Grounding and Enhancing Grid-based Models for Neural Fields

Zelin Zhao[1], Fenglei Fan[2*], Wenlong Liao[1,3], Junchi Yan[1]
[1] Shanghai Jiao Tong University
[2] The Chinese University of Hong Kong
[3] Cowa Tech. Ltd

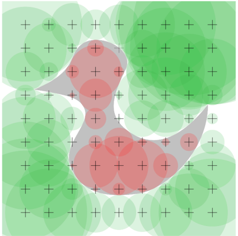
*Corresponding author.

Neural fields

Neural fields are coordinate-based networks representing a field, a continuous parameterization representing a physical quantity of an object or a scene. These fields have demonstrated significant success in various tasks in computer vision† and beyond:

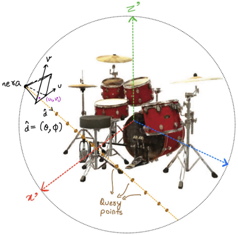
Image regression

(x, y) → RGB



SDF reconstruction

(x, y, z) → Signed distance

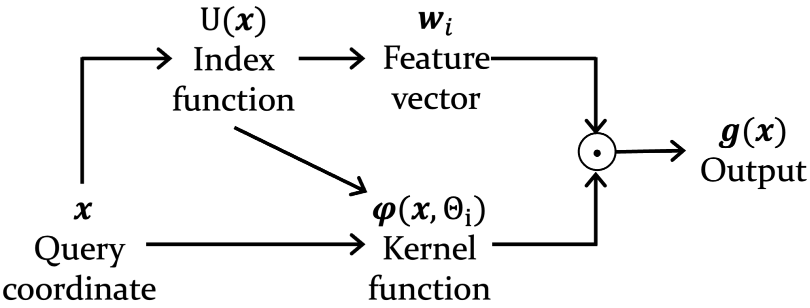


Novel view synthesis

(x, y, z) → RGB, density

†Ben Mildenhall et al. “Nerf: Representing scenes as neural radiance fields for view synthesis”. In: *European Conference on Computer Vision*. Springer. 2020, pp. 405–421.

Grid-based models



A grid-based model takes a query coordinate $\boldsymbol{x}$ as input, which is sent to an index function U to acquire a set of feature vectors $\boldsymbol{w}$ from the grid. Then, the model outputs a weighted average of the kernel function φ and the feature vectors $\boldsymbol{w}$.

Motivations of our theory

We wish to understand and enhance grid-based models:

1. How can we understand the training dynamics of a grid-based model?
2. How can we measure the generalization performance of a grid-based model?
3. How can we design a better grid-based model?

We propose a theoretical framework based on tangent kernels[‡] to address those questions.

[‡] Arthur Jacot, Franck Gabriel, and Clément Hongler. "Neural tangent kernel: Convergence and generalization in neural networks". In: *Advances in neural information processing systems* 31 (2018).

Grid tangent kernel (GTK)

Our theoretical results show that the approximation and generalization performances of grid-based models are related to the grid-tangent kernel (GTK), which is defined as a positive semidefinite matrix in the following form:

$$[\boldsymbol{G}_g(t)]_{i,j} = \left\langle \frac{\partial g\left(\boldsymbol{X}_i, \boldsymbol{w}(t)\right)}{\partial \boldsymbol{w}}, \frac{\partial g\left(\boldsymbol{X}_j, \boldsymbol{w}(t)\right)}{\partial \boldsymbol{w}} \right\rangle, \tag{1}$$

where g is a grid model parameterized by $\boldsymbol{w}(t)$, $\boldsymbol{X}$ is the dataset where $\boldsymbol{X}_i$ is the i-th data. The GTK measures the distance of two data points in the gradient space, which is dependent on the model and the data.

Derivation of the GTK

The model parameters evolve according to the following differential equation:

$$\frac{dg\left(\boldsymbol{X}_i, \boldsymbol{w}(t)\right)}{dt} = \frac{d\boldsymbol{w}(t)}{dt} * \frac{\partial g\left(\boldsymbol{X}_i, \boldsymbol{w}(t)\right)}{\partial \boldsymbol{w}}. \tag{2}$$

The gradient flow can be described as:

$$\frac{d\boldsymbol{w}(t)}{dt} = -\nabla \mathcal{L}(\boldsymbol{w}(t)) = -\sum_{i=1}^{n} \left(g\left(\boldsymbol{X}_i, \boldsymbol{w}(t)\right) - \boldsymbol{Y}_i\right) \frac{\partial g\left(\boldsymbol{X}_i, \boldsymbol{w}(t)\right)}{\partial \boldsymbol{w}}. \tag{3}$$

Substitute Equation (3) into Equation (2), we have:

$$\frac{dg\left(\boldsymbol{X}_i, \boldsymbol{w}(t)\right)}{dt} = -\sum_{j=1}^{n} g\left(\boldsymbol{X}_i, \boldsymbol{w}(t)\right) \boxed{\left\langle \frac{\partial g\left(\boldsymbol{X}_i, \boldsymbol{w}(t)\right)}{\partial \boldsymbol{w}}, \frac{\partial g\left(\boldsymbol{X}_j, \boldsymbol{w}(t)\right)}{\partial \boldsymbol{w}} \right\rangle}, \tag{4}$$

Theorem 1

(**Informal version**) When optimized by gradient descent, outputs of grid-based models evolve following an ordinary differentiable equation (ODE) related to the grid tangent kernel (GTK).

(**Formal version**) Let $\boldsymbol{O}(t) = (g(\boldsymbol{X}_i, \boldsymbol{w}(t)))_{1\leq i\leq n}$ be the outputs of a grid-based model g where $\boldsymbol{X} = (\boldsymbol{X}_i)_{1\leq i\leq n}$ is the input data at time t, and $\boldsymbol{Y} = (\boldsymbol{Y}_i)_{1\leq i\leq n}$ is the corresponding label. Then $\boldsymbol{O}(t)$ follows this evolution:

$$\frac{d\boldsymbol{O}(t)}{dt} = -\boldsymbol{G}_g(t)\cdot(\boldsymbol{O}(t)-\boldsymbol{Y}). \tag{5}$$

Theorem 2

(**Informal version**) The GTK of a grid-based model stays stationary during training.
(**Formal version**) The GTK of a grid-based model g, denoted by $\boldsymbol{G}_g$, stays stationary during training. Formally, this property can be written as:

$$\boldsymbol{G}_g(t) = \boldsymbol{G}_g(0), \tag{6}$$

where $\boldsymbol{G}_g(0)$ is the initial GTK of the grid-based model. This property holds for any size of the grid-based model.

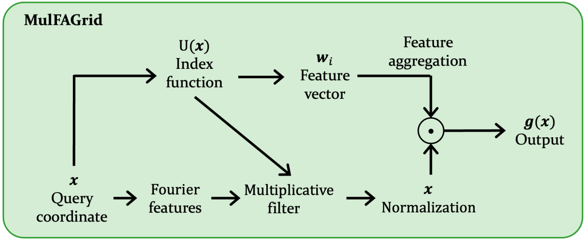
Theorem 3

(**Informal version**) The generalization bound of a grid-based model is determined by $\Delta = \boldsymbol{Y}^{\top}\boldsymbol{G}^{-1}\boldsymbol{Y}$.

(**Formal version**) Given a probability $\delta_p \in (0,1)$, suppose the dataset $S = (\boldsymbol{X}, \boldsymbol{Y})$ contains n i.i.d. samples from a distribution where $n \gg \log \frac{2}{\delta_p}$ and the minimum eigenvalue of the GTK, denoted by $\boldsymbol{G}$, is at least a constant λ_0: $\lambda_{min}(\boldsymbol{G}) \geq \lambda_0$. For any grid-based model g that is optimized by gradient descent with a learning rate η_l, and for any loss function $\mathcal{L} : \mathbb{R} \times \mathbb{R} \to [0,1]$, which is 1-Lipschitz in the first argument, we define the population loss as $\mathcal{L}_{\mathcal{D}}(t) = \mathbb{E}_{(\boldsymbol{X}_i, \boldsymbol{Y}_i) \sim \mathcal{D}}\left[\mathcal{L}\left(g\left(\boldsymbol{X}_i, \boldsymbol{w}(t)\right), \boldsymbol{Y}_i\right)\right]$. Then, with probability at least $1 - \delta_p$, a randomly initialized grid-based model trained by gradient descent for $t \geq \Omega\left(\frac{1}{\eta_l \lambda_0} \log \frac{n}{\delta_p}\right)$ iterations has a generalization bound:

$$\mathcal{L}_{\mathcal{D}}(t) \leq \sqrt{\frac{2\boldsymbol{Y}^{\top}\boldsymbol{G}^{-1}\boldsymbol{Y}}{n}} + O\left(\sqrt{\frac{\log \frac{2}{\delta_p}}{n}}\right). \tag{7}$$

A new grid-based model: MulFAGrid

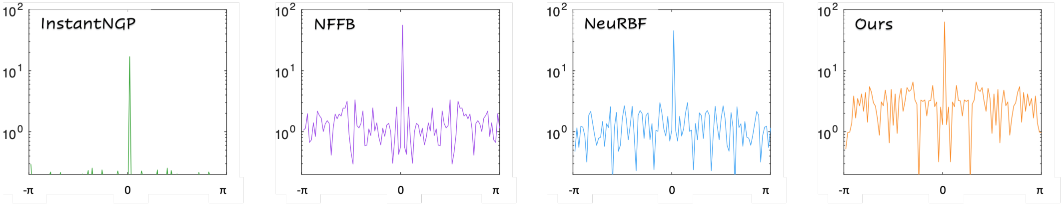


We use Fourier features[§] to boost the learning of high-frequency signals, and we adopt multiplicative filters[¶] to inform the model with node information.

[§]Matthew Tancik et al. "Fourier features let networks learn high frequency functions in low dimensional domains". In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 7537–7547.

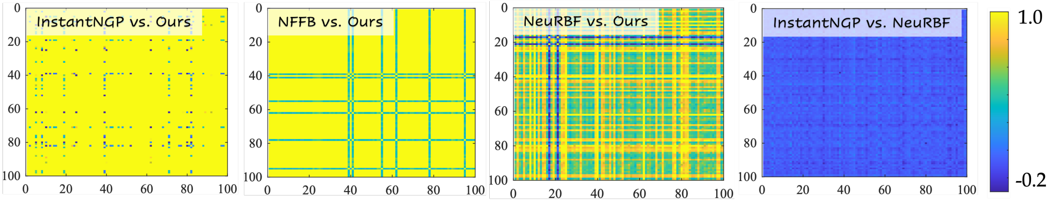
[¶]Rizal Fathony et al. "Multiplicative filter networks". In: *International Conference on Learning Representations*. 2020.

GTK-based analysis - (a) the GTK Fourier spectrums



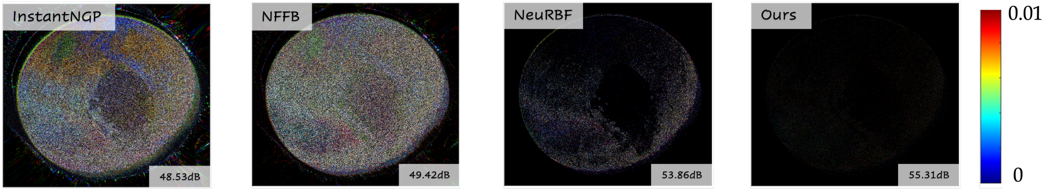
MulFAGrid has a wide spectrum, especially in the high-frequency domain, leading to faster convergence for high-frequency components.

GTK-based analysis - (b) the generalization bounds



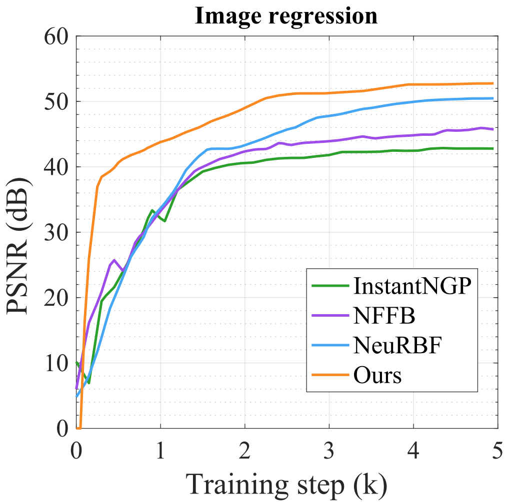
In this experiment, we construct a dataset, which only contains two data points with labels $\boldsymbol{Y} = (\boldsymbol{Y}_1, \boldsymbol{Y}_2)$, shown in the x-axis and y-axis correspondingly. MulFAGrid has a tighter (lower) generalization bound for most values of $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$.

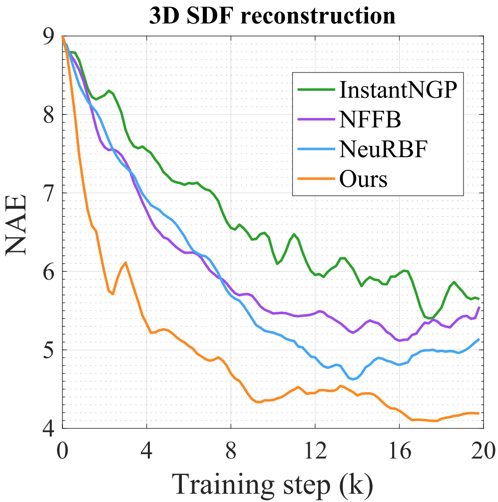
GTK-based analysis - (c) the error maps



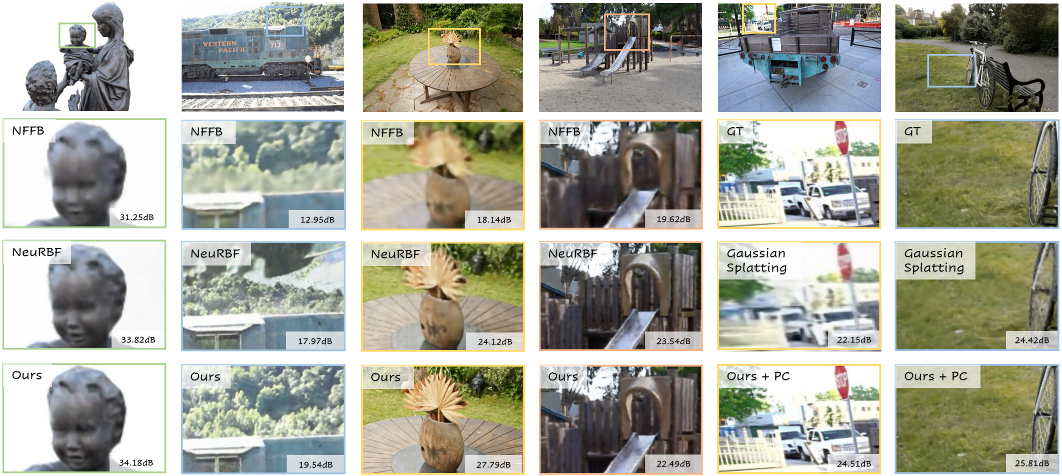
Error maps of the fitted images in comparison with ground truth ones.

2D image / 3D SDF fitting





Novel view synthesis



- Fathony, Rizal et al. "Multiplicative filter networks". In: *International Conference on Learning Representations*. 2020.
- Jacot, Arthur, Franck Gabriel, and Clément Hongler. "Neural tangent kernel: Convergence and generalization in neural networks". In: *Advances in neural information processing systems* 31 (2018).
- Mildenhall, Ben et al. "Nerf: Representing scenes as neural radiance fields for view synthesis". In: *European Conference on Computer Vision*. Springer. 2020, pp. 405–421.
- Tancik, Matthew et al. "Fourier features let networks learn high frequency functions in low dimensional domains". In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 7537–7547.

Thank You

Zelin Zhao[1], Fenglei Fan[2‖], Wenlong Liao[1,3], Junchi Yan[1]
[1] Shanghai Jiao Tong University
[2] The Chinese University of Hong Kong
[3] Cowa Tech. Ltd · Grounding and Enhancing Grid-based Models for Neural Fields