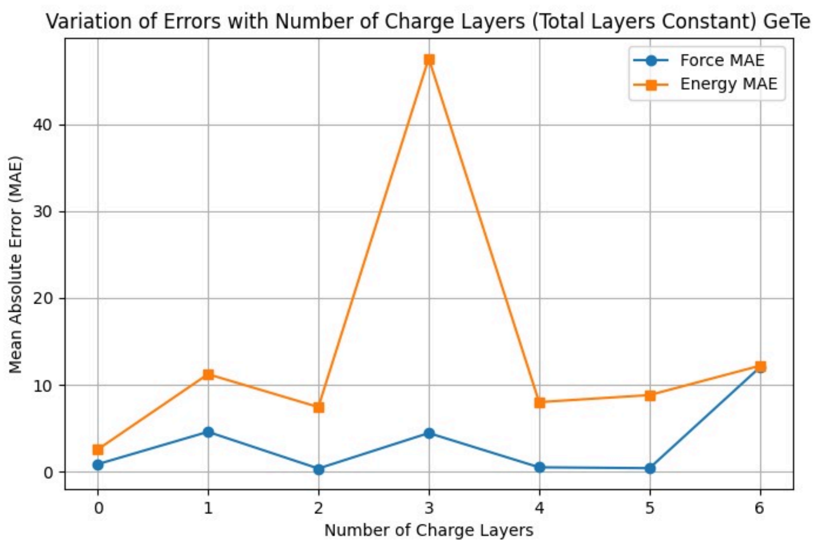
Comparative Analysis of NequIP and NequIP-LR - IV

GeTe

A) Effect of increasing the number of charge layers-total_layers6



- **Adding charge-encoding increases the errors**

- **GeTe has strong electron-lattice interactions, which means modeling charge effects is important . The fact that n_charge = 0 (no charge layers) still gives relatively low errors suggests that Nequip's base architecture captures some charge effects even without explicit charge layers**

Potential Reasons behind increasing errors in nequiplr-chagre_encoding

1) Since nequip-charge learns atomic charge distributions, it may struggle more with materials where charge is not well - localized.
2) GeTe is a narrow-bandgap semiconductor (~0.1-0.2 eV), meaning its electrons are more delocalized.
3) Increasing errors with added charge layers suggest that the long range model is overfitting on the movements of electrons, i.e, the model may learn non-physical charge distributions..
4) This is verified by the above plot.

Other examples:

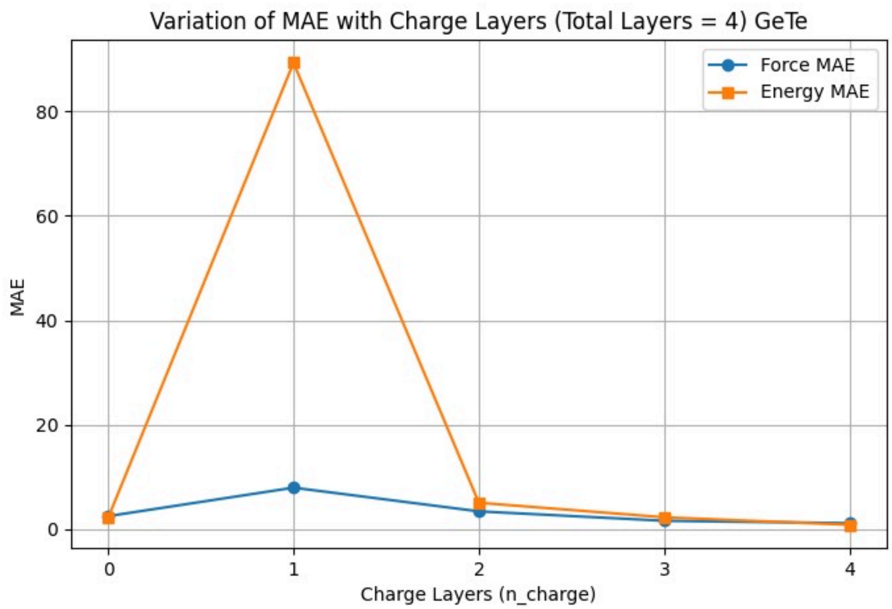
1) (MAPI) is a hybrid perovskite, which generally has more localized ionic and electronic charge distributions.
2) i) 3bpa is an organic molecule with localized electron density around covalent bonds.

   ii) Charges are well-defined on atoms (e.g., oxygen has a partial negative charge, hydrogen a partial positive charge).

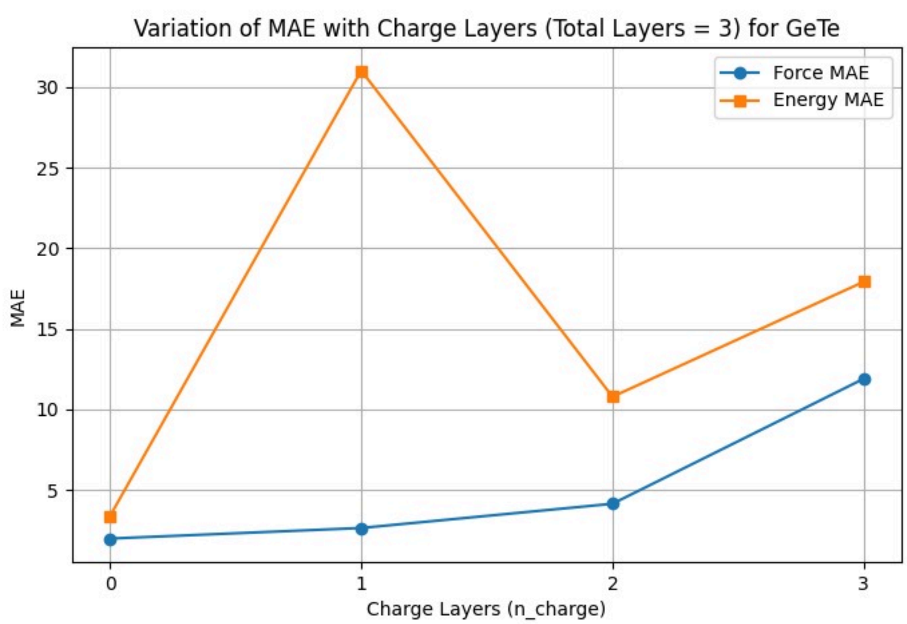
Nequip-charge can easily learn stable charge distributions because they are consistent across the dataset. Verifiable by the plots in the previous report

How to improve:
We can try adding some constraints on GeTe long range interactions (like a fixed number of charge categories to a predefined range) to improve performance.

B) Effect of increasing the number of charge layers-total_layers4



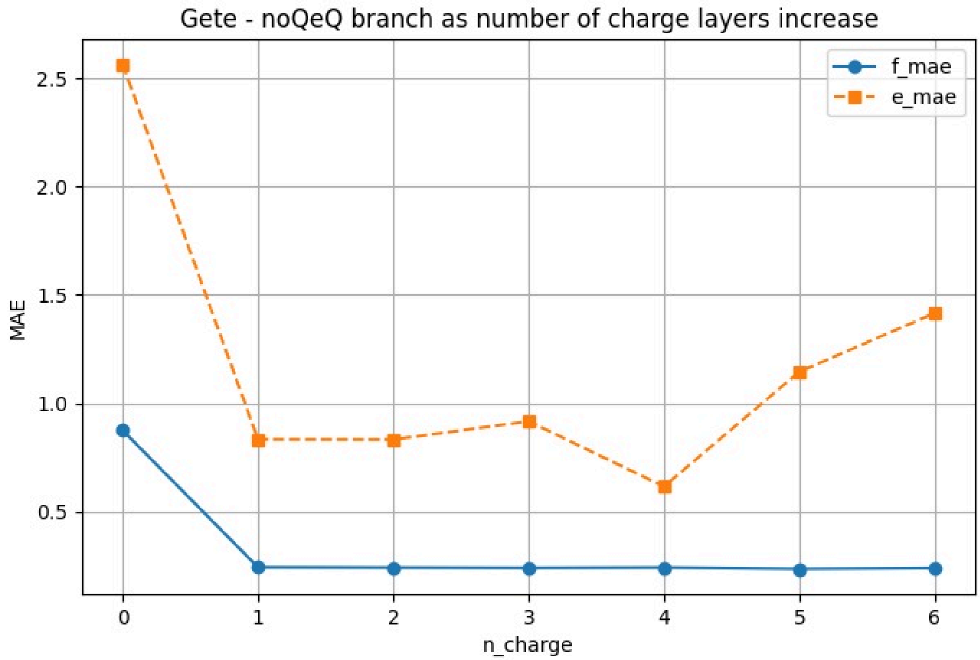
C) Effect of increasing the number of charge layers-total layers4

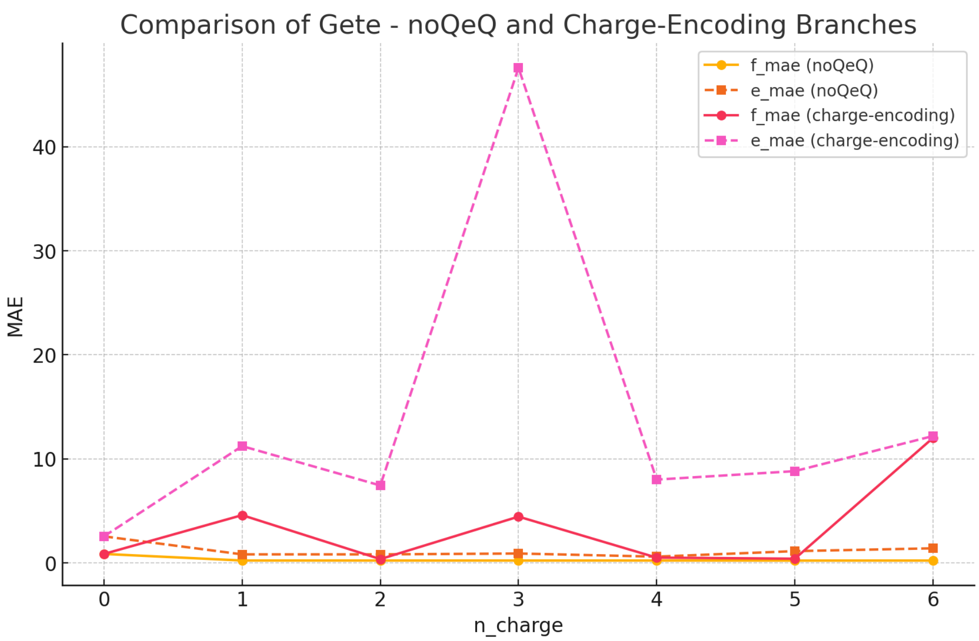


1) Again we do not observe any improvement with addition of charge layers due to similar reasons.
2) We observe that a spike is seen as soon as a single charge layer is introduced but it becomes stable afterwards.
3) N_charge_layer = 1 fails to differentiate between charge interactions and delocalisation effects while increasing the number of layers gives an intelligent prediction.

We theorise that a long range machine learning based system in which charge equilibration is not used and charges are not determined dynamically should perform better.

No-QeQ branch results for GeTe





We see a considerable improvement in No-QeQ over charge-encoding-nequiplr and Nequip. These plots align with our hypothesis.

RMAX Variation curves don't give any consistent information as already their errors surpass the base Nequip model