1 Second moment for SC[NuN]

Following Eq. (15) in the main text, the second moment for the two-layer DGP is given by,
\[
\frac{\sigma^2}{2} \left\{ 1 + \exp\left[2\left(e^{-(\alpha \vert \mathbf{x}_i \vert^2 + 2\beta \mathbf{x}_i \cdot \mathbf{x}_j - \alpha \vert \mathbf{x}_j \vert^2)} - e^{-2(\alpha - \beta)\vert \mathbf{x}_i \vert^2} - e^{-2(\alpha - \beta)\vert \mathbf{x}_j \vert^2}\right]\right\}
\]
where we have suppressed the factor of $\sigma_0^2/2\ell^2$ inside the bracket for convenience.

2 Experiments on two UCI datasets

Figure 1: Simulated regression using UCI data sets. House Price data has a total of 320 training and 94 testing data points. Abalone data has a total of 400 training and 400 testing data points.
3 Results using Doubly Stochastic Variational Inference for Deep Gaussian Process

Figure 2: Fit result using the method proposed by Salimbeni and Deisenroth (2017) (https://github.com/ICL-SML/Doubly-Stochastic-DGP). Panels (a), (b), and (c) correspond to Figure 1 (e), (f), and (g), respectively. Because of the existence of outliers, we used boxplots instead of mean ± se for the mean-square error in panel (b).

References