

# Supplementary Material for Investigating Learning in Deep Neural Networks using Layer-Wise Weight Change

November 2, 2020

## 1 RWC Plots

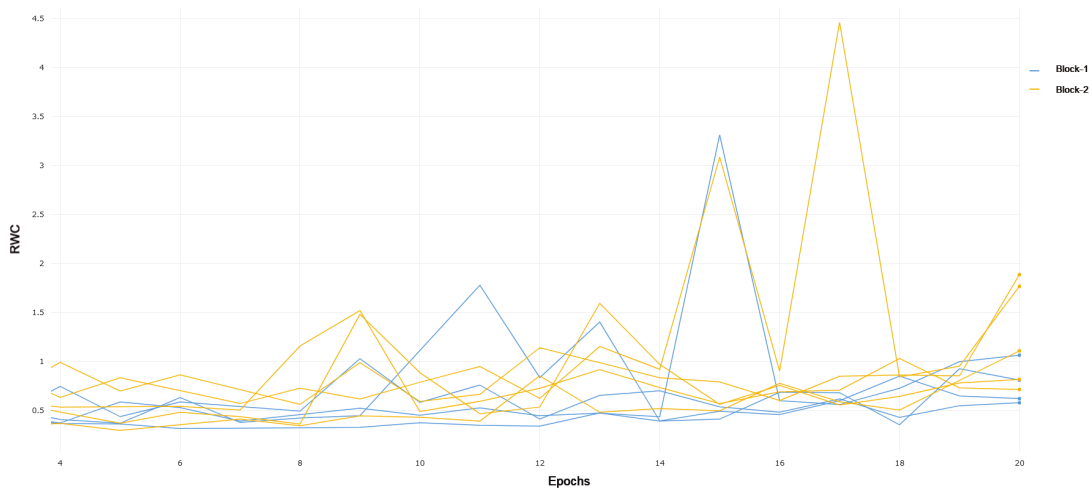


Figure 1: RWC for Resnet18 Blocks 1-2 on MNIST

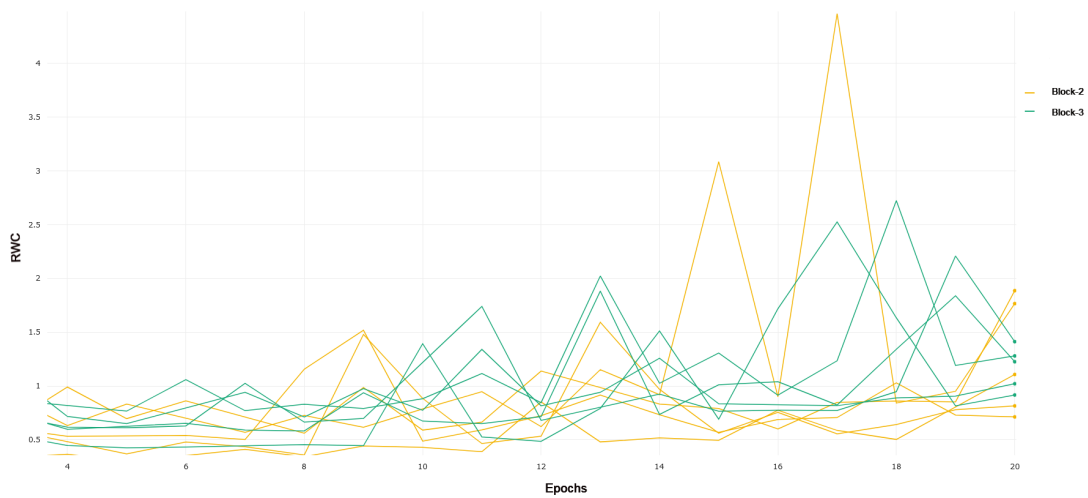


Figure 2: RWC for Resnet18 Blocks 2-3 on MNIST

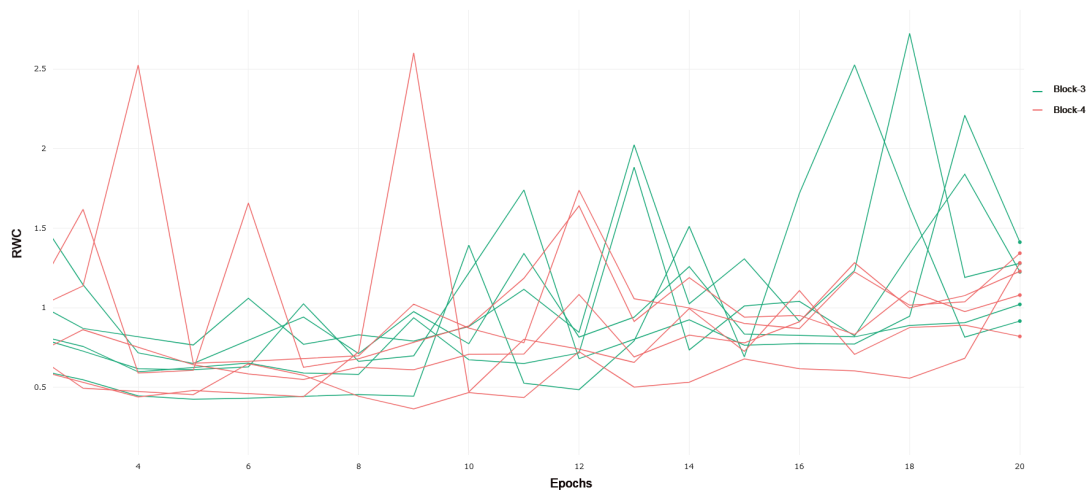


Figure 3: RWC for Resnet18 Blocks 3-4 on MNIST

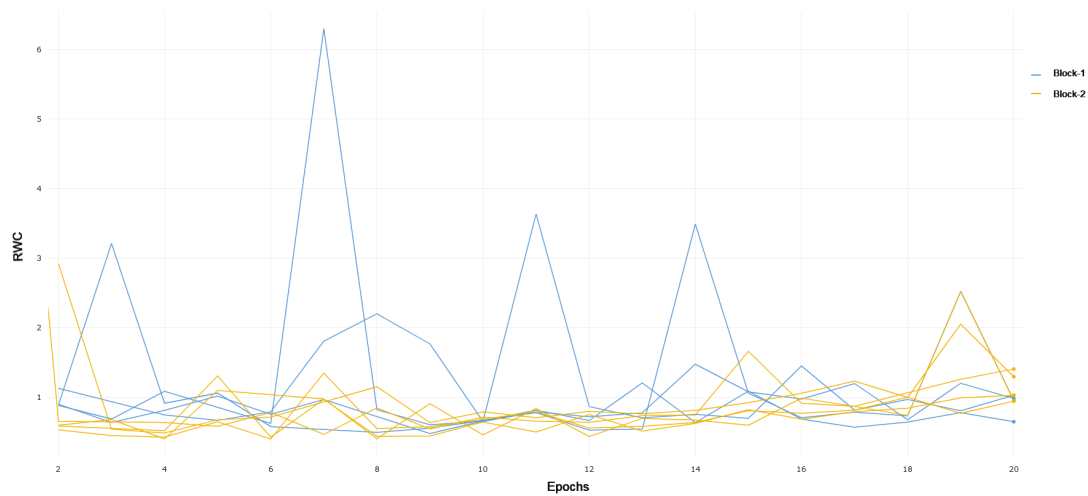


Figure 4: RWC for ResNet-18 Blocks 1-2 on FashionMNIST

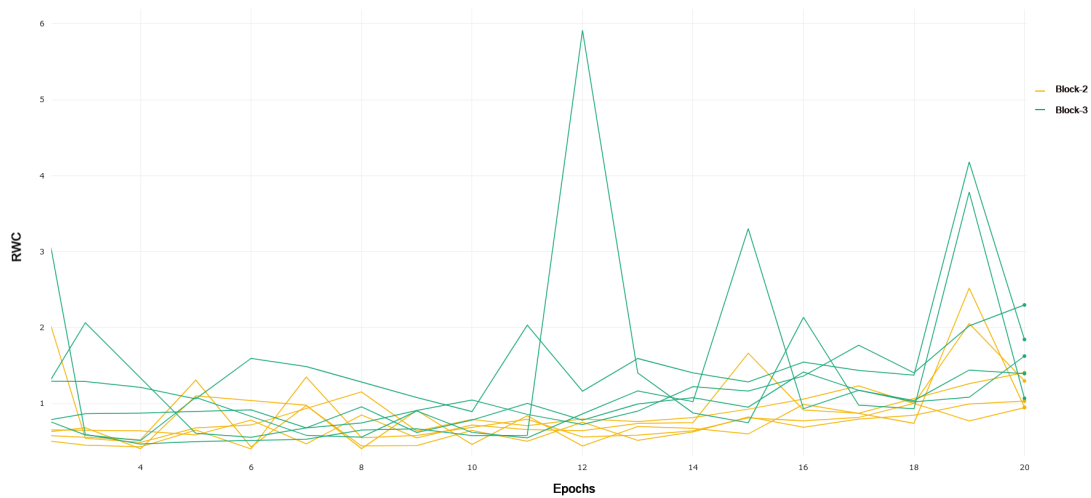


Figure 5: RWC for ResNet-18 Blocks 2-3 on FashionMNIST

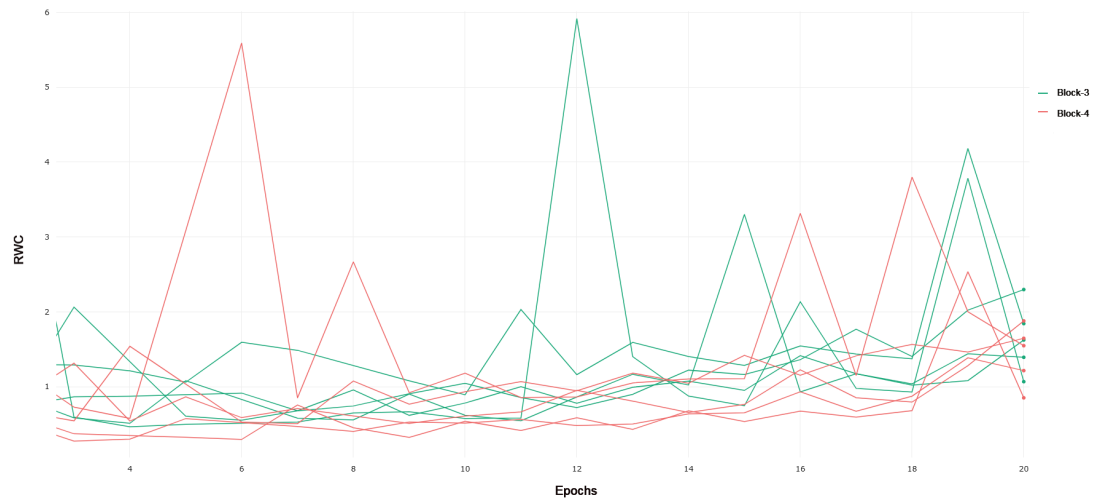


Figure 6: RWC for ResNet-18 Blocks 3-4 on FashionMNIST

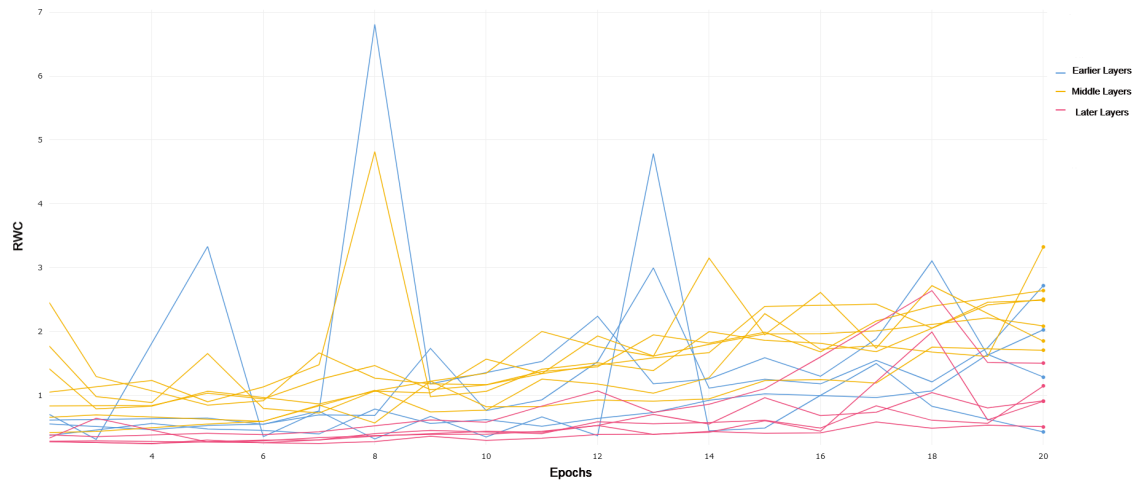


Figure 7: RWC for VGG on MNIST

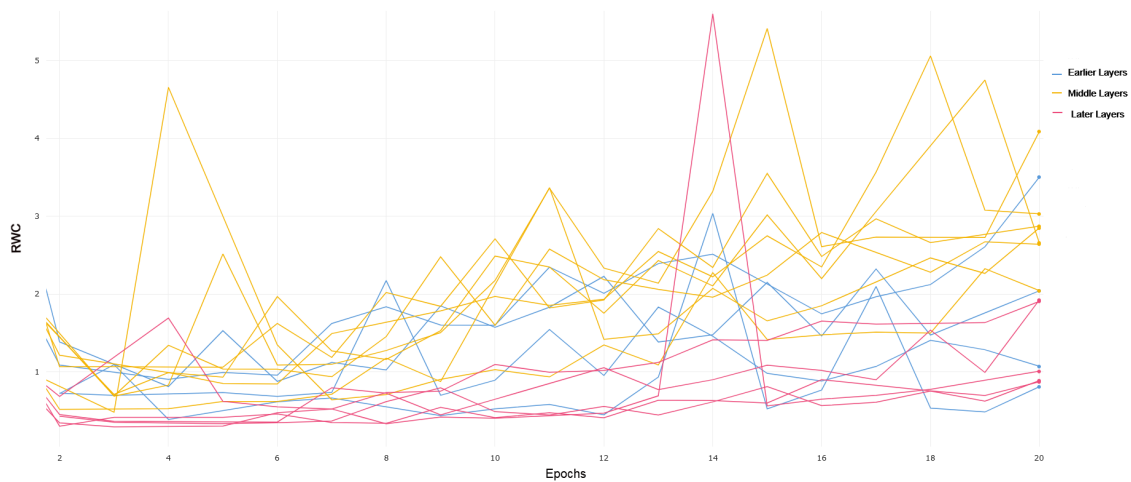


Figure 8: RWC for VGG on FashionMNIST

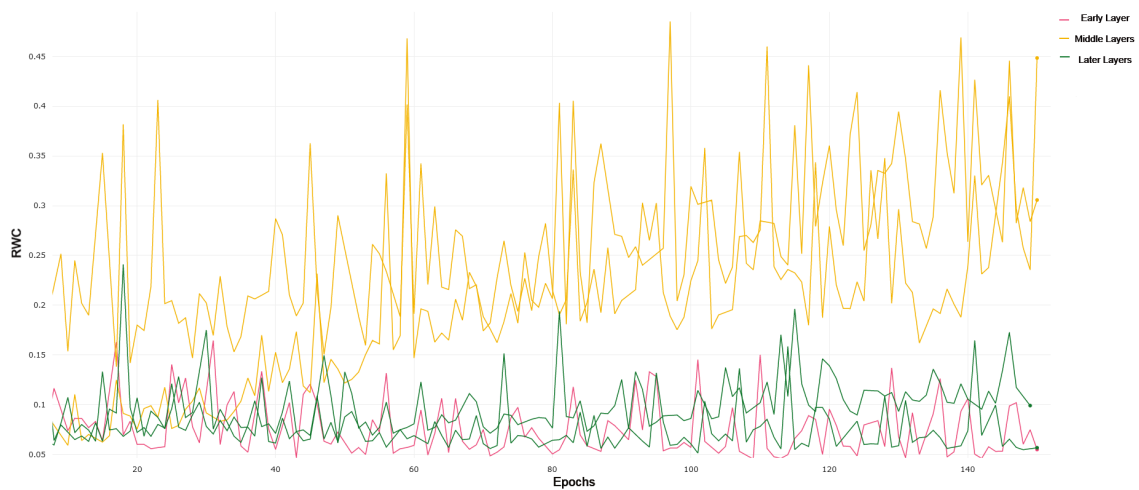


Figure 9: RWC for Alexnet on CIFAR10

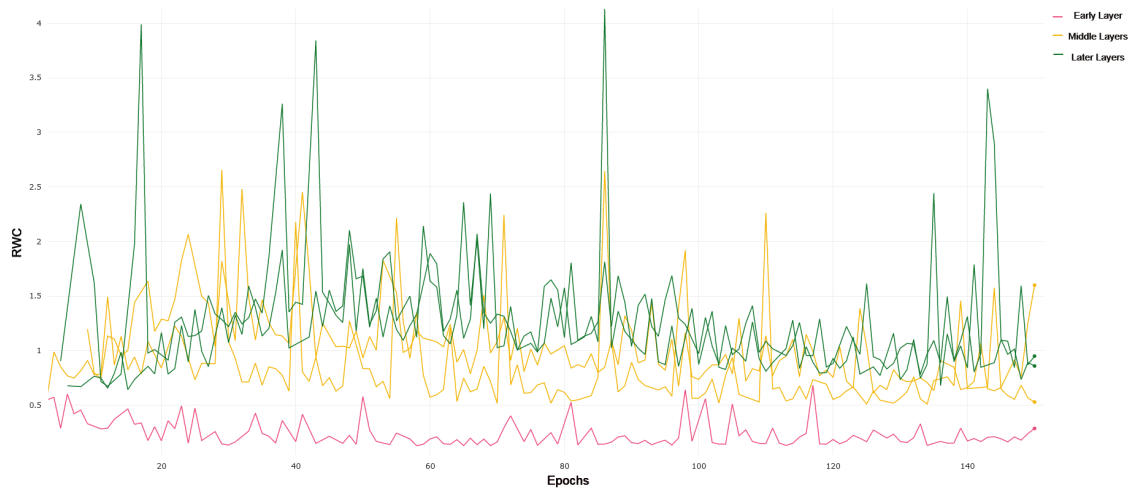


Figure 10: RWC for Alexnet on CIFAR100

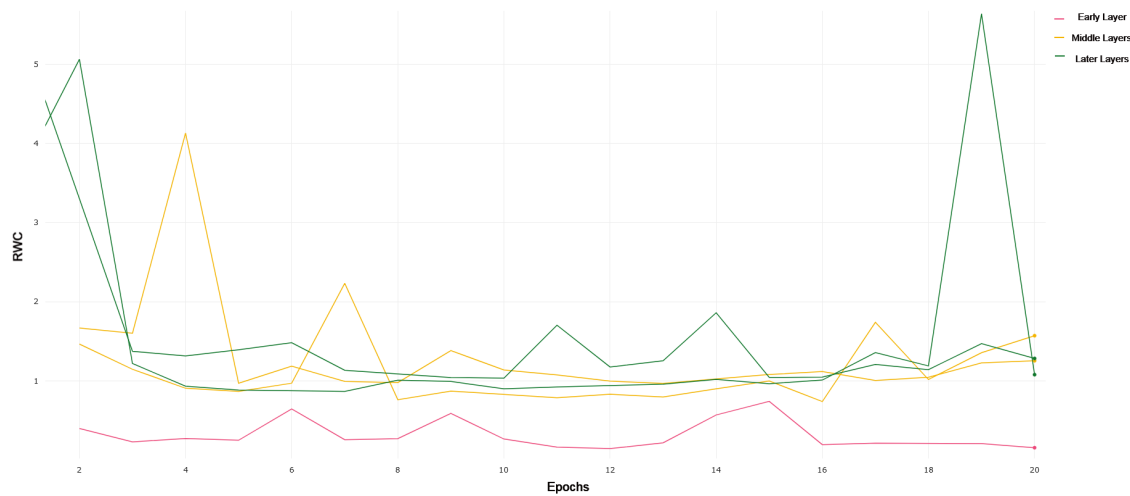


Figure 11: RWC for Alexnet on MNIST

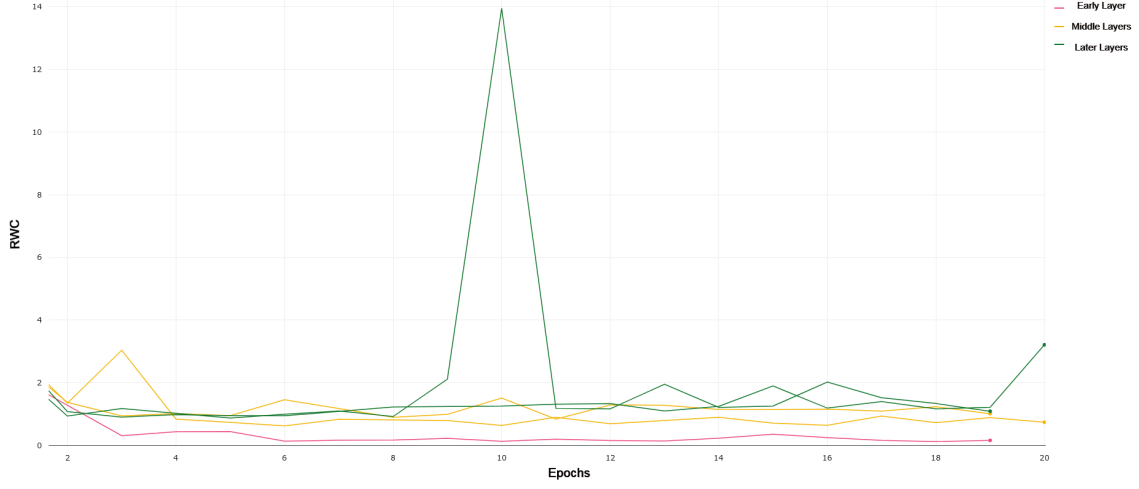


Figure 12: RWC for Alexnet on FashionMNIST

Table 1: Detailed hyperparameters used for training

Architecture	Datasets	LR	Momentum	Weight Decay
ResNet18	CIFAR-10	0.1	0.9	0.0001
	CIFAR-100	0.1	0.9	0.0001
	MNIST	0.1	0.9	0.0001
	FMNIST	0.1	0.9	0.0001
VGG19_bn	CIFAR-10	0.05	0.9	0.0005
	CIFAR-100	0.05	0.9	0.0005
	MNIST	0.05	0.9	0.0005
	FMNIST	0.05	0.9	0.0005
AlexNet	CIFAR-10	0.001	0.9	0.0001
	CIFAR-100	0.01	0.9	0.0001
	MNIST	0.1	0.9	0.0001
	FMNIST	0.1	0.9	0.0001