A data-driven method for modeling dissipation rates in stratified turbulence: Supplementary materials

Samuel F. Lewin[*], Stephen M. de Bruyn Kops[†], Colm-cille P. Caulfield[‡], Gavin D. Portwood[§]

[*]DAMTP, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK

[†]Department of Mechanical and Industrial Engineering, University of Massachusetts Amherst, Amherst, Massachusetts 01003, USA

[‡]Institute for Energy and Environmental Flows, University of Cambridge, Madingley Road, Cambridge CB3 0EZ, UK

[§]Lawrence Livermore National Laboratory, Livermore, California 94550, USA

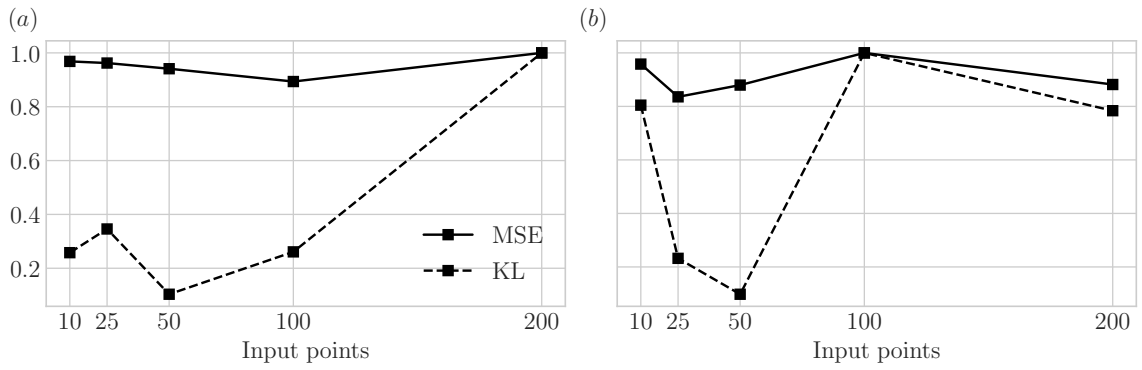

Figure 1: Values of MSE (solid lines) and KL-divergence (dashed lines) evaluated on five PCNN models with differing numbers of points m within the input columns for a) ε and b) χ. Each plot is normalised by its maximum value.

1 Sensitivity to input column height

As described in the manuscript, the local PCNN outputs for ε and χ depend non-locally on the inputs in a radius $m/2$ around the output point, where m is a user-specified parameter in the model. A value of $m = 50$ is chosen in the main manuscript based on the differential sensitivity analysis in §4. Here, we provide further evidence for this choice of m by evaluating the performance of five different models with different values of $m \in \{10, 25, 50, 100, 200\}$. The models are otherwise identical in their structure, and are trained and tested separately on the same dataset. For the purposes of an evaluation on all time steps from the DNS simultaneously, the test set here is slightly different to that used in the main manuscript in the sense that it is generated by randomly sampling approximately 10,000 columns from each time step of the simulation (which are importantly distinct from the training data columns). The metrics used for evaluation are the mean squared error (MSE) over the entire test set, and the Kullback-Leibler (KL) divergence (or relative entropy). The KL divergence is a measure of the difference between the predicted and true distributions of dissipation and is defined as

$$\int_{-\infty}^{\infty} P(y) \log\left(\frac{P(y)}{Q(y)}\right) dy, \tag{1}$$

where P and Q are the probability density functions (normalised frequency distribution) associated with the predicted and true data. As shown in figure 1, there is relatively little variation in the MSE between models, with more significant differences in the KL divergence. Ultimately, the choice of $m = 50$ corresponds to optimising the two error metrics simultaneously. As explained in the main manuscript, $m < 50$ will result in a loss of important non-local information about the background density gradient and therefore lead to worse predictions. We add here that larger values of m results in a network with many more trainable parameters than is necessary, potentially more likely to converge on a sub-optimal function for calculating dissipation. Thus $m = 50$ is an intermediate optimal balance.