ICNN - JKO.

- outer loop learning rate η_1; inner loop learning rate η_2; JKO discretization steps T_2. $\varphi_\theta \leftarrow$ basic ICNN model with pretrained parameters.; h is the JKO step size.; n: # of samples from μ.
- for $t_1 = 1, 2, \dots, T_1$

  set $\varphi'_\theta \leftarrow \varphi_\theta$ (previous ICNN result)

  for $t_2 = 1, 2, \dots, T_2$

    Sample $(z_i)_{i=1}^n \sim \mu$ (μ is the reference Dist) $\bar{\mu}(\cdot)$ is the p.d.f of ref Dist.

    W_2^2 or L_2^2

    $W_2^2 \leftarrow W_2^2(\nabla\varphi'_\theta(z), \nabla\varphi_\theta(z))$.

    $\pi_n(\cdot)$ p.d.f of posterior.

    $$F \leftarrow \frac{1}{n}\sum_{z\in Z}\left[\log\det\nabla^2\varphi'_\theta(z)\cdot\bar{\mu}(z) - \log\left(\pi_n(\nabla\varphi'_\theta(z))\right)\right].$$

    $$\mathcal{L} \leftarrow F + \frac{1}{2h}W_2^2$$

    Perform a gradient step over θ by using $\frac{\partial\hat{\mathcal{L}}}{\partial\theta}$ to update φ'_θ

  $\varphi_\theta \leftarrow \varphi'_\theta$

e.g. L_2 SGD $m=0.9$ $lr = 5e\text{-}2$, $n=100$, $(2, 25)$ jko_t = 3.5 ini $N(0,2)$.

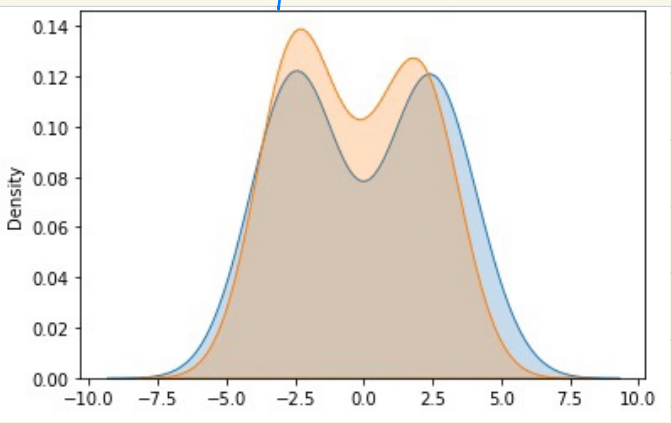
is using L_2^2 instead of W_2^2, optimizer SGD with momentum 0.9, learn rate 5e-2.

in the inner loop each time generate $n=100$ samples, $T_1=2$, $T_2=25$, jko_t is $h = 3.5$.

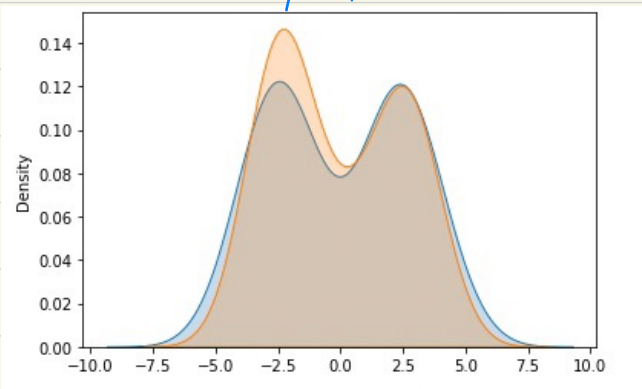
The $\varphi_\theta^{(0)}$ the initinal ICNN pre-trained to map: $z\sim N(0,1)$, $\nabla\varphi_\theta^{(0)}(z) \sim N(0,2)$.

L2 SGD $m=0.9$ $lr=5e-2$, $n=100$, (2,25) iko-t=3.5 ini N(0.2).
(4,25)
(6,25)
(8,25)
(10,25)

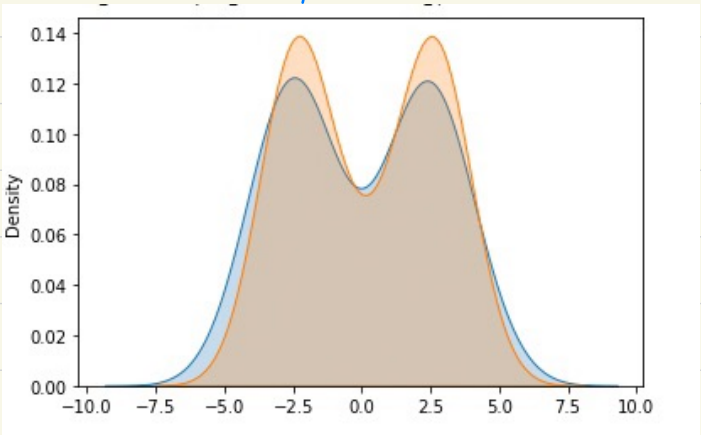
epoch 2



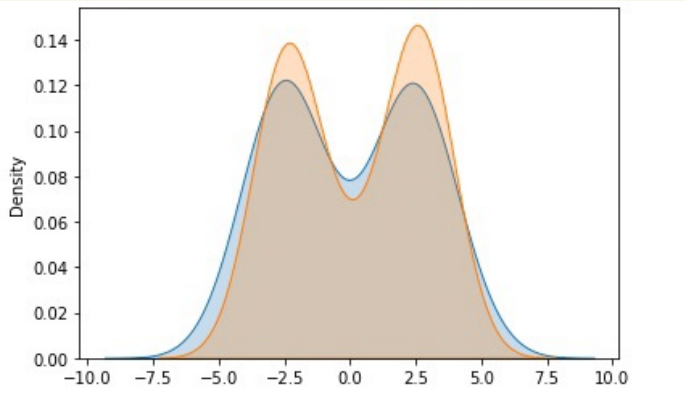
epoch 4



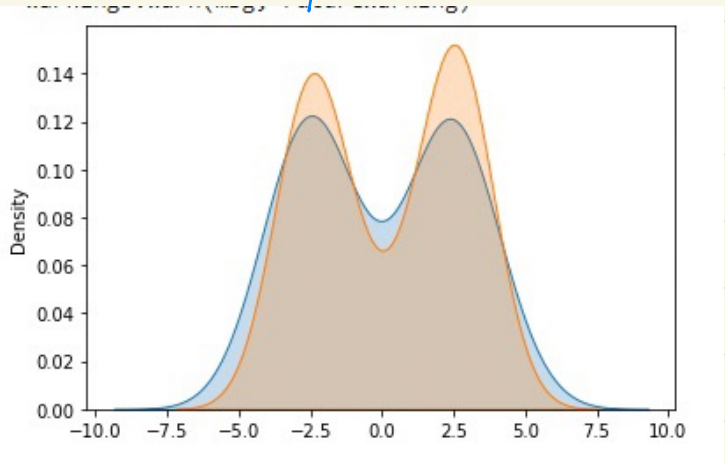
epoch 6



epoch 8

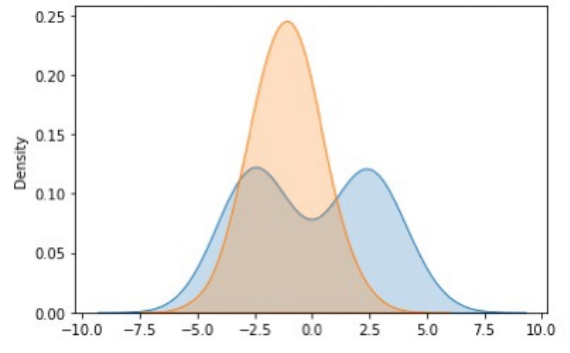


epoch 10.

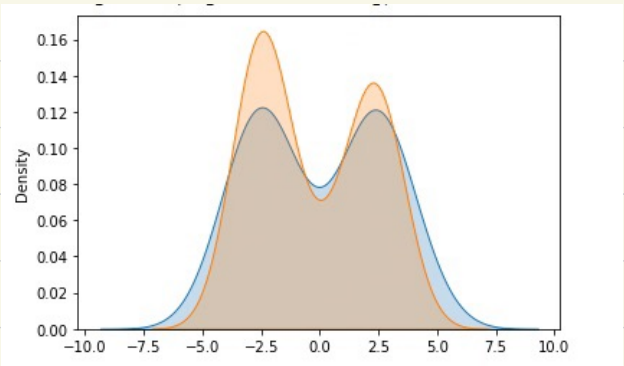


L_2 SGD $m=0.9$ $lr=5e\text{-}2$, $n=100$, $(10,25)$ jko_t = {0.2, 0.5, 1, 3.5, 5, 10} ini $N(10,2)$.

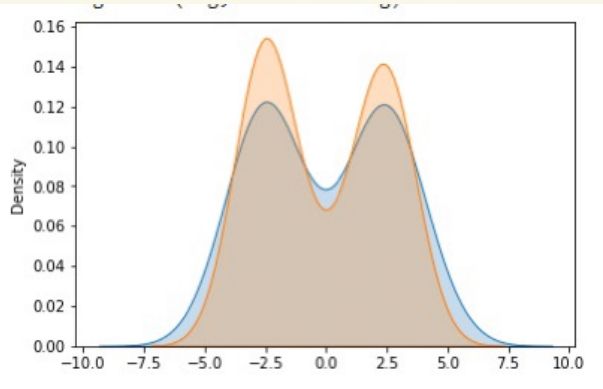
jko_t = 0.2



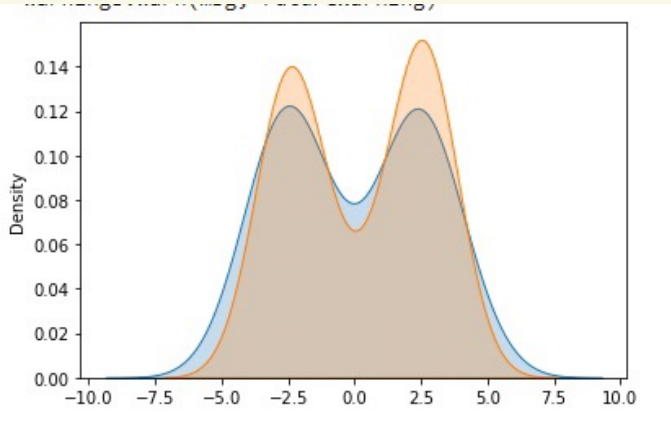
jko_t = 0.5



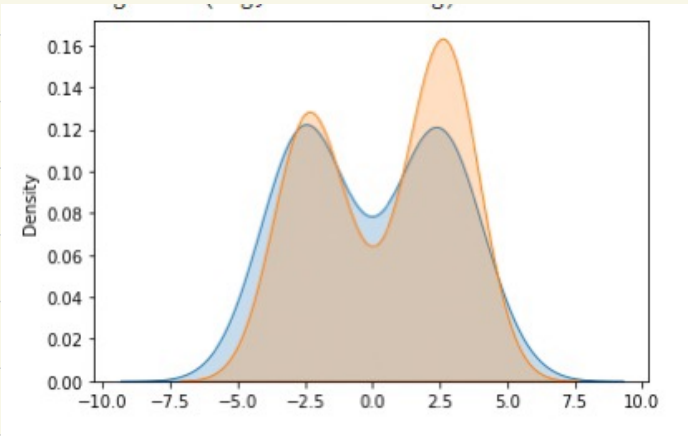
jko_t = 1



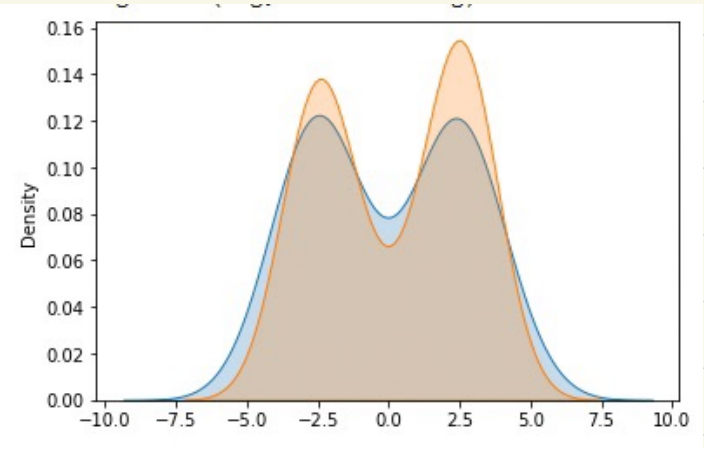
jko_t = 3.5



jko_t = 5



jko_t = 10.

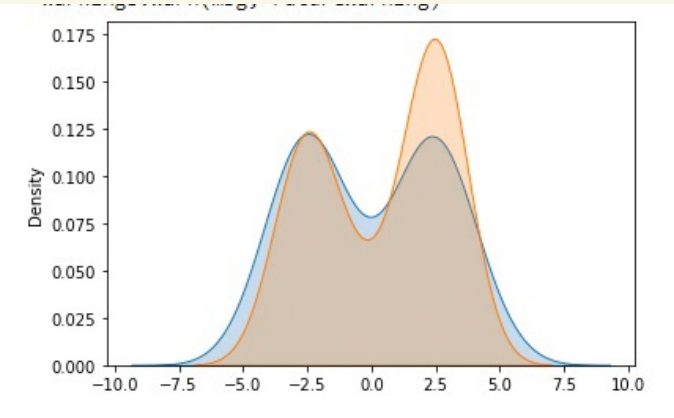


L2 SGD $m=0.9$ $lr=5e-2$, $n=100$, $(10, 25)$ iko_t=3.5 ini $N(0,2)$.

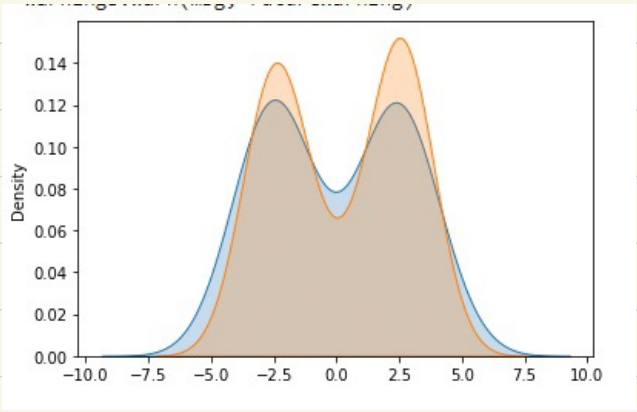
ini $N(0,1)$

ini WGAN.

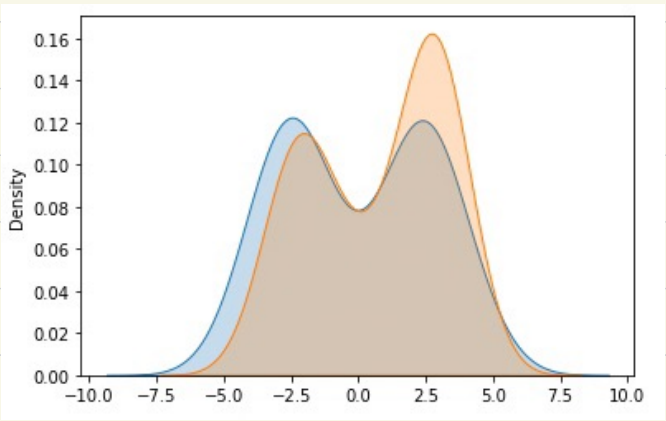
ini $N(0,1)$.



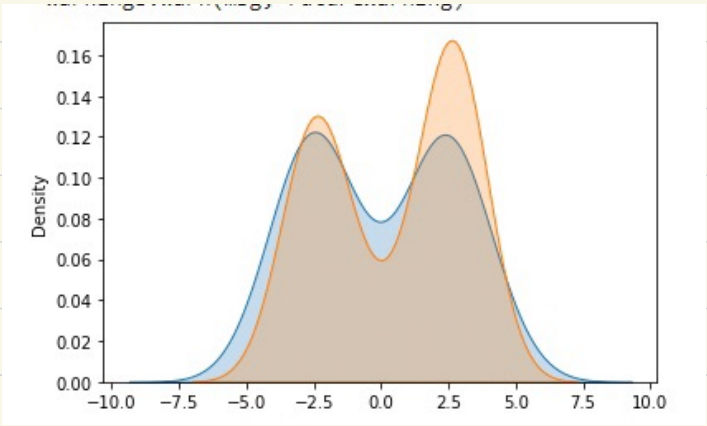
ini $N(0,2)$



WGAN GP result baseded on 10 samples + noise.

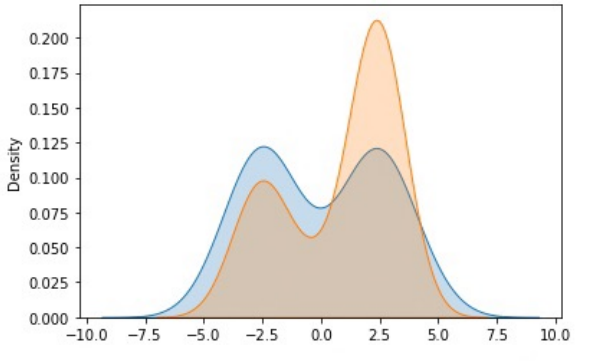


WGAN GP ini

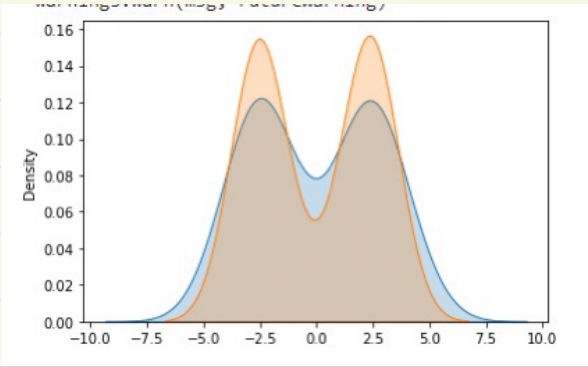


L2 SGD M=0.9 lr = 5e-2, n=30 (10, 25) jko-t=3.5 ini N(0,2).
(10, 50)
(10, 75)
(10, 100).

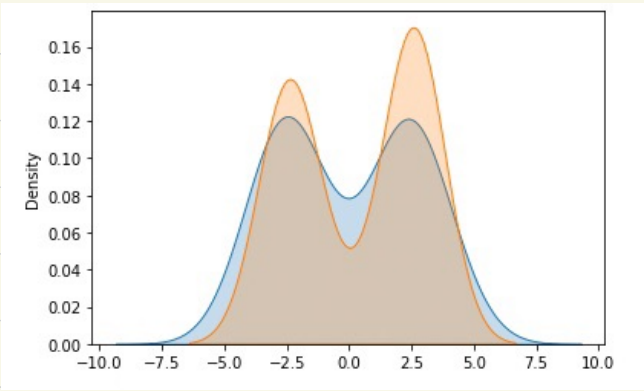
(10, 25)



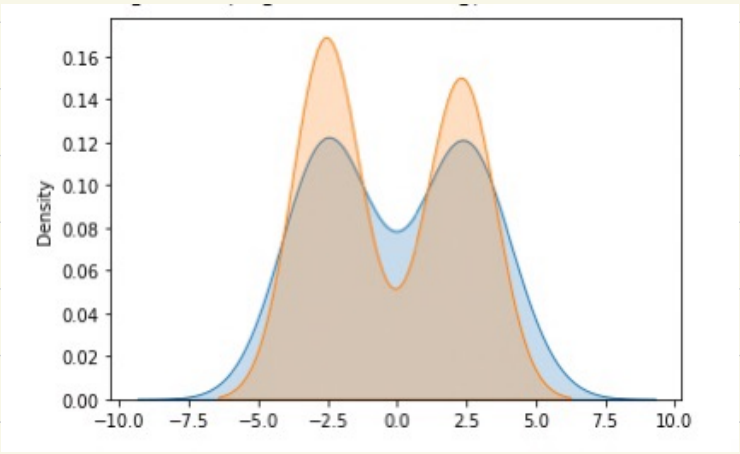
(10, 50)



(10, 75).

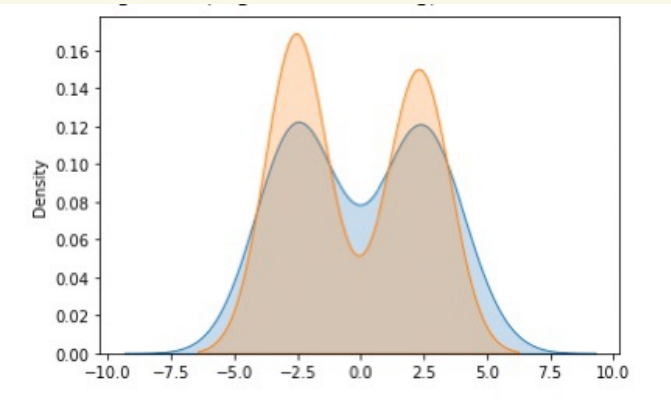


(10, 100).

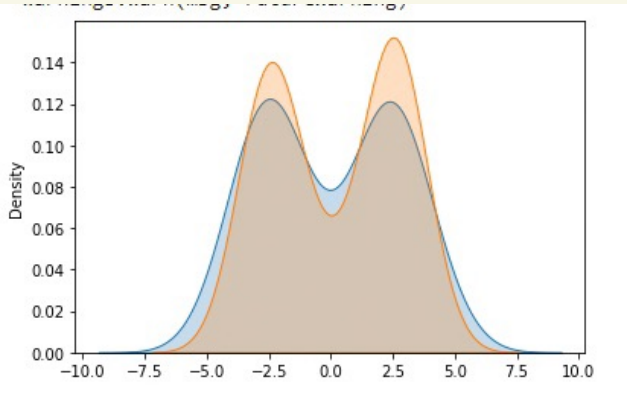


Compare $n=30$, $(10, 100)$; $n=100$, $(10, 25)$; $n=1000$, $(10, 15)$.

$n=30$, $(10, 100)$.



$n=100$ $(10, 25)$.



$n=1000$, $(10, 15)$.