problem)
  - consider a wide range of parameters
    without limiting the search space (in contrast to conventional methods)

Procedure of Learning

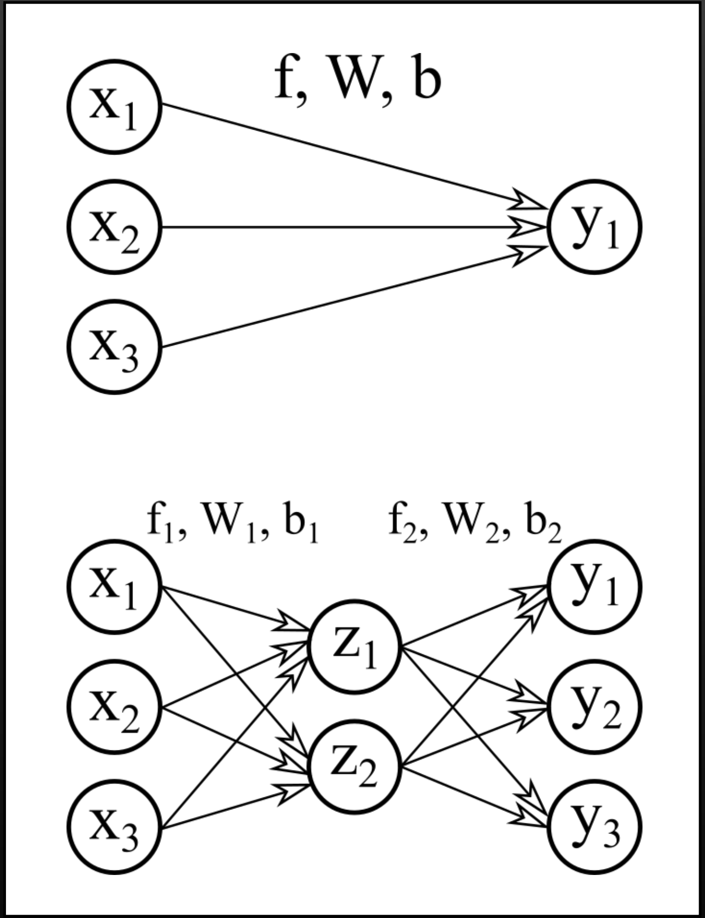
- Generating parameters $G = \{\tau, \alpha, \beta\}$ (α, β are real)

  Calculated observables as labels $L = \left\{\frac{m_u}{m_t}, \frac{m_c}{m_t}, \frac{m_d}{m_b}, \frac{m_s}{m_b}, \left|U_{ij}^{\mathrm{CKM}}\right|, J^q\right\}$

  For various time step t,

1) G is prepared randomly, and related label L is calculated
2) A random fluctuation ΔG (noise) is added as $G_t = \frac{T-t}{T} G + \frac{t}{T} \Delta G$
3) A neural network is trained to predict ΔG from G_t & L

Neural Network（1）

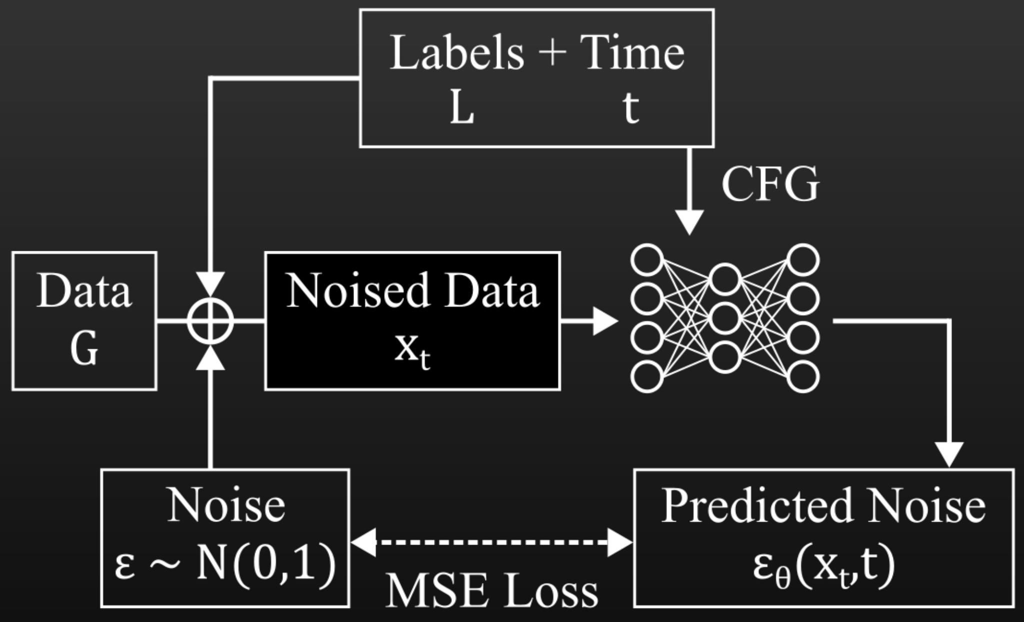


- The mathematical model of brains.
  It has input layers (x) & output layers (y).

- Outputs are calculated from the inputs
  using parameters W, b (weight & bias)
  activation function f

$$y_1 = f(W_1 x_1 + W_2 x_2 + W_3 x_3 + b)$$

Neural Network（2）



- In the diffusion model,
  input： noised data & label info
  output： predicted noise

- Weight W & bias b are updated to decrease the difference between added noise and predicted noise.

Example of Learning（1）

1） G is prepared randomly, and related label L is calculated

$$G:\{\tau,\alpha,\beta\}=\{1+2i,3,4\} \quad \rightarrow \quad L:\left\{\frac{m_u}{m_t},\frac{m_c}{m_t}\right\}=\{10^{-2},10^{-1}\}$$

2） A random fluctuation ΔG（noise）is added

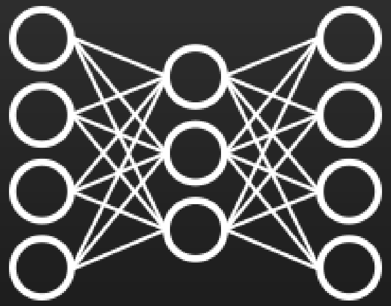
$$\Delta G:\{0.8-0.6i,-0.4,0.2\} \quad \rightarrow \quad G_t:\frac{G+\Delta G}{2}=\{0.9+0.7i,1.3,2.1\}$$

Example of Learning（2）

3） A neural network is trained to predict ΔG from G_t & L

G_t: $\{0.9 + 0.7i, 1.3, 2.1\}$

L: $\{10^{-2}, 10^{-1}\}$



ΔG_{NN}: $\{0.7 + 0.1i, -0.2, 0.1\}$

update parameters of NN
to minimize the difference

ΔG: $\{0.8 - 0.6i, -0.4, 0.2\}$

Procedure of Generating

- Generating parameters $G = \{\tau, \alpha, \beta\}$

  Physical observables as labels $L_{\text{exp}} = \left\{\frac{m_u}{m_t}, \frac{m_c}{m_t}, \frac{m_d}{m_b}, \frac{m_s}{m_b}, \left|U_{ij}^{\text{CKM}}\right|, J^q\right\}$

1） Pure noise G_T is prepared from Gaussian distribution

2） A predicted fluctuation ΔG is denoised as $G_t = \frac{T-t}{T} G_{t+1} - \frac{t}{T} \Delta G$

3） New data G_0 is generated along experimental label L_{exp}

Example of Generating（1）

- We adopted real values of observables during generating.

$$L_{\mathrm{exp}}: \left\{\frac{m_u}{m_t}, \frac{m_c}{m_t}\right\} = \{5.4 \times 10^{-6}, 2.8 \times 10^{-3}\}$$

1） Pure noise G_T is prepared from Gaussian distribution

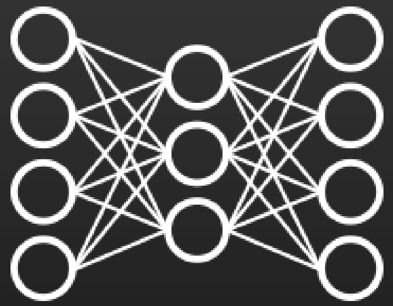
$$G_T: \{-0.3 + 0.7i, -1.2, 0.5\} \quad \text{from} \quad N(0,1)$$

Example of Generating（2）

2）A predicted fluctuation ΔG is denoised

G_T:

$\{-0.3 + 0.7i, -1.2, 0.5\}$



ΔG:

$\{0.08 - 0.02i, -0.06, 0.1\}$

$$G_{T-1}: \frac{G_T - \Delta G}{2} = \{-0.19 + 0.36i, -0.57, 0.20\}$$

3）New data G_0 is generated along experimental label L_{exp}

Progress of Generating（1）

- From the parameters found by DM, 14 observables are calculated.

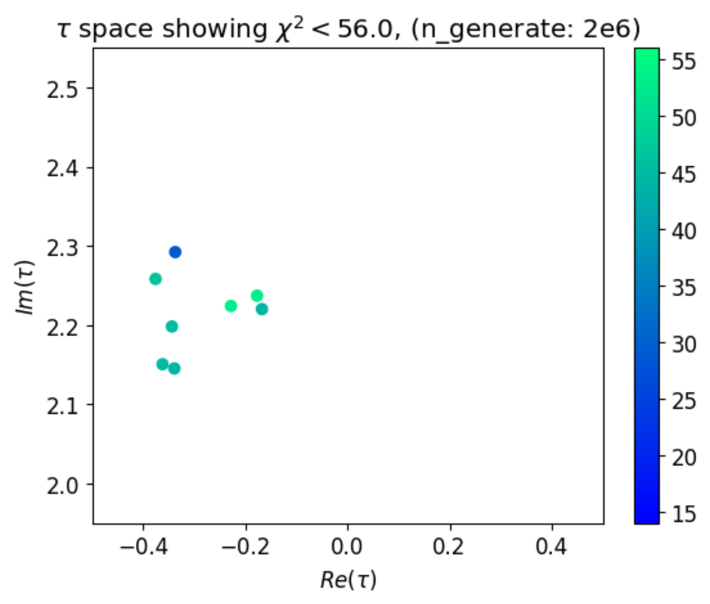
$$\left\{\frac{m_u}{m_t}, \frac{m_c}{m_t}, \frac{m_d}{m_b}, \frac{m_s}{m_b}, \left|U_{ij}^{\mathrm{CKM}}\right|, J^q\right\}$$

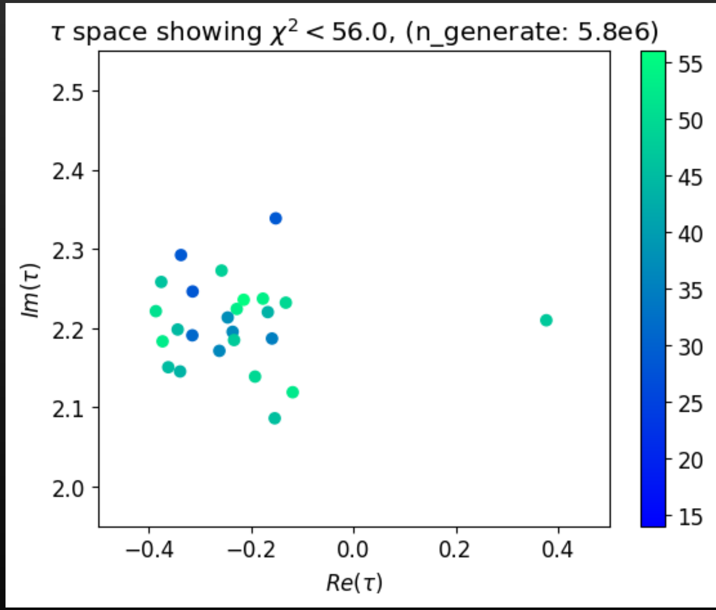
- χ^2 is defined from calculated & experimental values $v_{\mathrm{cal}}, v_{\mathrm{exp}}$

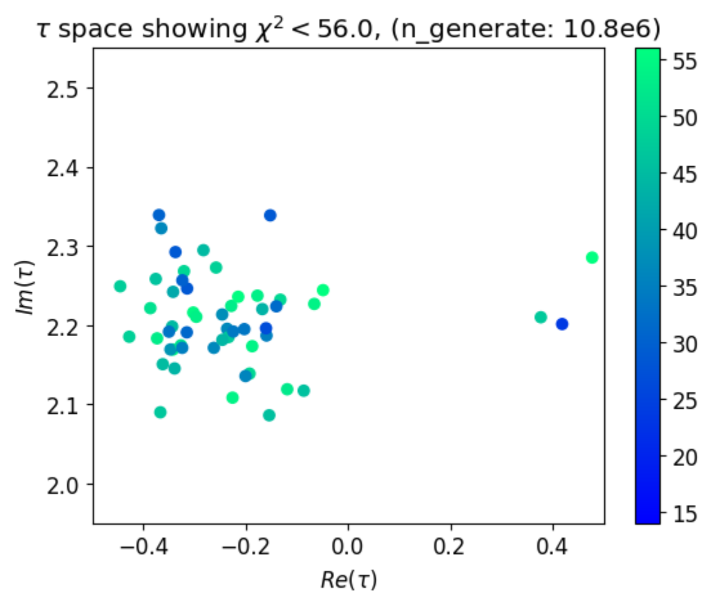
$$\chi^2 = \sum \chi_i^2 = \sum \frac{\left(v_{i,\mathrm{cal}} - v_{i,\mathrm{exp}}\right)^2}{v_{i,\mathrm{error}}^2}$$

Progress of Generating（2）

- The distributions of moduli τ with high accuracy are shown.
  (with an increase in the data generated by the DM)



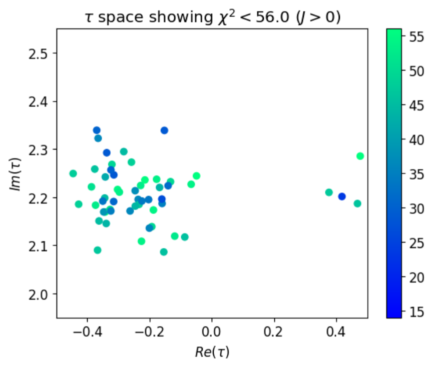




Result：Various Solutions

- The flavor structure of S'_4 model is sensitive to $\mathrm{Im}\,\tau$. It is an arduous task to find good parameter regions. Analytically, large $\mathrm{Im}\,\tau$ is preferred. (Ex：$\mathrm{Im}\,\tau = 2.8$)

- We found various solutions with smaller $\mathrm{Im}\,\tau \sim 2.2$ even without strict limitations for param. region.

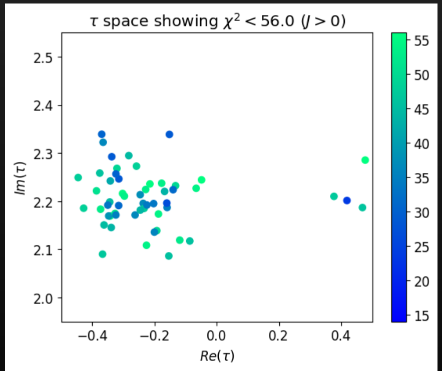


Result : Spontaneous CP Violation

- Coefficients α, β are real. We found $\mathrm{Re}\,\tau$ reproduces Jarlskog invariant, so spontaneous CP violation is derived.

$$J_{\mathrm{DM}} \sim 2.70 \times 10^{-5}, \qquad J_{\mathrm{exp}} = 2.87 \times 10^{-5}$$

- In previous works, complex coefficients are often introduced to reproduce CP violation.
  It is interesting to find the parameters with SCPV.



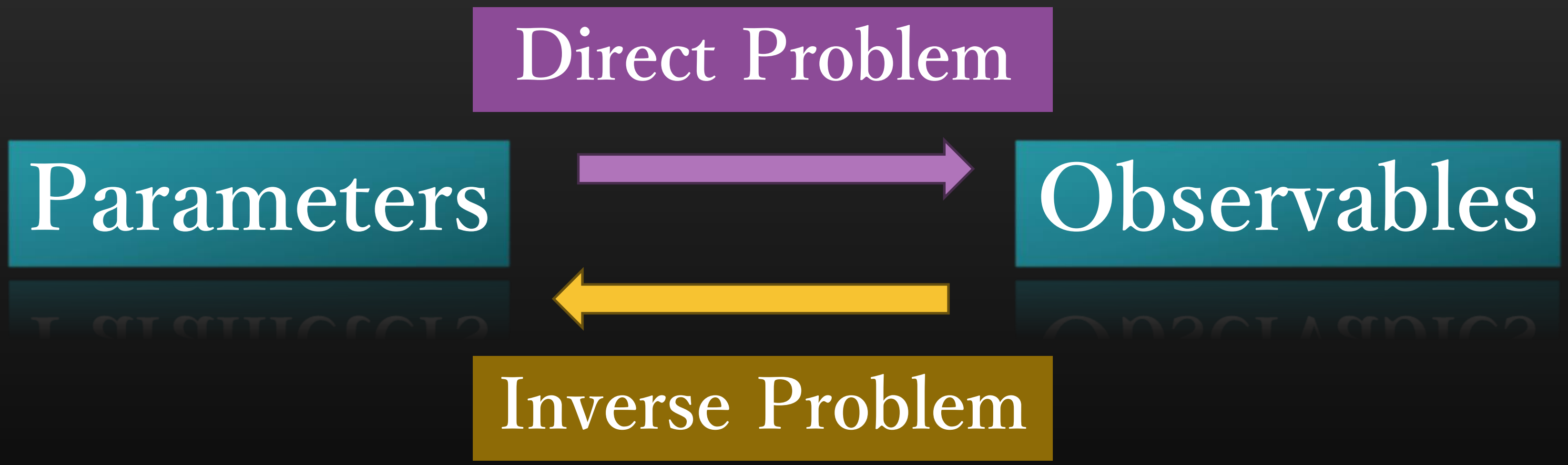
Summary : application of DM for flavor physics

- We focused on the S'_4 modular flavor model,
  and searched parameters with diffusion model.

- We found various parameters with smaller $\mathrm{Im}\,\tau$ which reproduce the flavor structure of quark sector in an observable-driven way.
  This approach do not need constraints on searching region.

Summary : application of DM for flavor physics

- Even when coefficients of Yukawa are real, the S'_4 model occurs the spontaneous CP violation.



Outlook

- Will spontaneous CP violation be visible with A_4 or other modular symmetries? Will non-trivial parameters be discovered?

- We hope to construct a framework that get predictions from flavor models by simply specifying the experimental values.