FINER: Flexible spectral-bias tuning in Implicit NEural Representation by Variable-periodic Activation Functions: Supplemental Material

Zhen Liu[1,†], Hao Zhu[1,†], Qi Zhang[2], Jingde Fu[3], Weibing Deng[3], Zhan Ma[1], Yanwen Guo[4], Xun Cao[1]
[1] School of Electronic Science and Engineering, [3] Department of Mathematics,
[4] Department of Computer Science and Technology, Nanjing University, Nanjing 210023, China
[2] AI Lab, Tencent Company, Shenzhen 518054, China
[†] Equal contribution. Corresponding author: caoxun@nju.edu.cn

1. 2D Image Fitting

1.1. Comparisons with baselines

Fig. 1 provides more qualitative comparisons between the FINER and baselines on fitting images.

1.2. Effect of number of hidden neurons

Fig. 2 shows the PSNR on fitting image with varying number of hidden neurons. In all cases, the number of hidden layers were fixed to be two. FINER significantly outperforms baselines with the increase of hidden neurons.

2. 3D Shape Representation

Fig. 3 shows qualitative comparisons on representing the signed distance field of Beared Man, Dragon, Thai Statue and Lucy.

3. Neural Radiance Fields Optimization

Fig. 4 shows more close-up and full-image comparisons on the Blender dataset [5].

4. Heuristic NTK Analysis for FINER network

We use the Neural tangent kernel (NTK) theory [3] to analysis the convergence behavior of FINER. Consider a FINER network with 1D input, 1D output and 1 hidden layer with n neurons, then it can be written as

$$f(x;\theta) = \sum_{k=1}^{n} c_k \sigma(w_k x + b_j) \tag{1}$$

where $x \in \mathbb{R}$ is the input, $\theta = \{\boldsymbol{w} \in \mathbb{R}^{n\times 1}, \boldsymbol{b} \in \mathbb{R}^n\}$ is the parameters of the first layer, $c \in \mathbb{R}^n$ is the output weight, and $\sigma(x) = \sin((|x|+1)x)$ is the activation function. To simplify, we follow [2] that only varies θ and keeps c fixed, specifically, $c_k = \pm\frac{1}{\sqrt{n}}$. Obviously, the derivative of $\sigma(x)$ is

$$\sigma'(x) = (2|x|+1)\cos((|x|+1)x) \tag{2}$$

Thus,

$$\begin{aligned}\nabla_{\boldsymbol{w}} f(x;\theta) &= \Big[x c_1 \sigma'(w_1 x + b_1), \cdots, x c_n \sigma'(w_n x + b_n)\Big] \\ &= x\Big[c_1 \sigma'(w_1 x + b_1), \cdots, c_n \sigma'(w_n x + b_n)\Big]\end{aligned} \tag{3}$$

and

$$\nabla_{\boldsymbol{b}} f(x;\theta) = \Big[c_1 \sigma'(w_1 x + b_1), \cdots, c_n \sigma'(w_n x + b_n)\Big] \tag{4}$$

According to [3], the NTK matrix for the shallow FINER network is $\mathbf{K} = [\mathbf{K}(x_i, x_j)]_{i,j=1,\cdots,N}$, where N is number of inputs, and

$$
\begin{aligned}
\mathbf{K}(x_i, x_j) &= \mathbb{E}_\theta \langle \nabla_\theta f(x_i;\theta), \nabla_\theta f(x_j;\theta) \rangle \\
&= \mathbb{E}_\theta \langle \nabla_{\boldsymbol{w}} f(x_i;\theta), \nabla_{\boldsymbol{w}} f(x_j;\theta) \rangle + \mathbb{E}_\theta \langle \nabla_{\boldsymbol{b}} f(x_i;\theta), \nabla_{\boldsymbol{b}} f(x_j;\theta) \rangle \\
&= (x_i x_j + 1)\mathbb{E}_\theta \sum_{k=1}^{n} c_k^2 \sigma'(w_k x_i + b_k)\sigma'(w_k x_j + b_k) \\
&= (x_i x_j + 1)\mathbb{E}_\theta \sum_{k=1}^{n} c_k^2 \sigma'(g_k(x_i))\sigma'(g_k(x_j)) \\
&= (x_i x_j + 1)\mathbb{E}_\theta \sum_{k=1}^{n} c_k^2 \underbrace{(2|g_k(x_i)| + 1)(2|g_k(x_j)| + 1)}_{\text{Scale term}} \\
&\quad \underbrace{\cos\left((|g_k(x_i)| + 1)g_k(x_i)\right)\cos\left((|g_k(x_j)| + 1)g_k(x_j)\right)}_{\text{Sign term}}
\end{aligned} \tag{5}
$$

where $g_k(x_i) = w_k x_i + b_k$.

It is observed that, the scale terms are approximately proportional to the absolution of bias b_k for all nodes of the kernel, however the change rule of sign terms for diagonal elements differs significantly from non-diagonal elements due to the high-frequency oscillation of $\cos((|x| + 1)x)$ when x is large. Specifically, the sign terms are almost always positive for diagonal elements $\mathbf{K}(x_i, x_i)$ while could be either positive or negative for non-diagonal elements $\mathbf{K}(x_i, x_j)$ with nearly equal probability for different $k(k = 1, 2, \ldots, n)$, leading to cancellation between these terms. As a result, the diagonal elements $\mathbf{K}(x_i, x_i)$ are increased with the increase of $|\boldsymbol{b}|$, while the non-diagonal elements $\mathbf{K}(x_i, x_j)$ can be very small, appearing as a diagonal enhanced kernel. Furthermore, a matrix with a stronger diagonal structure has a greater number of large eigenvalues. By NTK theory [3, 8], the convergence of FINER with a larger $|\boldsymbol{b}|$, that is, with larger NTK eigenvalues, is expected to be accelerated, compared with classical MLP and SIREN [7].

5. Comparisions with SOTAs

While [1, 4, 6] have already explored the idea of variable frequencies and amplitudes in SIREN, FINER differs significantly from them:

1) We propose a new activation function with inherent variable frequency characteristics, improving the precision and compactness of INR.

2) The SOTA methods require additional (trainable) parameters but exhibit limited control capability over frequency and amplitude variations. Modulated Periodic Activation lacks direct control over frequency. The frequency modulation of MetaSin is constrained by the number of sinusoidal components. NeuRBF employs sinusoidal composition only in the first layer of the MLP and could be viewed as a variant of positional encoding [8], with a fixed and immutable modulation frequency.

In Tab. 1, we provide comparisons on 2D image fitting between FINER, Modulated Periodic Activation with an auto-encoder setting (MPA), MLP with MetaSin activation (MS-MLP), and MLP with sinusoidal composition in NeuRBF (SC-MLP), where the natural datase [8] is adopted. Note that, only the components of controlling frequency/amplitudes are used for fair comparisons. FINER outperforms the baselines.

Table 1. Quantitative comparisons on image fitting.

Metrics	MPA	MS-MLP	SC-MLP	FINER
PSNR ↑	25.96	30.69	26.96	40.76
SSIM ↑	0.7334	0.8850	0.7518	0.9790
LPIPS ↓	3.62e-1	1.266e-1	1.89e-1	2.56e-3

6. Visualizations of the trained frequencies

Fig. 5 shows the activation statistics of the first two layers of FINER. The maximum frequency grows with the increase of layers, additionally, larger k results in larger frequencies in different layers.

References

[1] Zhang Chen, Zhong Li, Liangchen Song, Lele Chen, Jingyi Yu, Junsong Yuan, and Yi Xu. Neurbf: A neural fields representation with adaptive radial basis functions. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4182–4194, 2023. 2

[2] Simon S. Du, Xiyu Zhai, Barnabás Póczos, and Aarti Singh. Gradient descent provably optimizes over-parameterized neural networks. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6-9, 2019*. OpenReview.net, 2019. 1

[3] Arthur Jacot, Franck Gabriel, and Clement Hongler. Neural tangent kernel: Convergence and generalization in neural networks. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2018. 1, 2

[4] Ishit Mehta, Michaël Gharbi, Connelly Barnes, Eli Shechtman, Ravi Ramamoorthi, and Manmohan Chandraker. Modulated periodic activations for generalizable local functional representations. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 14214–14223, 2021. 2

[5] Ben Mildenhall, Pratul P Srinivasan, Matthew Tancik, Jonathan T Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In *European conference on computer vision*, pages 405–421. Springer, 2020. 1

[6] Farnood Salehi, Tunç Aydin, André Gaillard, Guglielmo Camporese, and Yuxuan Wang. Empowering convolutional neural nets with metasin activation. *Advances in Neural Information Processing Systems*, 36, 2023. 2

[7] Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. *Advances in neural information processing systems*, 33:7462–7473, 2020. 2

[8] Matthew Tancik, Pratul Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems*, 33:7537–7547, 2020. 2

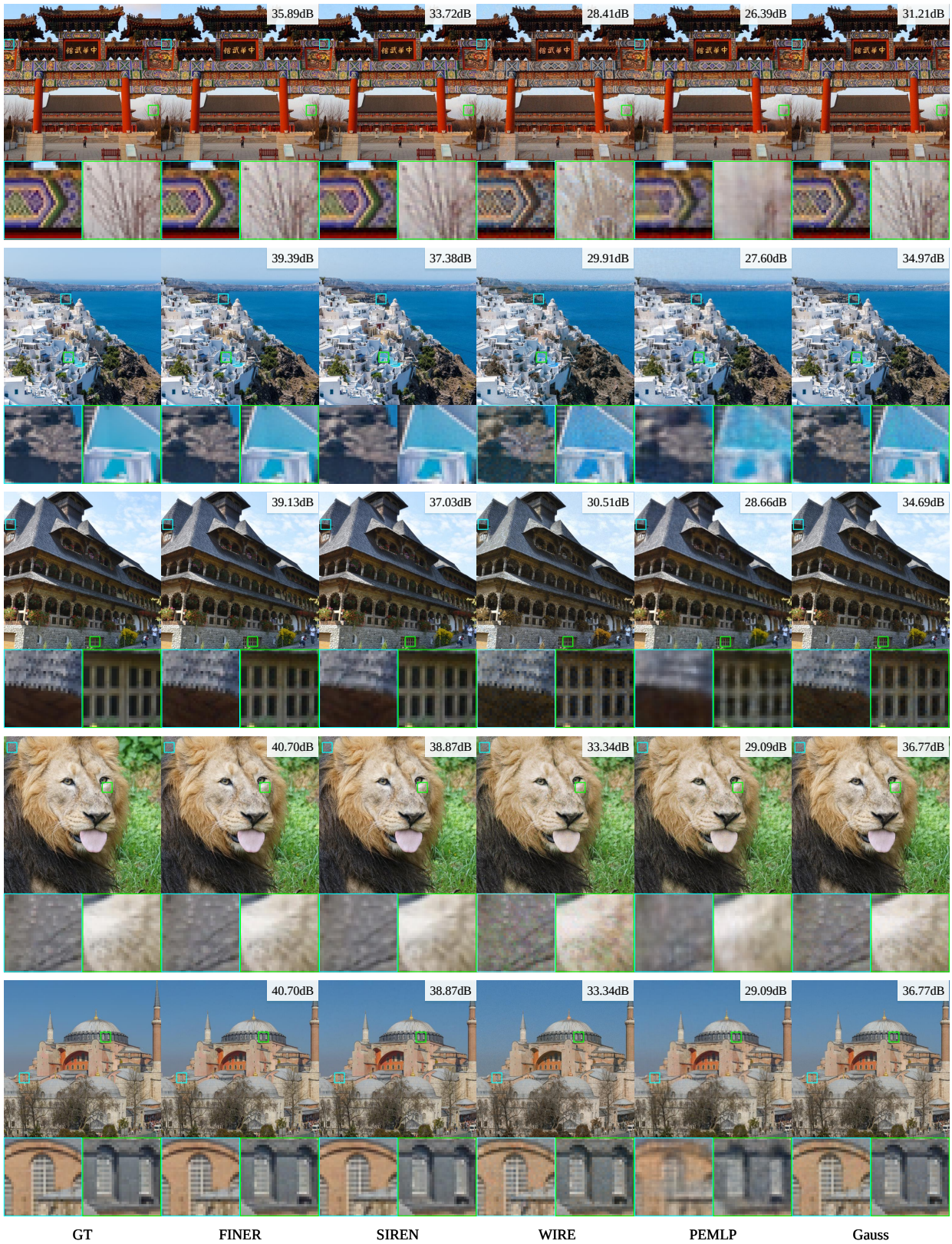

Figure 1. Qualitative comparisons between the FINER and baselines on fitting images.

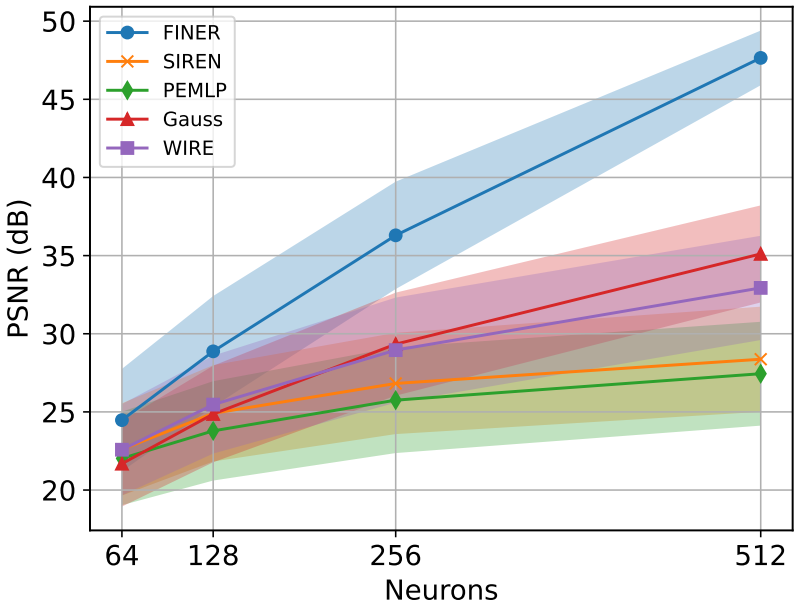

Figure 2. Qualitative comparisons between the FINER and baselines on fitting images with different number of hidden neurons.

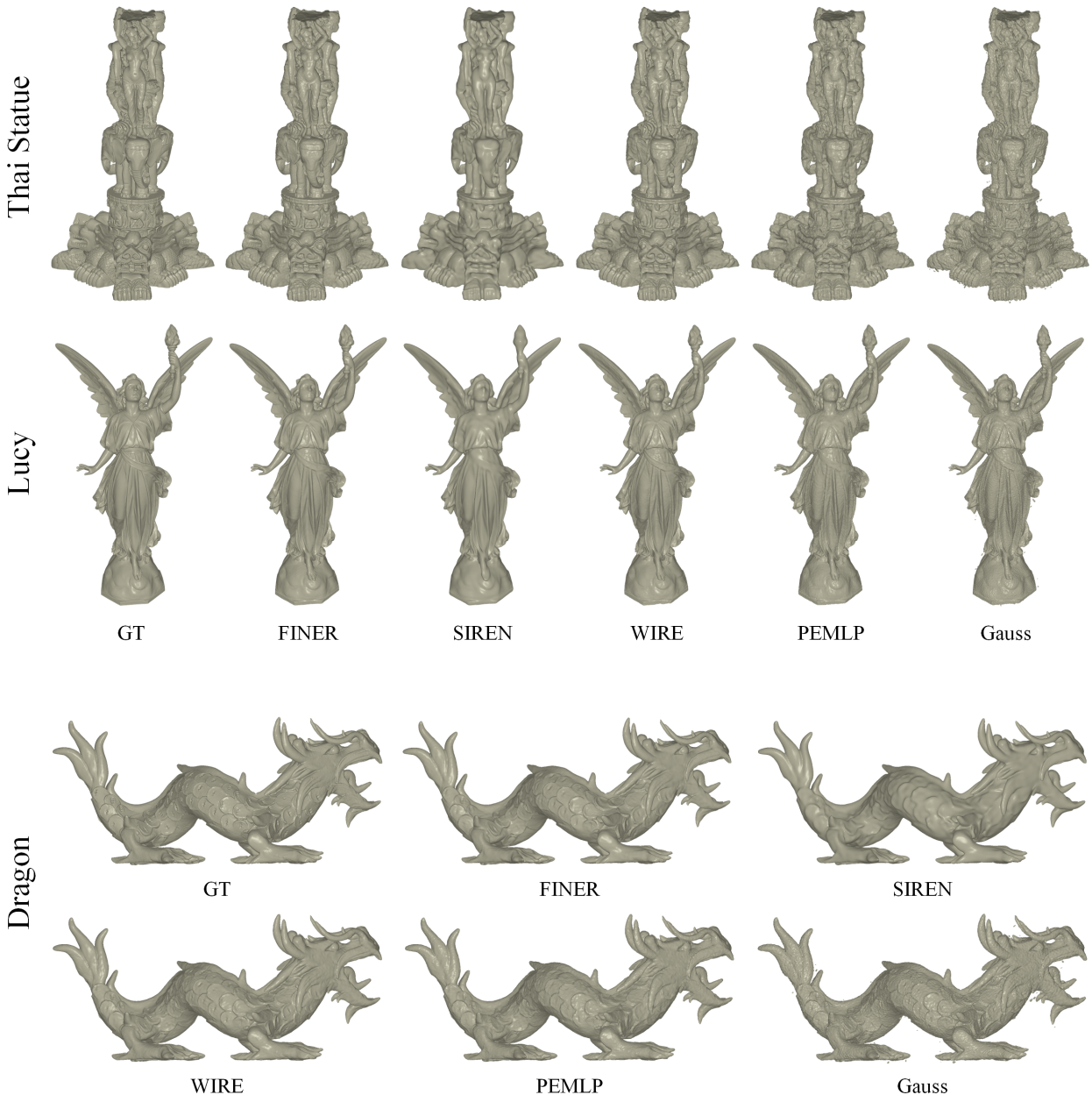

Figure 3. Qualitative comparisons between the FINER and baselines on representing the signed distance field.

Figure 4. Qualitative comparisons between the FINER and baselines on NeRF.

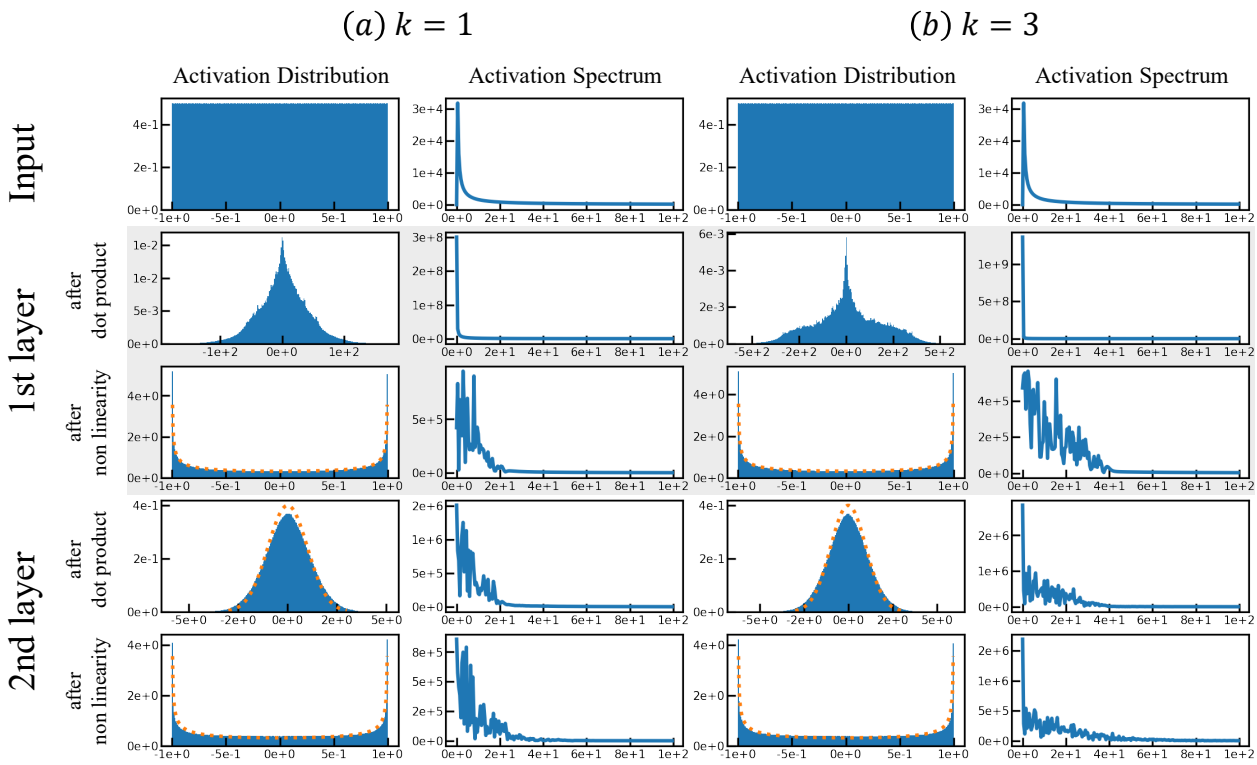

Figure 5. Activation statistics at initialization for a 2-layer FINER. **Better viewed on screen with zooming in.**