Appendix

A.1

Why can we not including damping into the Hamiltonian? Let us take the following damped system:

$$\ddot{\mathbf{q}} = -\mathbf{q} - \delta\dot{\mathbf{q}} \tag{1}$$

where we know $\mathbf{p} = m\dot{\mathbf{q}}$ which implies $\dot{\mathbf{q}} = m^{-1}\mathbf{p}$.

Then, the integral of the right hand side with respect to q will give us:

$$\frac{\mathbf{q}^2}{2} + \delta\mathbf{q}\dot{\mathbf{q}} \tag{2}$$

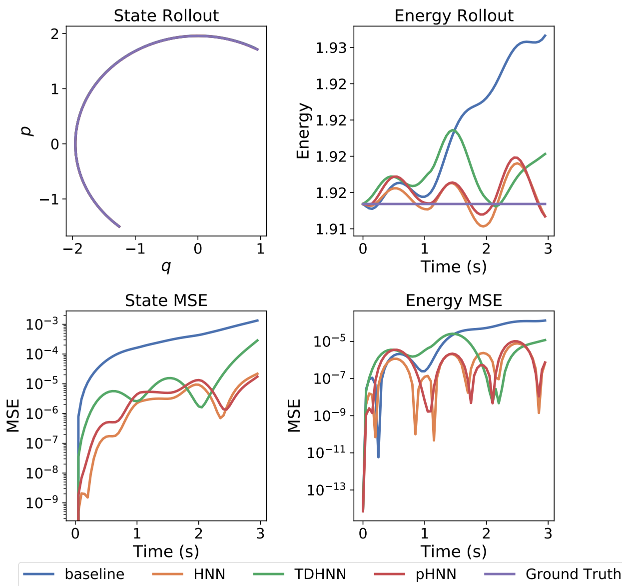
The equation above looks like a modified potential function which can be combined with a kinetic energy term to give a Hamiltonian s.t.:

$$\mathcal{H} = \frac{\mathbf{p}^2}{2m} + \frac{\mathbf{q}^2}{2} + \delta\mathbf{q}\dot{\mathbf{q}} \tag{3}$$

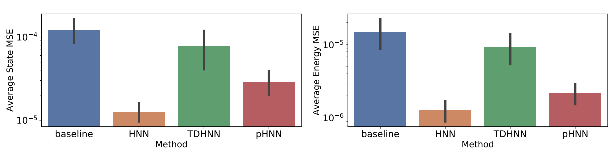
However, although we can recover the differential equation for $\ddot{\mathbf{q}}$ by $-\frac{\partial\mathcal{H}}{\mathrm{d}\mathbf{q}} = -\mathbf{q} - \delta\dot{\mathbf{q}} = \ddot{\mathbf{q}}$, we violate the rule that $\dot{\mathbf{q}} = m^{-1}\mathbf{p}$ since $\frac{\partial\mathcal{H}}{\mathrm{d}\mathbf{p}} = \dot{\mathbf{q}} + \delta\mathbf{q} \neq \dot{\mathbf{q}}$.

Indeed an alternative formulation that uses exponentiated time can allow us to represent a Hamiltonian for the damped system but this is not within the scope of our study as the formalism does not explain how to include external forcing.

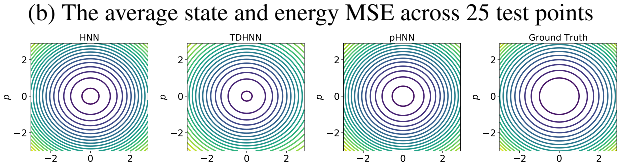
B



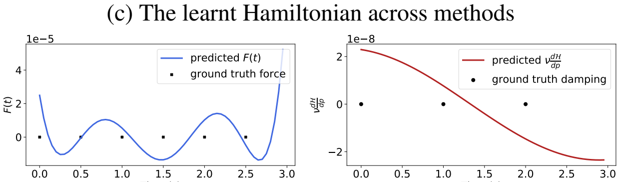
(a) State and energy rollout of an initial condition from the test set



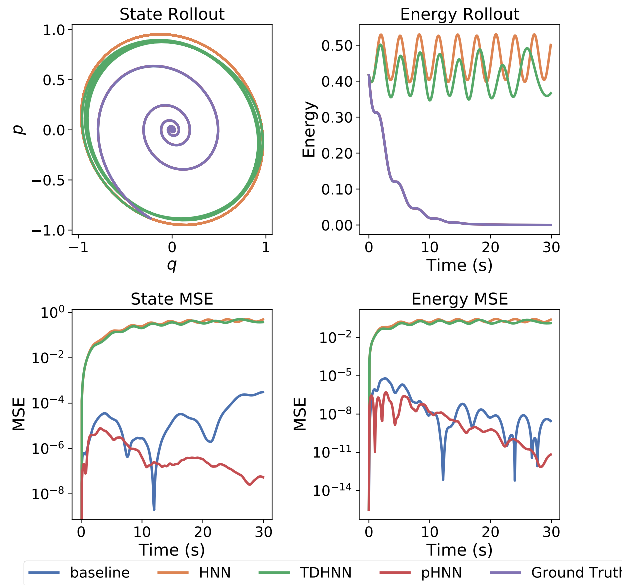
(b) The average state and energy MSE across 25 test points



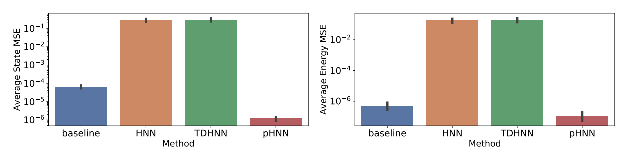
(c) The learnt Hamiltonian across methods



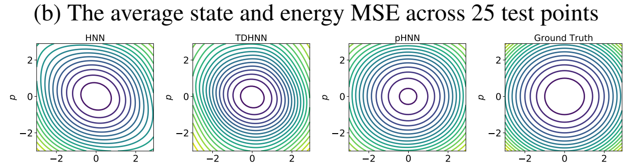
(d) The learnt force and damping of TDHNN4



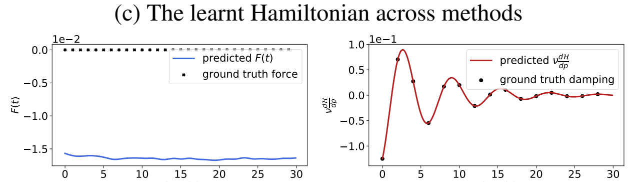
(a) State and energy rollout of an initial condition from the test set



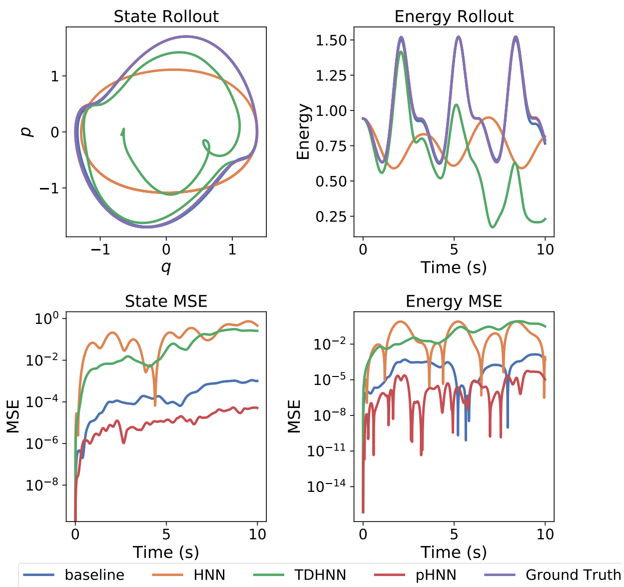
(b) The average state and energy MSE across 25 test points



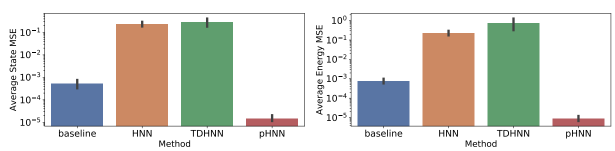
(c) The learnt Hamiltonian across methods



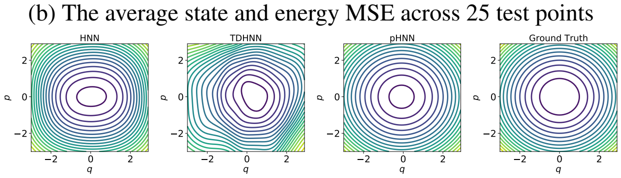
(d) The learnt force and damping of TDHNN4



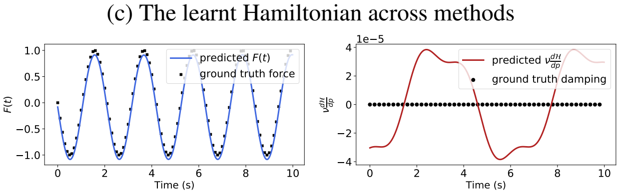
(a) State and energy rollout of an initial condition from the test set



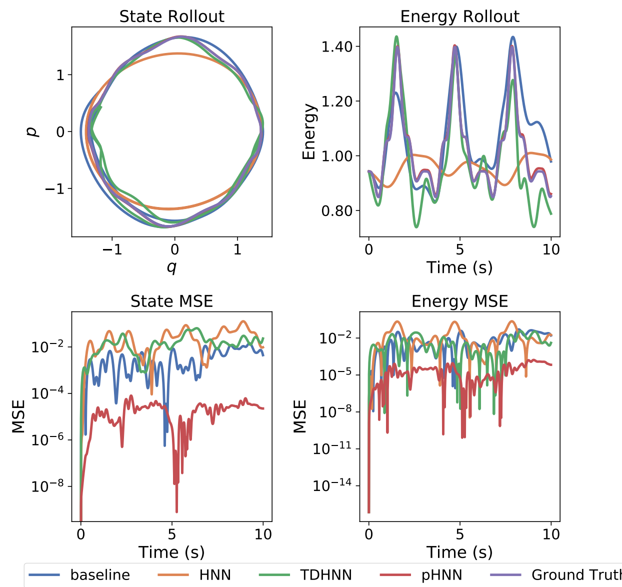
(b) The average state and energy MSE across 25 test points



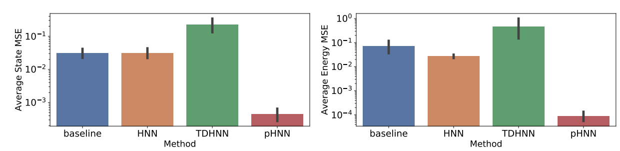
(c) The learnt Hamiltonian across methods



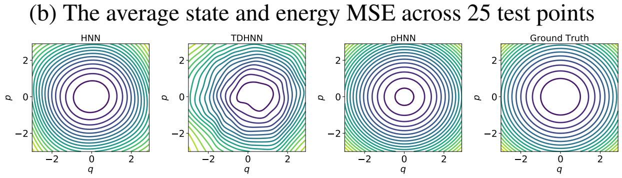
(d) The learnt force and damping of TDHNN4



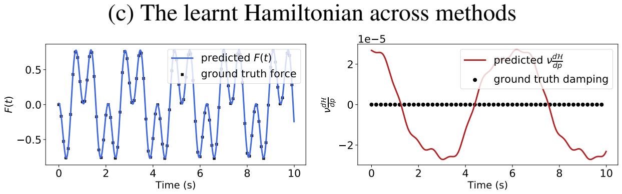
(a) State and energy rollout of an initial condition from the test set



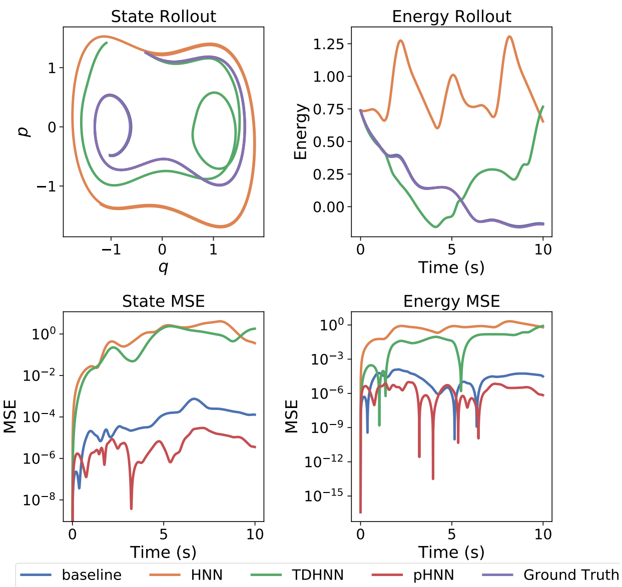
(b) The average state and energy MSE across 25 test points



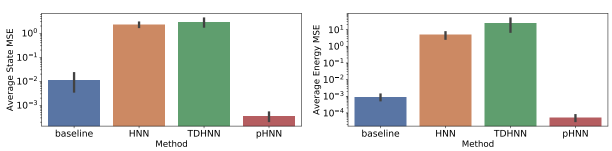
(c) The learnt Hamiltonian across methods



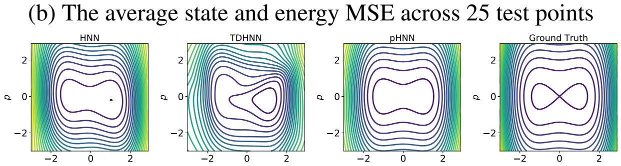
(d) The learnt force and damping of TDHNN4



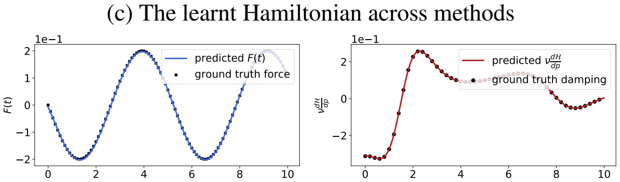
(a) State and energy rollout of an initial condition from the test set



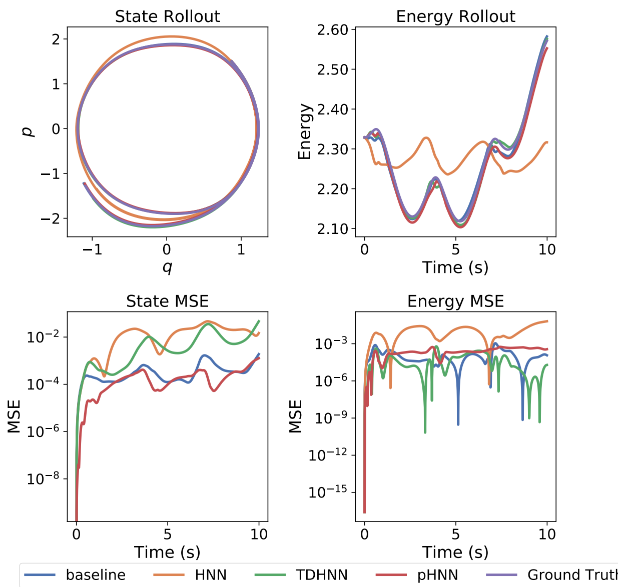
(b) The average state and energy MSE across 25 test points



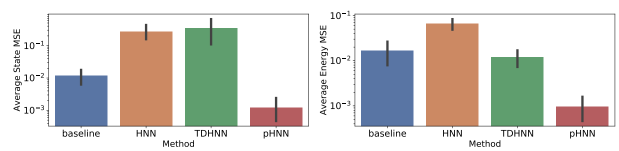
(c) The learnt Hamiltonian across methods



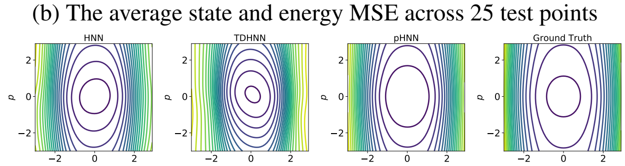
(d) The learnt force and damping of TDHNN4



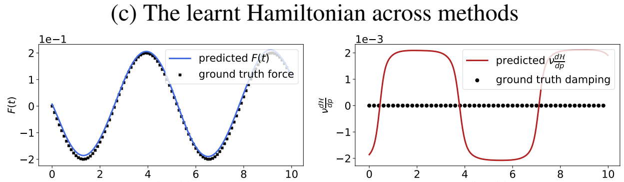
(a) State and energy rollout of an initial condition from the test set



(b) The average state and energy MSE across 25 test points



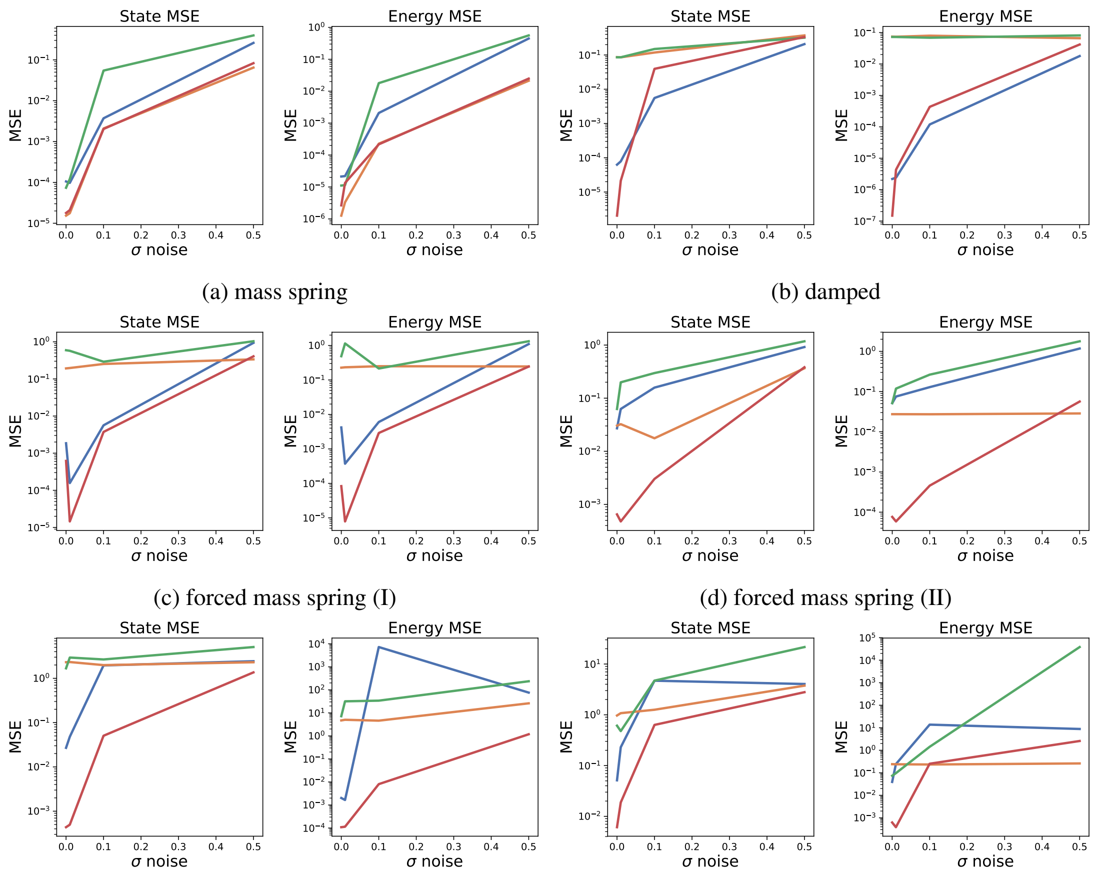
(c) The learnt Hamiltonian across methods



(d) The learnt force and damping of TDHNN4

C

During the training of HNN, the authors add gaussian noise with a standard deviation $\sigma = 0.1$ to the input state vector data. The reason this is done is to ensure the model is robustly trained. We run a set of experiments to test the robustness to this 'noisy' input.



(a) mass spring (b) damped

(c) forced mass spring (I) (d) forced mass spring (II)

(e) duffing (f) relativity

D

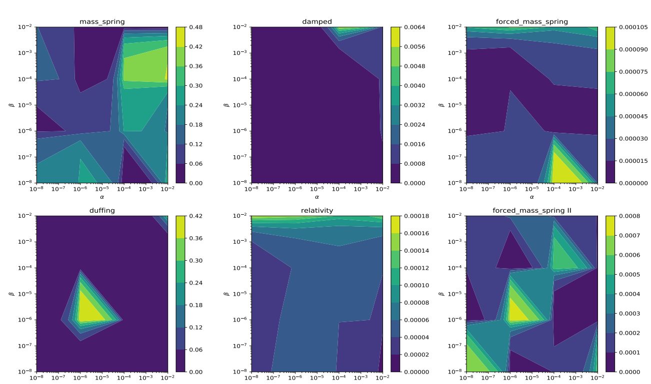


Figure 8. Hyperparameter Optimization for TDHNN4. We plot, for each system, the validation loss as a function of the α and β parameters from the loss in eqn. 5