
THE LOSS SURFACES OF NEURAL NETWORKS WITH GENERAL ACTIVATION FUNCTIONS: SUPPLEMENTARY MATERIAL

Nicholas P Baskerville
School of Mathematics
University of Bristol
Bristol, UK
n.p.baskerville@bristol.ac.uk

April 9, 2020

This document contains figures from experiments in Section 2.4 of [Bas+20].

References

- [Bas+20] Nicholas P. Baskerville et al. *The Loss Surfaces of Neural Networks with General Activation Functions*. 2020. arXiv: 2004.03959 [math.PR].

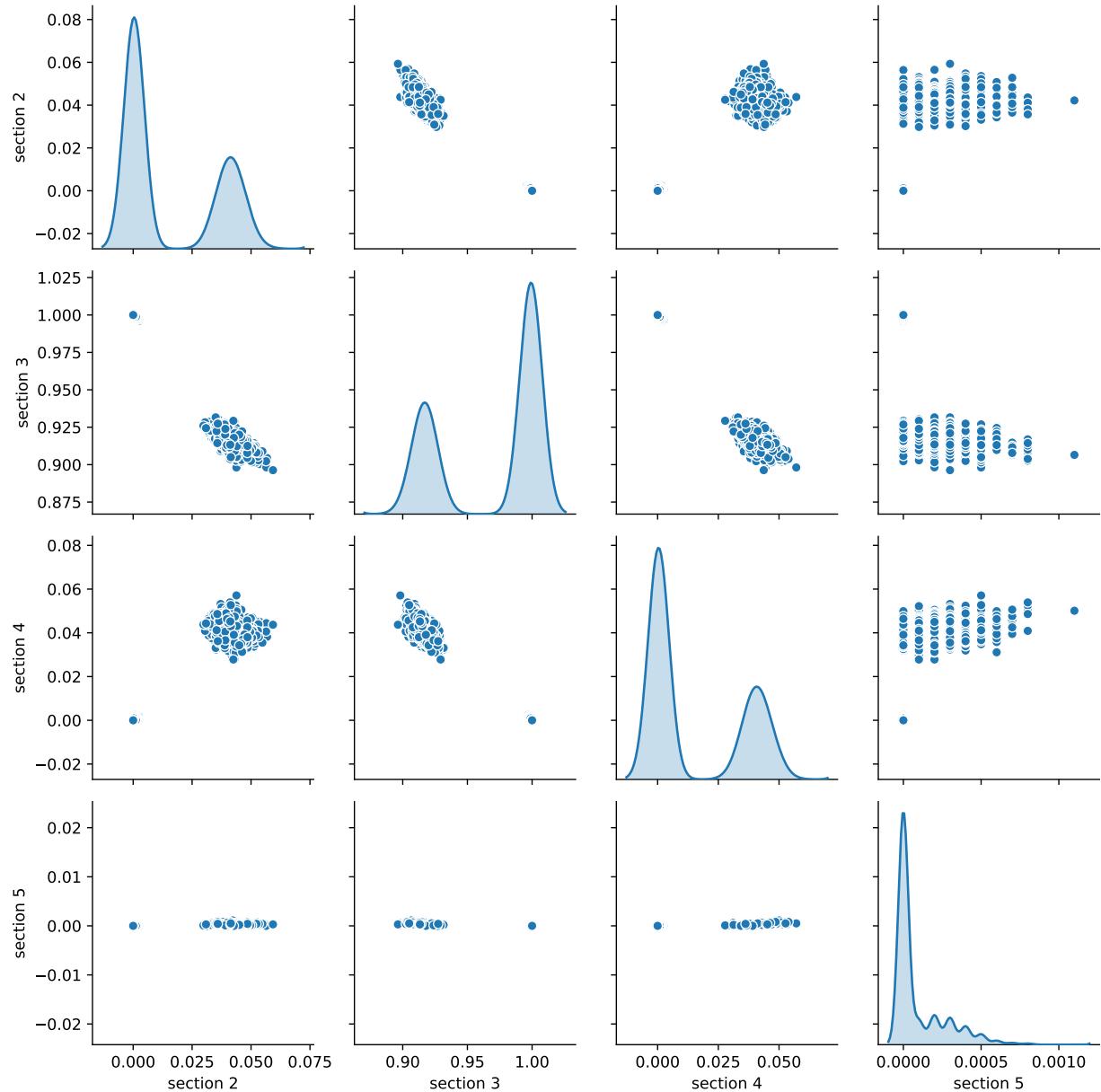


Figure 1: Experimental distribution of (R_2, R_3, R_4, R_5) (data averaging; each sample is a single neuron) for random MLP 5-piece network, and i.i.d. normal data.

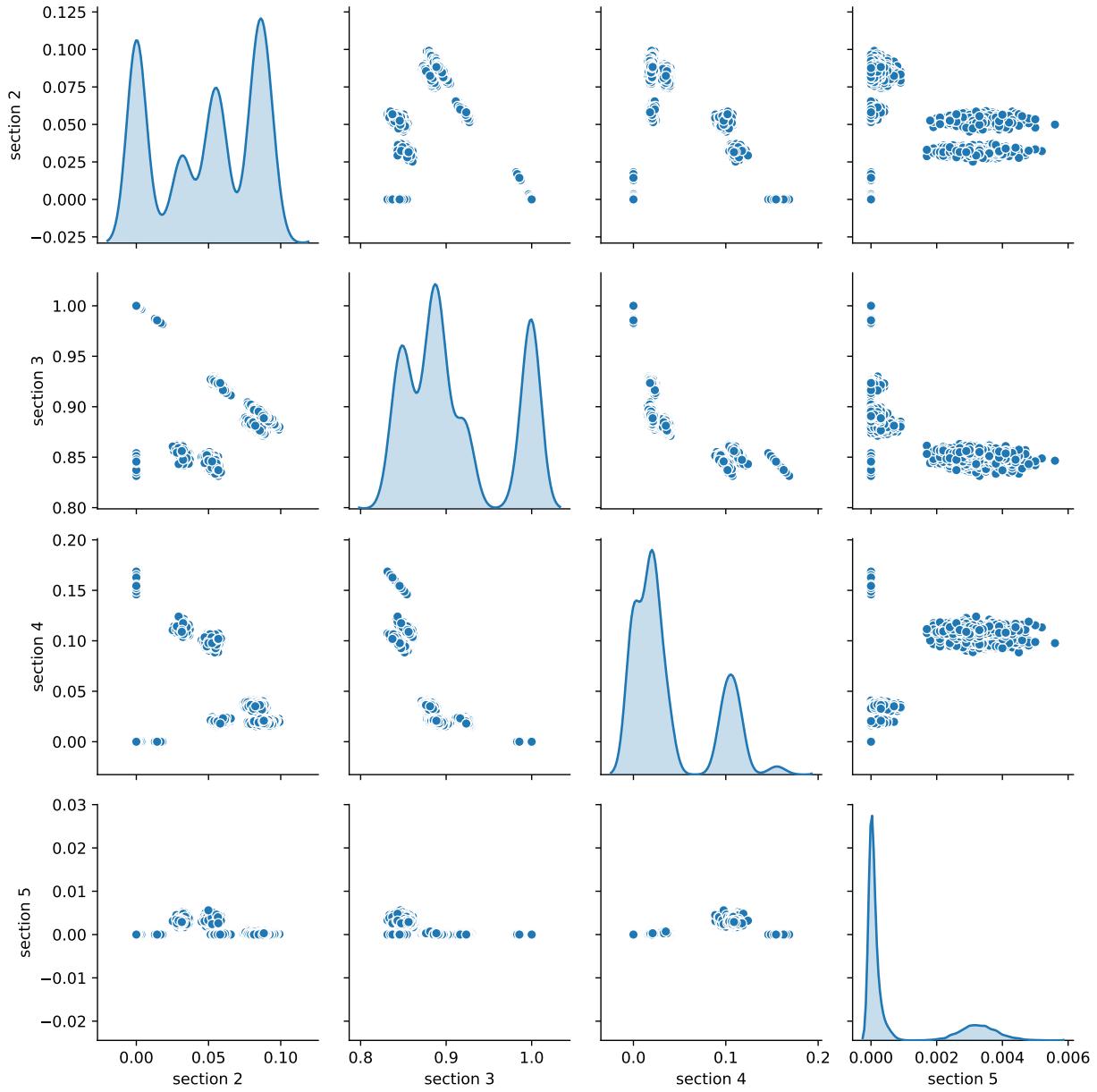


Figure 2: Experimental distribution of (R_2, R_3, R_4, R_5) (data averaging; each sample is a single neuron) for random LeNet 5-piece network, and i.i.d. normal data.

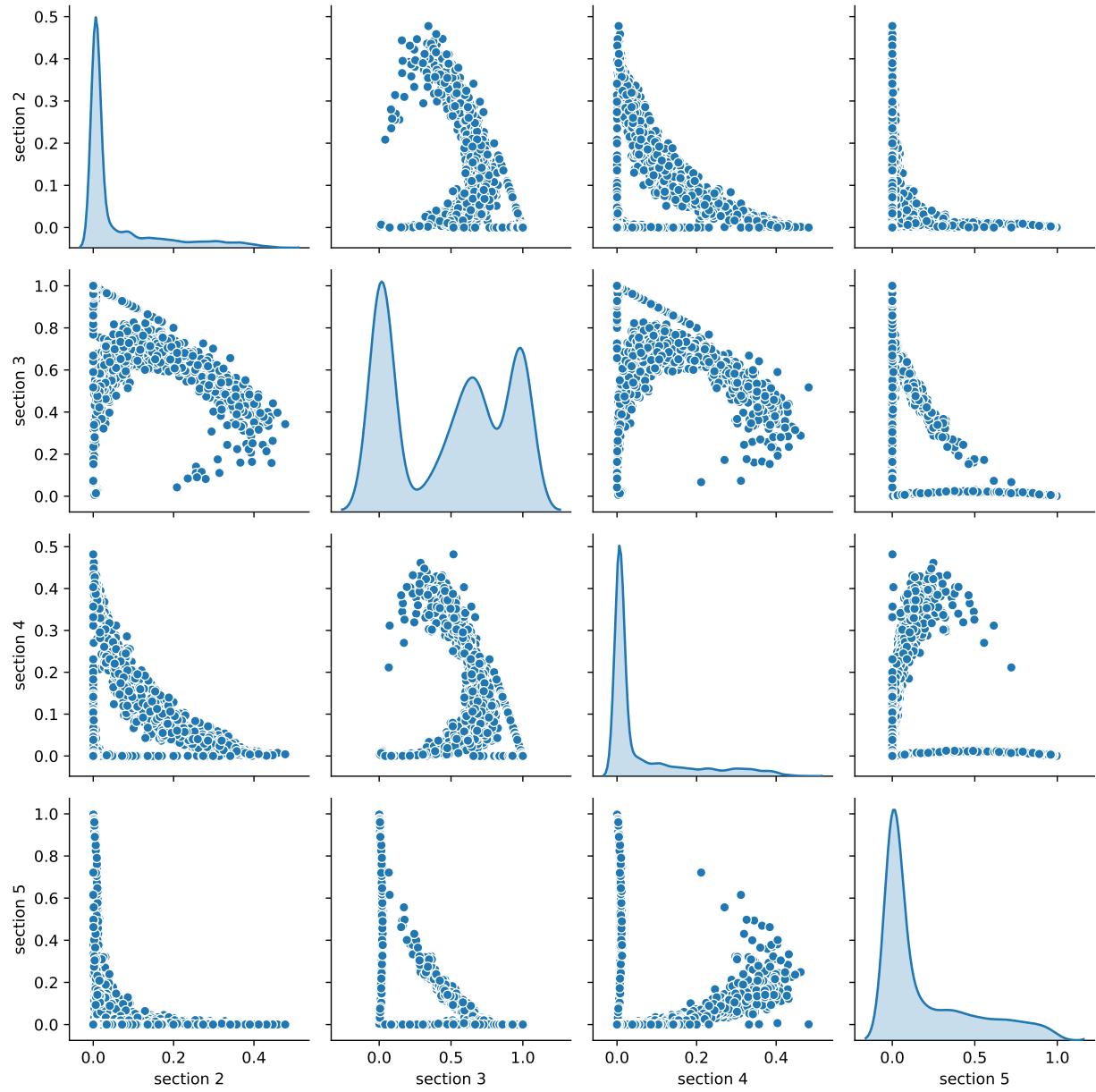


Figure 3: Experimental distribution of (R_2, R_3, R_4, R_5) (data averaging; each sample is a single neuron) for random MLP 5-piece network, and MNIST data.

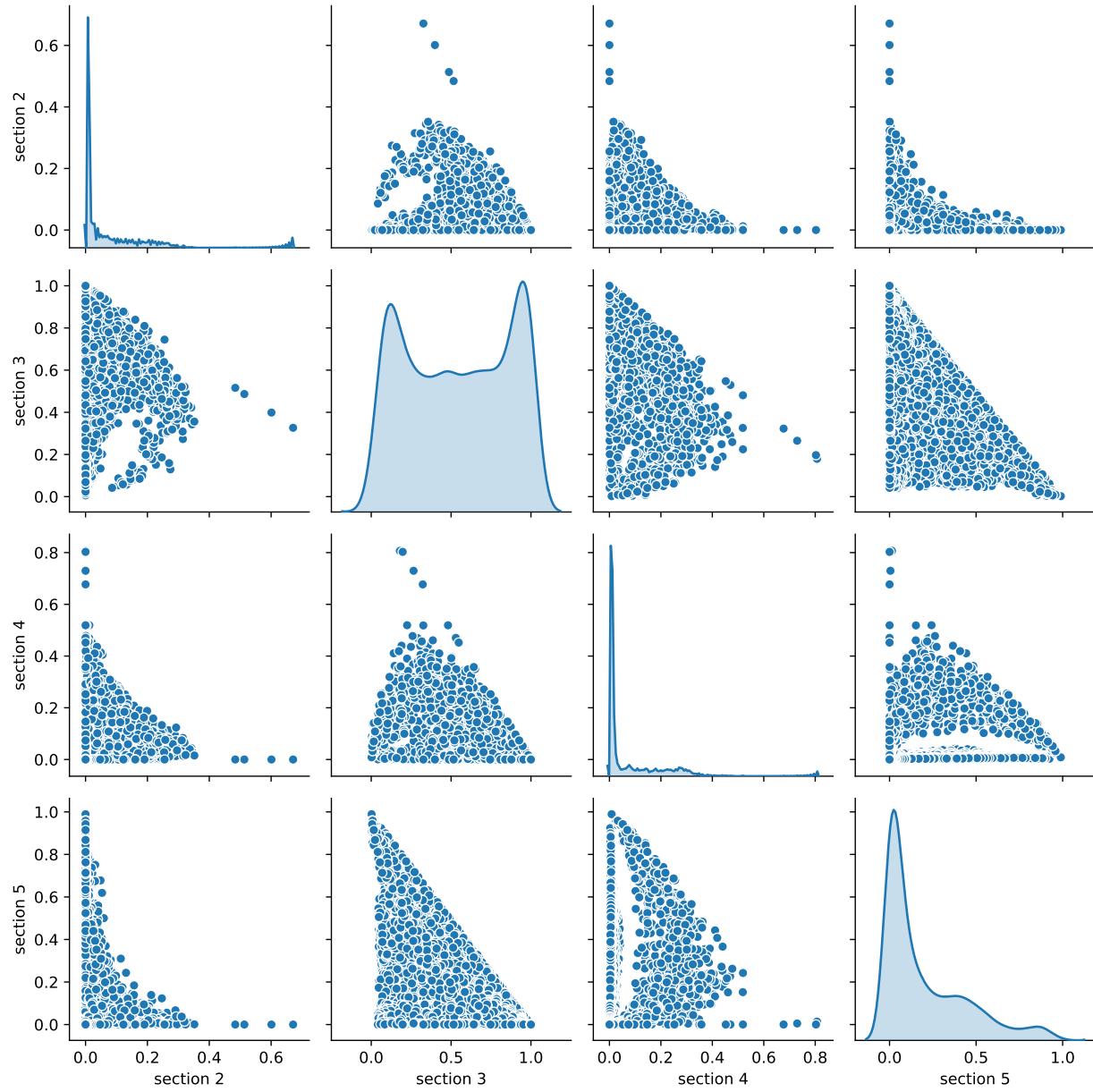


Figure 4: Experimental distribution of (R_2, R_3, R_4, R_5) (data averaging; each sample is a single neuron) for random LeNet 5-piece network, and MNIST data.

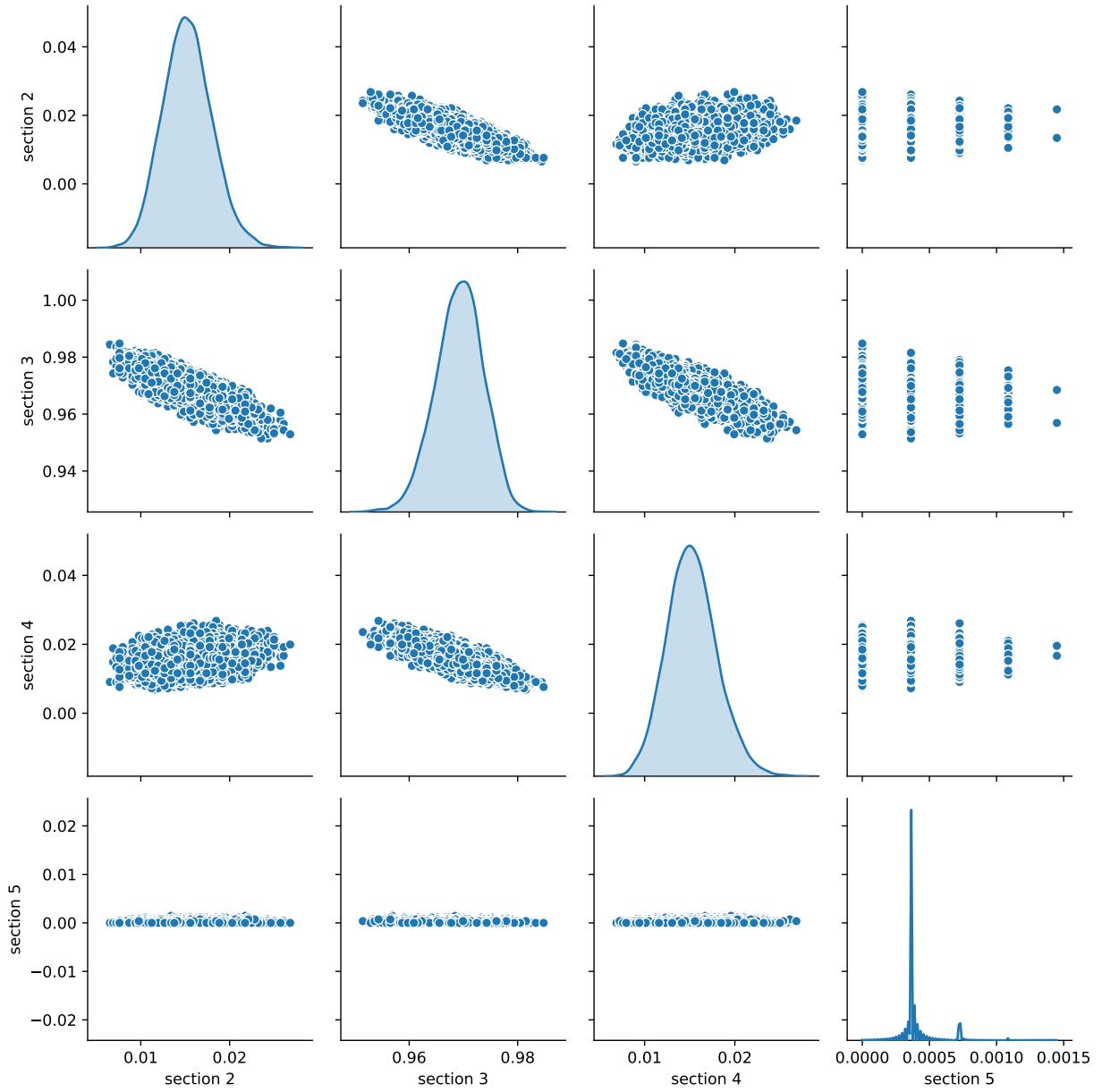


Figure 5: Experimental distribution of $(\bar{R}_2, \bar{R}_3, \bar{R}_4, \bar{R}_5)$ (neuron averaging; each sample is a single datum) for random 5-piece MLP network and i.i.d. normal data.

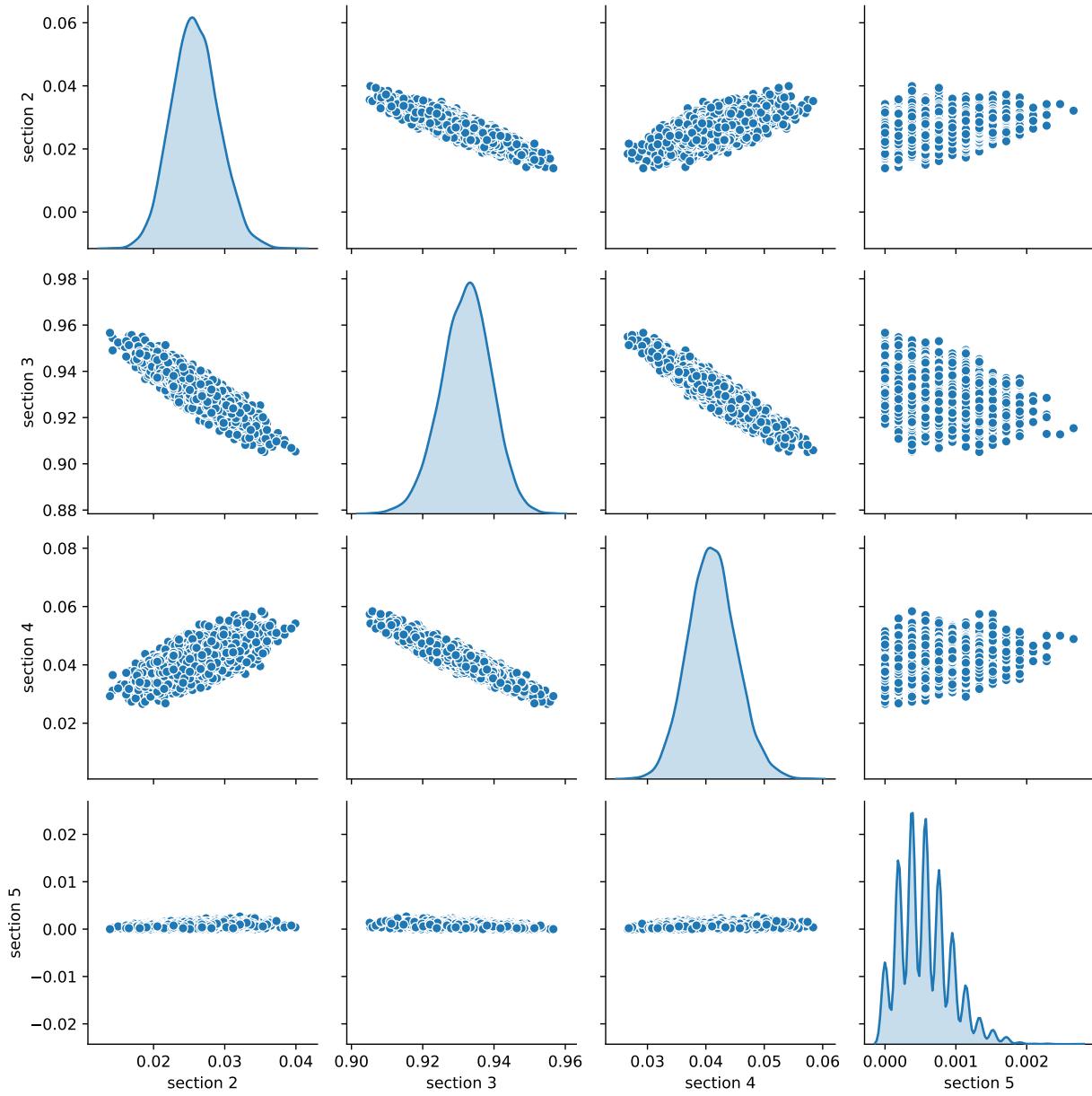


Figure 6: Experimental distribution of $(\bar{R}_2, \bar{R}_3, \bar{R}_4, \bar{R}_5)$ (neuron averaging; each sample is a single datum) for random 5-piece LeNet network and i.i.d. normal data.

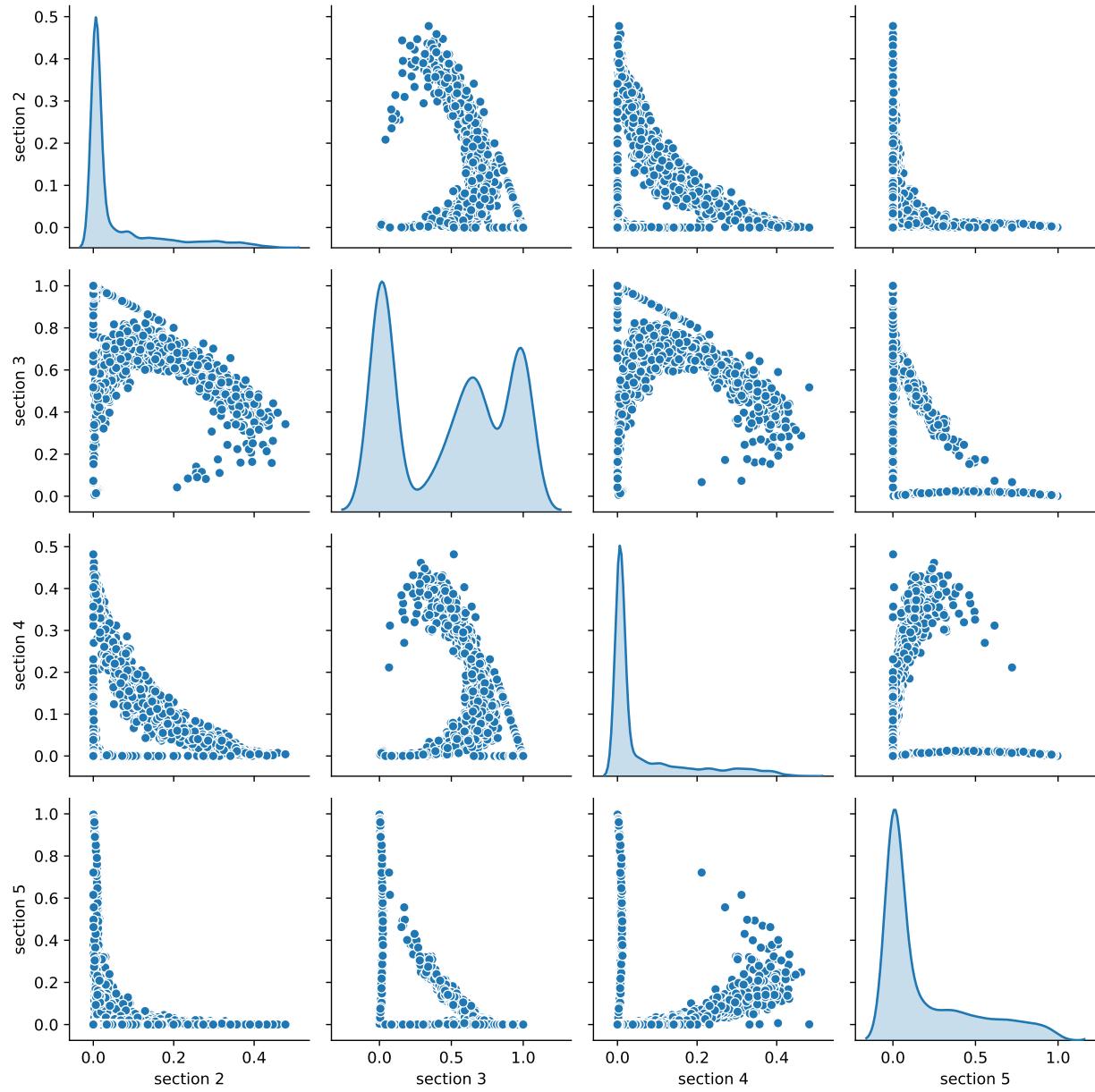


Figure 7: Experimental distribution of $(\bar{R}_2, \bar{R}_3, \bar{R}_4, \bar{R}_5)$ (neuron averaging; each sample is a single datum) for random 5-piece MLP network and MNIST data.

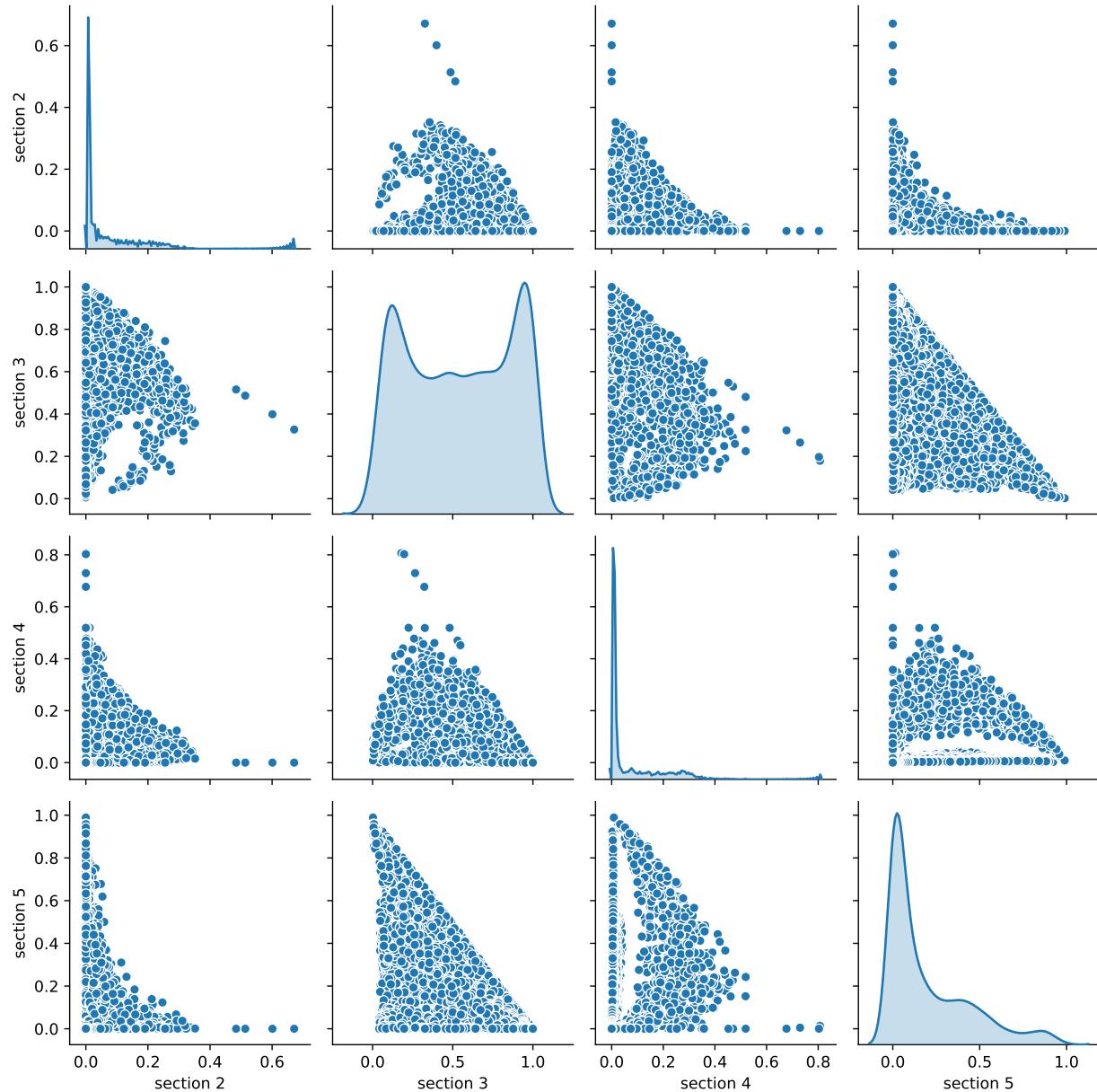


Figure 8: Experimental distribution of $(\bar{R}_2, \bar{R}_3, \bar{R}_4, \bar{R}_5)$ (neuron averaging; each sample is a single datum) for random 5-piece LeNet network and MNIST data.

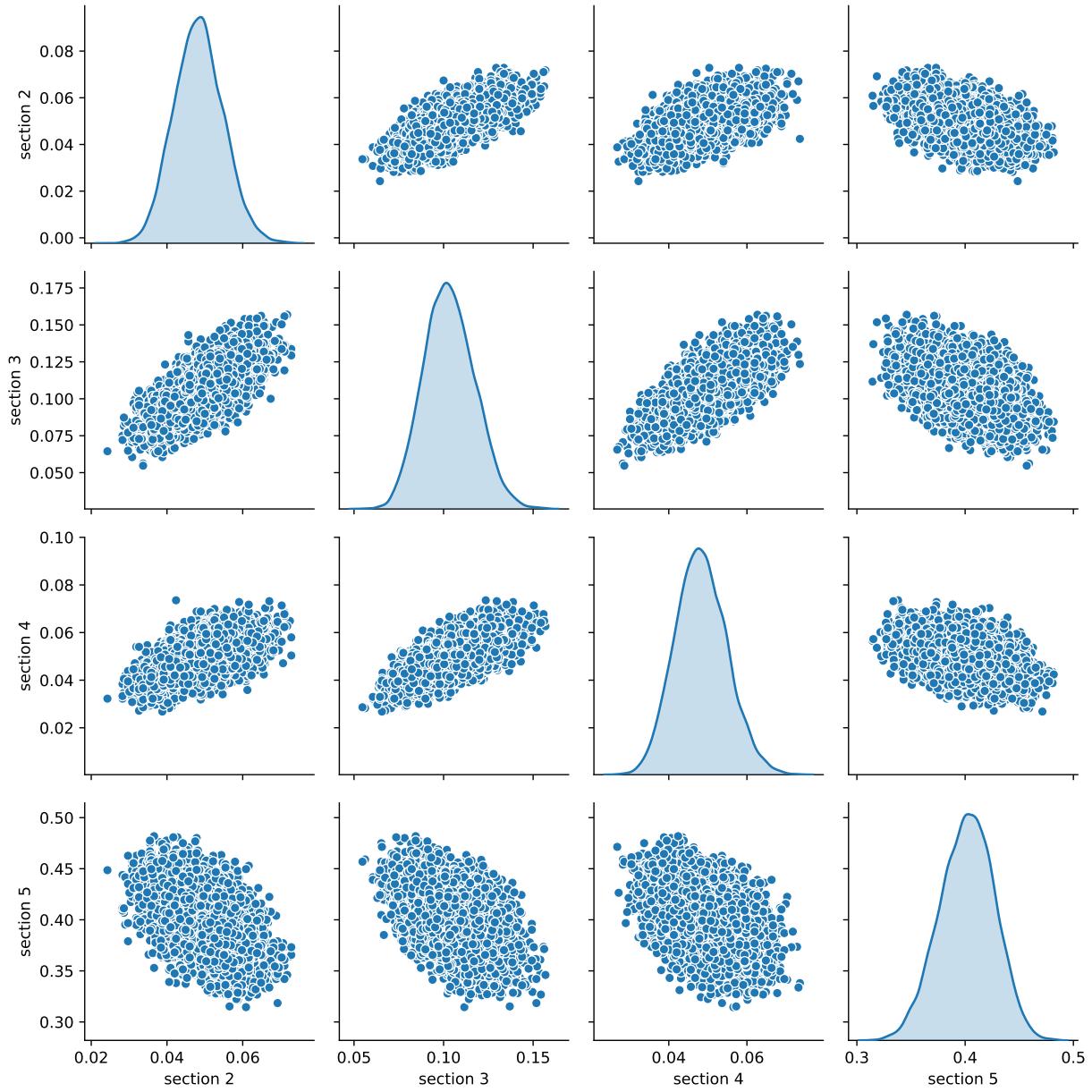


Figure 9: Experimental distribution of (R_2, R_3, R_4, R_5) (data averaging; each sample is a single neuron) for MLP 5-piece network trained to high validation accuracy on MNIST, and evaluated on i.i.d. normal data.

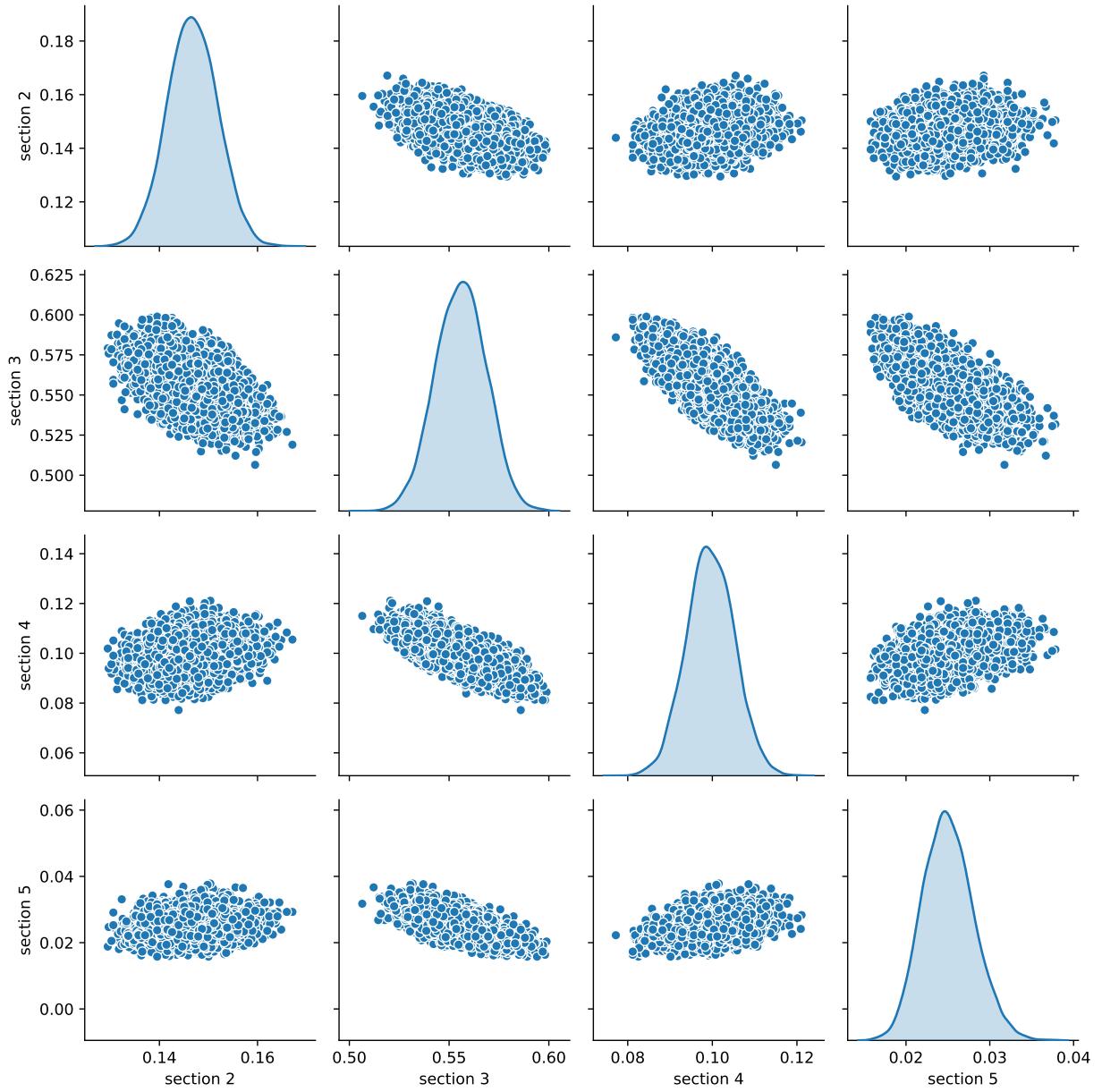


Figure 10: Experimental distribution of (R_2, R_3, R_4, R_5) (data averaging; each sample is a single neuron) for LeNet 5-piece network trained to high validation accuracy on MNIST, and evaluated on i.i.d. normal data.

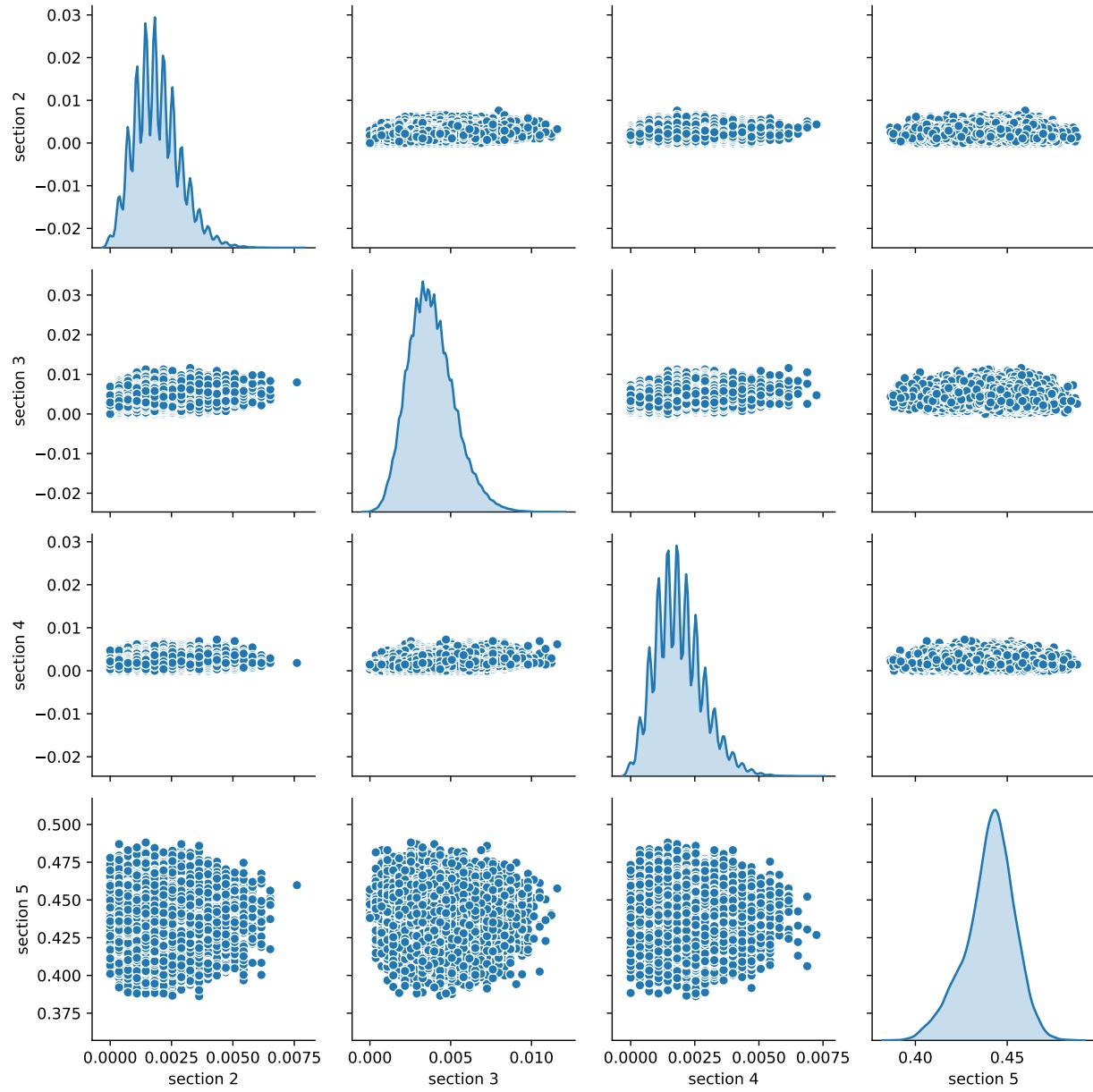


Figure 11: Experimental distribution of (R_2, R_3, R_4, R_5) (data averaging; each sample is a single neuron) for MLP 5-piece network trained to high validation accuracy on MNIST, and evaluated on MNIST data.

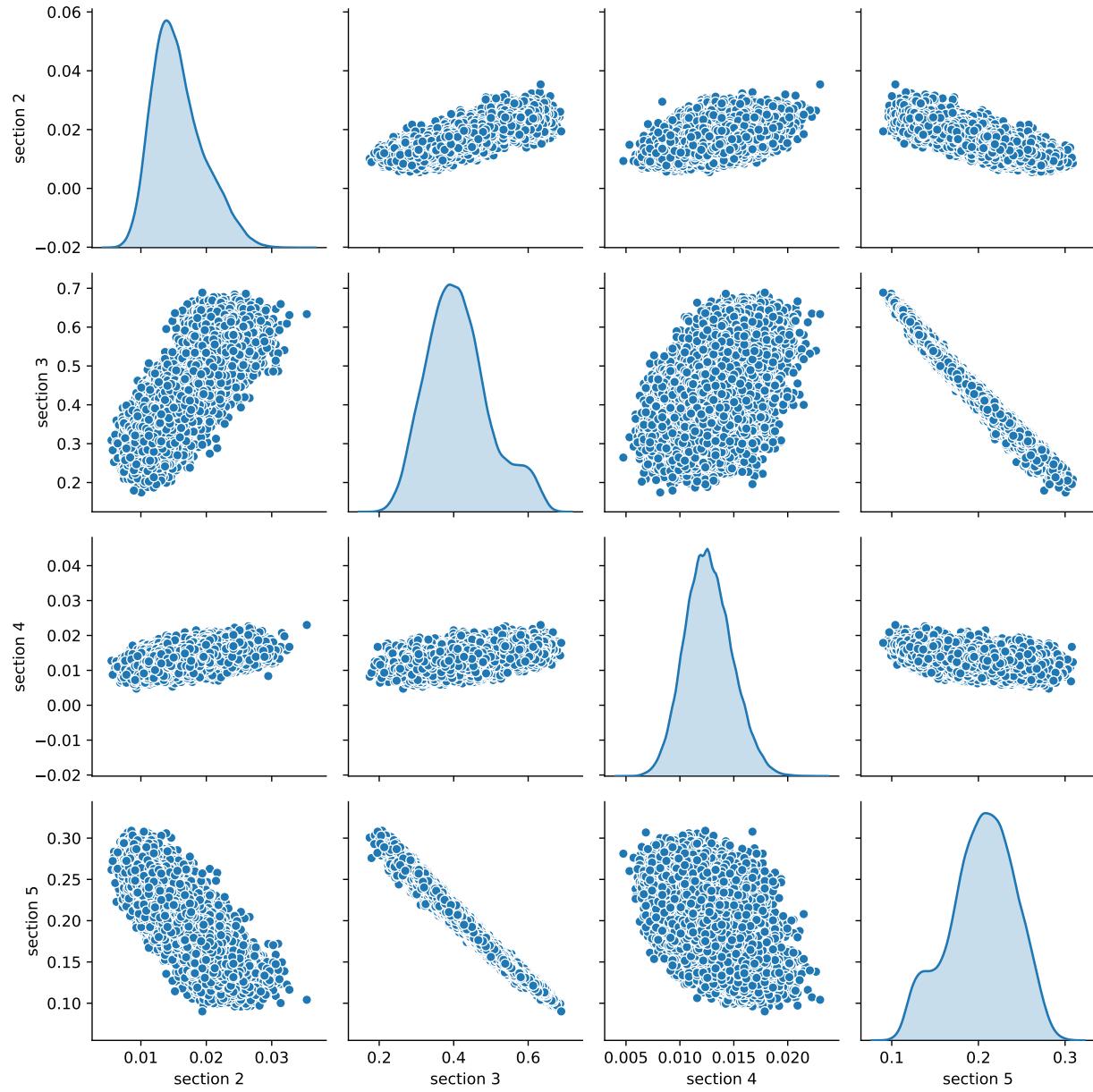


Figure 12: Experimental distribution of (R_2, R_3, R_4, R_5) (data averaging; each sample is a single neuron) for LeNet 5-piece network trained to high validation accuracy on MNIST, and evaluated on MNIST data.

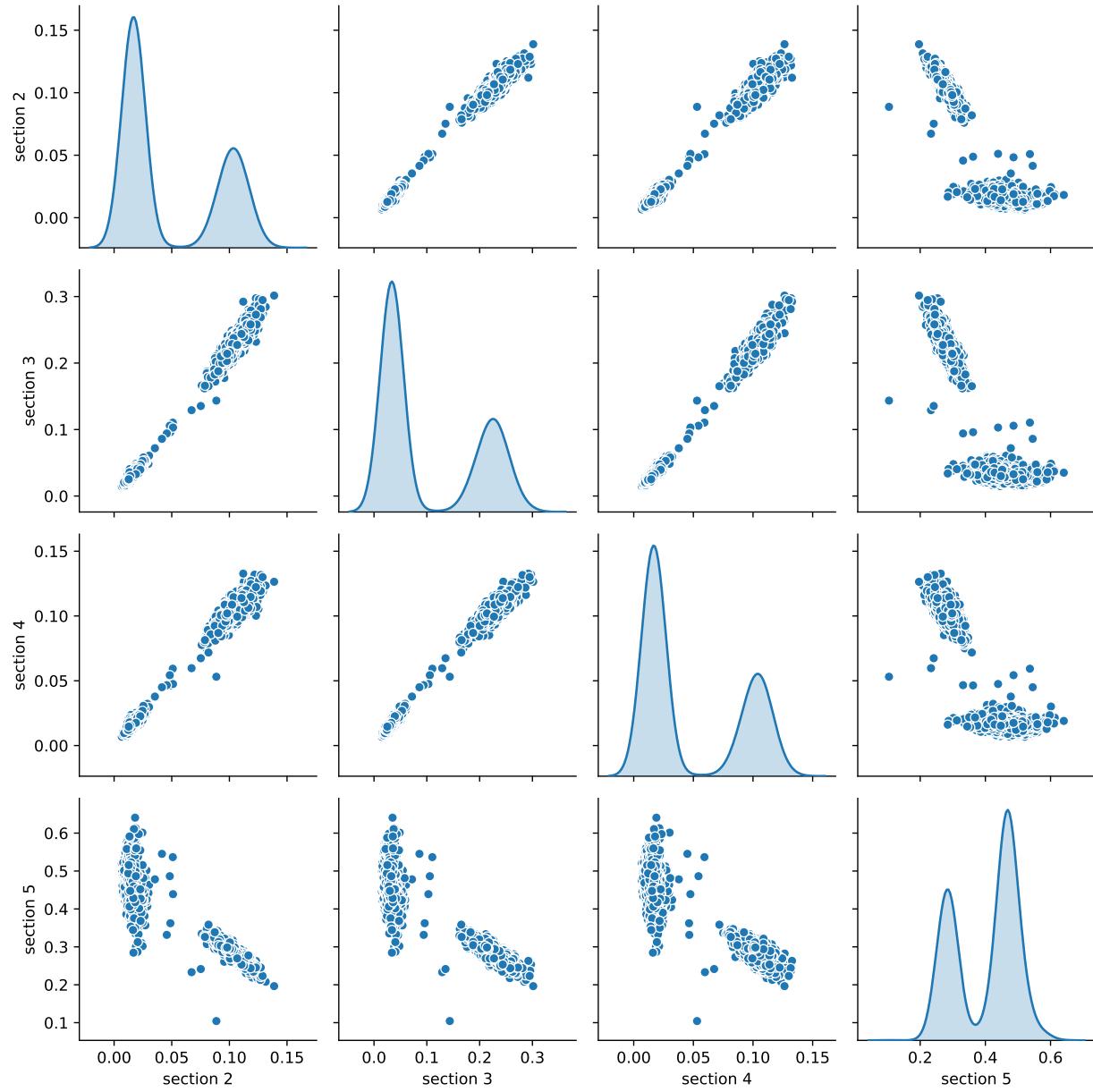


Figure 13: Experimental distribution of $(\bar{R}_2, \bar{R}_3, \bar{R}_4, \bar{R}_5)$ (neuron averaging; each sample is a single datum) for MLP 5-piece network trained to high validation accuracy on MNIST, and evaluated on i.i.d. normal data.

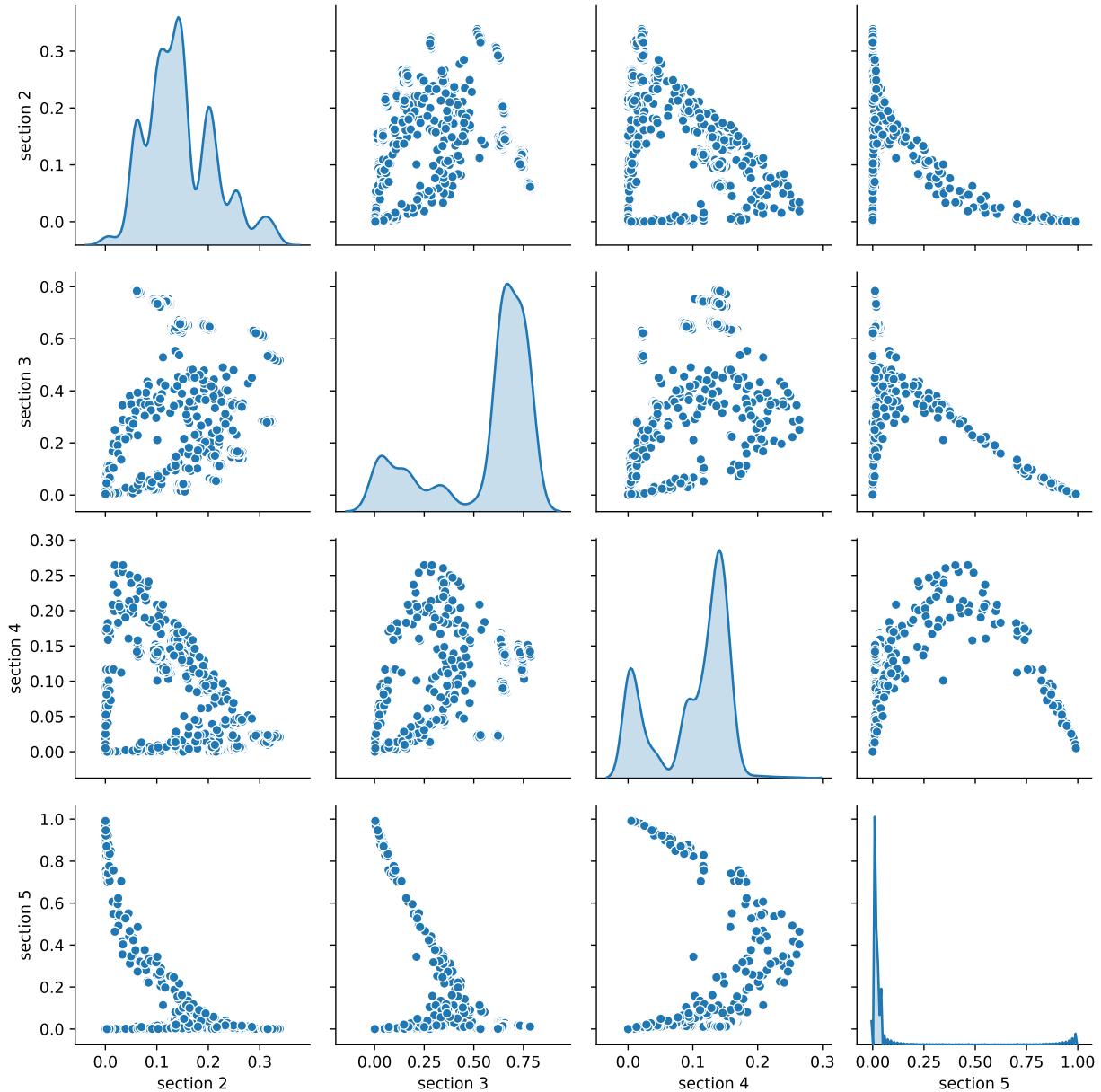


Figure 14: Experimental distribution of $(\bar{R}_2, \bar{R}_3, \bar{R}_4, \bar{R}_5)$ (neuron averaging; each sample is a single datum) for LeNet 5-piece network trained to high validation accuracy on MNIST, and evaluated on i.i.d. normal data.

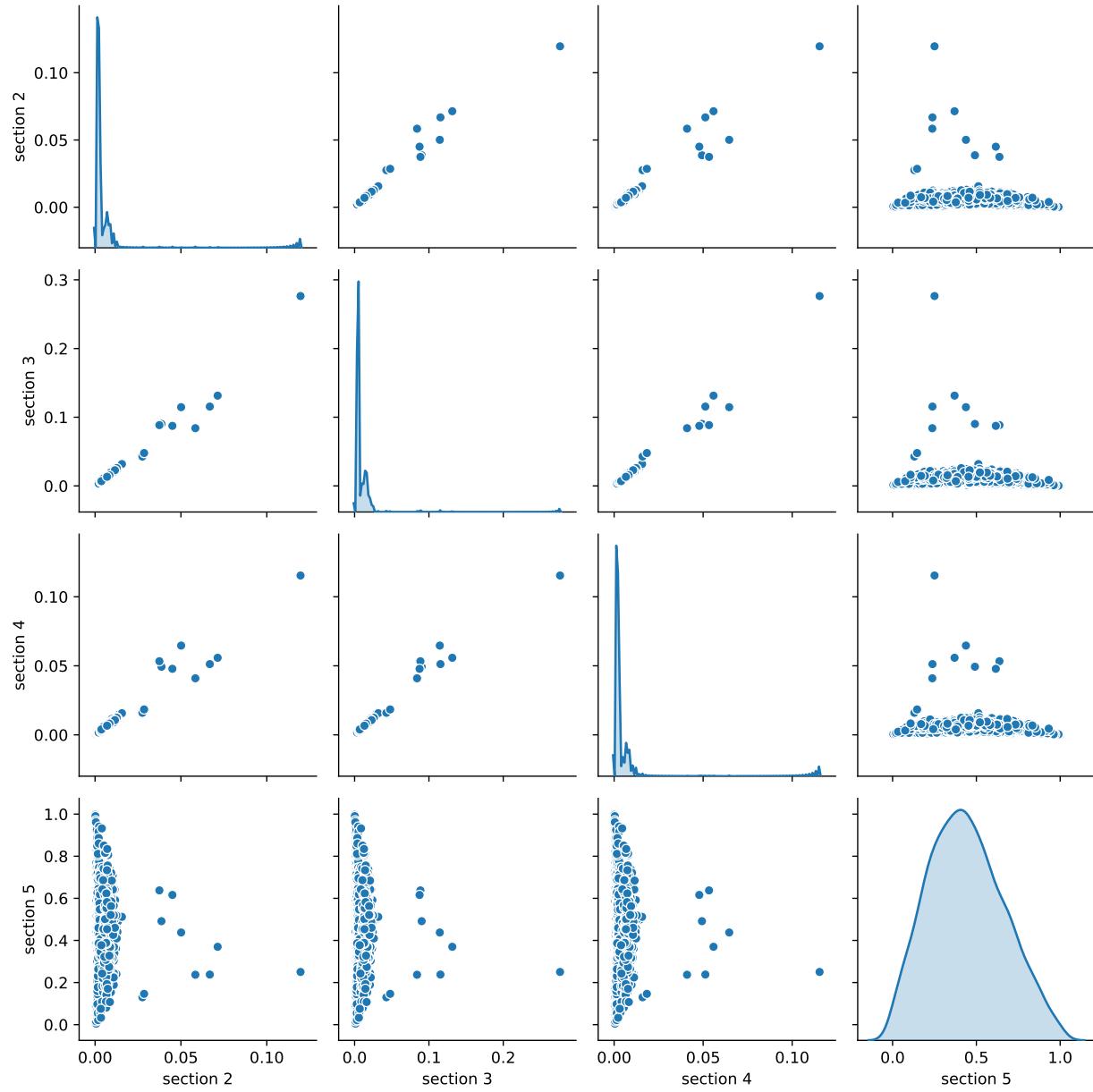


Figure 15: Experimental distribution of $(\bar{R}_2, \bar{R}_3, \bar{R}_4, \bar{R}_5)$ (neuron averaging; each sample is a single datum) for MLP 5-piece network trained to high validation accuracy on MNIST, and evaluated on MNIST data.

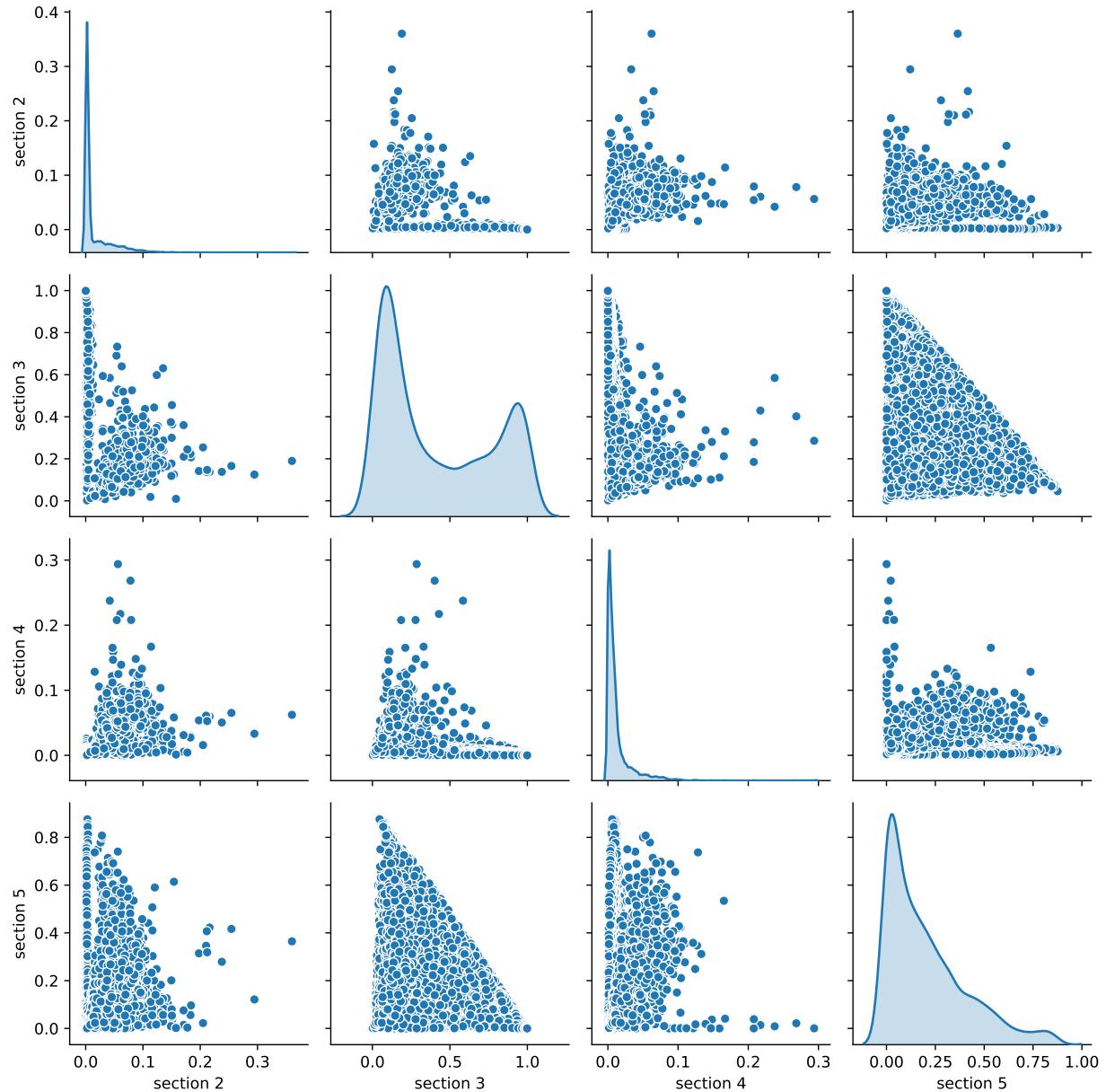


Figure 16: Experimental distribution of $(\bar{R}_2, \bar{R}_3, \bar{R}_4, \bar{R}_5)$ (neuron averaging; each sample is a single datum) for LeNet 5-piece network trained to high validation accuracy on MNIST, and evaluated on MNIST data.