Out-of-Distribution Generalization

$r=1$ $r=2$ $r=5$

MSE (y-axis), Number of training samples (n) (x-axis)

Legend:
- $\sigma=0$
- $\sigma=0.25$
- $\sigma=0.5$
- $\sigma=1$
- without extra layers
- with extra layers
- σ^2, irreducible error