Simplicity Bias

"NN are fundamentally Bayesian and DNN are biased, at init, towards simple functions"
(Chris Mingard)

Fix a DNN and be Θ space of its parameters, each point $\theta \in \Theta$ is a complete set of weight and biases of the network

$$F : \Theta \longrightarrow \mathcal{F}$$

↳ space of Function

$f \in F$ is a specific function that the model can learn

→ Complexity of f : $C(f) \in [0, +\infty)$

$$\Theta \xrightarrow{F} \mathcal{F} \xrightarrow{C} [0, +\infty)$$

Given a P on Θ, probability measure, we can pick θ at random according to the measure P (RANDOM INIT OF NETWORK)

We can also think of F as a random variable and indeed $C(F)$. Then, simplicity bias

$$C(f) < C(g) \Rightarrow P(F=f) > P(F=g)$$

"POINTWISE"