1 Preliminary Results

$\tilde{\rho}_{\mathrm{mse}}$

n

$\mathrm{var}(\tilde{\rho}_{\mathrm{pred}})$

$\Delta\tilde{\rho}$

(a) Number of discarded predictions n vs. mean squared error $\tilde{\rho}_{\mathrm{mse}} = \frac{1}{N}\sum_{i=1}^{N}(\tilde{\rho}_{\mathrm{pred},i} - \tilde{\rho}_{\mathrm{exp},i})^2$

(b) Absolute error $\Delta\tilde{\rho} = |\tilde{\rho}_{\mathrm{pred}} - \tilde{\rho}_{\mathrm{exp}}|$ vs. model uncertainty $\mathrm{var}(\tilde{\rho}_{\mathrm{pred}})$

$\tilde{\rho}_{\mathrm{pred}}$

$\tilde{\rho}_{\mathrm{exp}}$

$\mathcal{N}(\tilde{\rho}_{\mathrm{pred}})$

$\mathcal{N}(\tilde{\rho}_{\mathrm{exp}})$

(c) Experimental vs. predicted $\tilde{\rho}$

(d) Exp. vs. pred. $\tilde{\rho}$ normalized to $\mathcal{N}(0,1)$
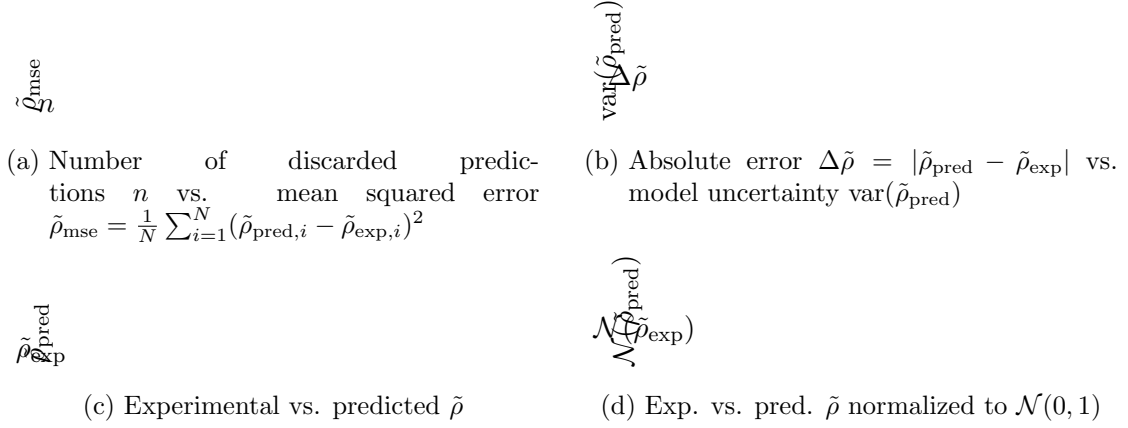
Figure 1: Preliminary results for predicting normalized logarithmic resistivity $\tilde{\rho} = \log[\rho/(\Omega\,\mathrm{cm})]$

?? shows preliminary results for a neural network that combines aleatoric (observational) and epistemic (model) uncertainty estimation. It predicts the logarithm of the electrical resistivity $\log(\rho)$, which together with Seebeck coefficient S and thermal conductivity κ is one of three material properties necessary to compute the thermoelectric figure of merit (T is the absolute temperature),

$$zT = \frac{S^2 T}{\rho\,\kappa}. \tag{1}$$

For brevity, we show results only for ρ but our neural nets perform comparably on S and better on κ.

In **??**, the mean squared error (MSE) drops off monotonically as we successively discard the model's most uncertain predictions (blue line), demonstrating that the estimated model uncertainty is strongly correlated with the error it made. The orange line was plotted by successively removing points of largest model error. It marks the ideal that would be reached if the model knew its own error perfectly.

?? provides evidence that our model is unlikely to predict false positives. Most points lie on or above the parity line ($y = x$). There are no points with large absolute error that the model feels certain about. Points in the top-left corner have large uncertainty despite incurring only small errors. If those materials are in fact good thermoelectrics, they would be false negatives.

???? show model predictions vs. experimentally measured ground truth. The error bars indicate model uncertainty, demonstrating that the model is less certain about outliers and more so in regions of material space with ample training data. **??** is normalised, i.e. with labels rescaled to zero-mean and unit-variance, **??** non-normalised. We see good agreement between prediction and experiment. Moreover, the uncertainty

of predictions indicated by the error bars is significantly larger for outliers than for clustered data indicating that the model correctly estimates high epistemic uncertainty for those predictions.