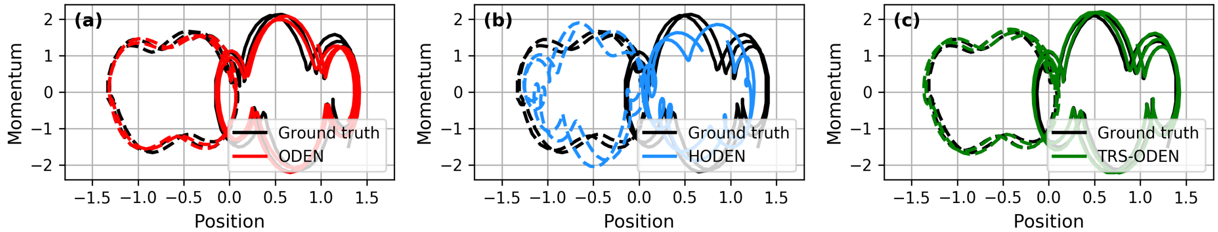

Figure A: Test trajectory of oscillators (solid: mass 1 / dashed: mass 2) predicted by (a) ODEN, (b) HODEN, and (c) TRS-ODEN.

We thank all reviewers for their valuable comments. In particular, we appreciate that reviewers agree on the importance of the time-reversal symmetry (TRS) [R1, R2, R4] and clear writing [R1, R2, R3, R4]. Followings are our responses.

Q1. [R1, R3, R4] Empirical evaluation on real dataset. During the rebuttal period, we performed an experiment with real-world data [35]. This data consists of a trajectory of real double oscillators, which are neither conservative nor reversible due to damping and other non-ideal effects. We set the first 3/5 of the trajectory for training, and the remains for test. We used same hyper-parameters in Experiment IV. **Table A** and **Figure A** show our model outperforms baselines, especially HODEN. It reveals 1) while enforcing the conservation may not be good for real world, 2) guiding symmetry with TRS loss is helpful. We will add this result in the final draft.

Table A: Experimental results (repeated 5 times).

Model	Mass 1 MSE	Mass 2 MSE
ODEN	1.00 ± 0.22	0.37 ± 0.05
HODEN	38.13 ± 2.16	32.60 ± 1.96
TRS-ODEN	$\mathbf{0.36 \pm 0.06}$	$\mathbf{0.15 \pm 0.01}$

Q2. [R1] Balance between two losses. Because the total loss is given by $\mathcal{L}_{\text{ODE}} + \lambda \cdot \mathcal{L}_{\text{TRS}}$, higher (lower) λ leads stricter (looser) symmetry. While λ is treated as a constant generally, one can deal with it as a function of $(\mathbf{q}, \mathbf{p})$ and t, with the assumption that the irreversible term in ODE is also a function of them. We premise it gives more precise balance between two losses, especially when addressing irreversible systems. We will clarify this part in the final draft.

Q3. [R2] Guarantees on TRS. We agree with the comment that regularizers do not guarantee the perfect TRS solution. Nevertheless, one can force the solution be almost symmetric by increasing λ. To confirm this, we evaluated the relative error between forward and backward trajectories of TRS-ODENs that trained with varying λ (**Figure B**). It shows large $\lambda = 10^3$ guarantees lower than 10^{-3} relative error, without performance degradation. We believe the flexibility to control the degree of TRS is rather a strong merit of TRS-ODEN, as mentioned in line 56-61 and 129-131. We will add the respective discussion to the final draft.

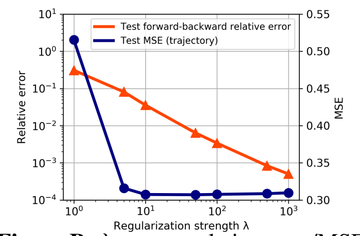

Figure B: λ *vs.* test relative error/MSE.

Q4. [R2] The reversing operator. We used the reversing operator $R : R(\mathbf{q}, \mathbf{p}) = (\mathbf{q}, -\mathbf{p})$, which we will state explicitly in the final draft. It is quiet general for classical mechanics, whose $\mathbf{p}$ is naturally negated by the time-reversal, as described in line 141-145. However, there are some particular systems that such R does not work well, e.g., chaotic strange attractors (see **Q8** for more information). While one can set proper R for this case (as we did in **Q8**), automatic search of R from unknown data is an interesting future work. We will add a related discussion in the final draft.

Q5. [R2] Impact on training time. TRS-ODENs require approximately two times larger training time than default ODENs because the backward as well as forward evolutions need to be calculated. We will state it in the final draft.

Q6. [R2] Previous works. The regularizers mentioned in line 27-29 use very specific models such as Navier-Stokes equation, thus can be applied only when governing laws are exactly known. We will clarify this part in the final draft.

Q7. [R3] Motivation and benefit of TRS loss. As described in line 39-47 and pointed by R1, R2, and R4, TRS is an important symmetry in physics. It motivates us to design the TRS regularizer. As you pointed out, it is important to inform readers what kinds of systems that TRS works well. To do this, we stated the target systems and expected benefits of TRS for them in line 111-131. Clearly, TRS is powerful to model reversible systems, which are natural in classical dynamics. Thus, they are the primary target of our proposed method. In addition, we would like to emphasize that even for irreversible systems TRS is helpful for model generalization, as shown in Experiment IV and real-world experiment (**Q1**, **Table A** and **Figure A**). We analyze the reason as twofold: 1) the irreversible terms in ODE can be (partially) negligible during dynamics, and 2) TRS regularizer flexibly guides the symmetry rather than enforcing it.

Q8. [R3] Strange attractors. Some strange attractors show TRS under non-trivial reversing operators R, according to "Sprott, J. C. *International Journal of Bifurcation and Chaos*, 25(05):1550078, 2015". During the rebuttal period, we conducted an experiment with $\dot{x} = 1 + yz, \dot{y} = -xz, \dot{z} = y^2 + 2yz$, a reversible strange attractor under $R : R(x, y, z) = (-x, -y, -z)$. Since it is not straightforward to set Hamilton's equation for this system, HODEN is not evaluated. We generated 1,000 and 50 trajectories for training and test, respectively. As a result, we found TRS-ODEN can achieve smaller test MSE than ODENs, e.g., from 17.0 to 13.2 (see **Figure C**). We will add this result to the final draft.

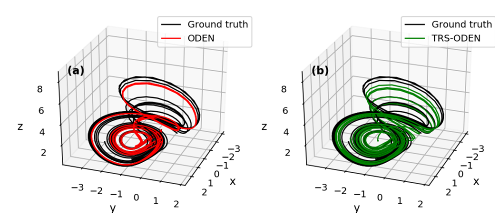

Figure C: Sampled test trajectories of the strange attractors predicted by (a) ODEN and (b) TRS-ODEN.

Q9. [R2, R4] Suggestions on clarity. We will revise typos and grammatical errors according to reviewers' suggestions.