Formal Verification of Neural ODE for Safety Evaluation in Autonomous Vehicles

Abdelrahman Sayed Sayed

Univ Gustave Eiffel, COSYS-ESTAS, F-59657 Villeneuve d'Ascq, France

Email: `abdelrahman.ibrahim@univ-eiffel.fr`

Université Gustave Eiffel — LABORATOIRE ESTAS ÉVALUATION DES SYSTÈMES DE TRANSPORTS AUTOMATISÉS ET DE LEUR SÉCURITÉ

Co-funded by the European Union

Motivation & Societal Impact

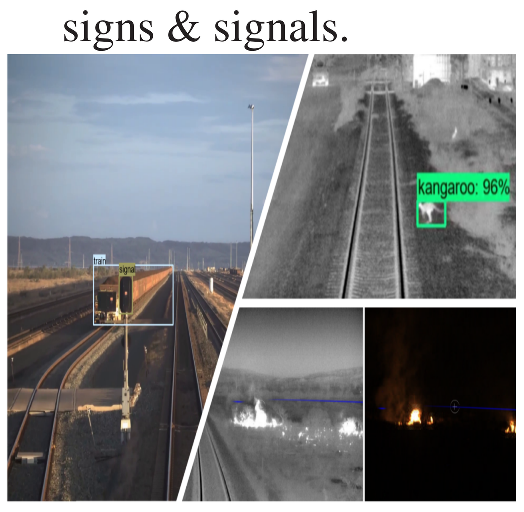
- **Higher Autonomy** is a primary goal for transportation systems of future cities.
- **Problem:** Most autonomy systems rely on **AI models** for critical **perception tasks**.
- **AI in Transportation Trade-off:**

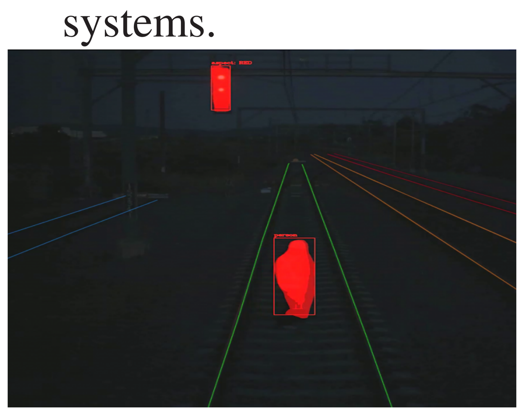
Promises of AI Perception

- ✓ High performance in complex urban settings.
- ✓ Superior recognition of signs & signals.

Risks of Deployment

- ⚠ Unpredictability in edge-cases.
- ⚠ lack of of formal verification.
- ⚠ Major risks for deployment on safety-critical systems.

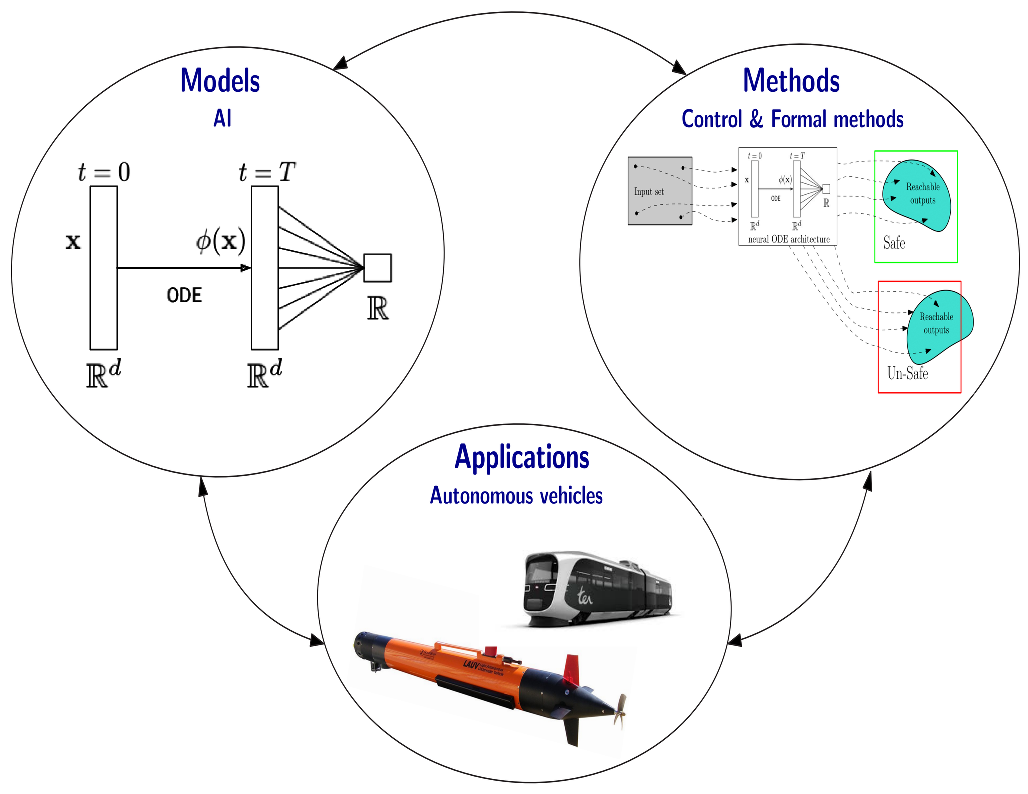




My research is motivated by the critical need to bridge the gap between AI-based perception modules and their safety verification when deployed in autonomous vehicles.

- **Neural ODE** *(Chen et al. 2018)* are commonly described as a **continuous-depth** generalization *(Sayed et al. 2025a)* of a discrete **ResNet** *(He et al. 2016)*.
- **Neural ODE** have gained prominence in **time-series** modeling over discrete neural networks *(kidger, 2021)*.
- **Neural ODE Verification Gap:** Current methods are **under-developed** *(Lopez et al. 2022)*, and existing methods are **computationally intensive** *(Sayed et al. 2025b)*.

> – We provide novel methods and light weight tools based on reachability analysis to **analyze and verify safety properties of neural ODE.**
> – We combine **continuous AI/neural ODE** models, **reachability analysis** methods to analyze such models, with the final objective of **verifying safety of AI mdoels utilized by autonomous vehicles.**



Future Research Directions

- **Scalability to other architectures** Extending the framework in **R.D 1** to handle more neural network architectures (e.g., **RNN and CNN**).
- **Uncertainty quantification** Moving from deterministic ODEs to **Stochastic Differential Equations (SDEs)** to model sensor noise and environmental unpredictability.
- **Closed-loop verification** Verifying the safety of the entire autonomous loop by integrating the neural ODE perception module with the vehicle's control policy.
- **Real-time verification** Optimizing verification algorithms to run efficiently on embedded hardware for **adaptive mission planning**.

References

[1] Sayed, A.S., Meyer, P.J., Ghazel, M.: Bridging neural ode and resnet: A formal error bound for safety verification. In: International Symposium on AI Verification. pp. 97–114. Springer (2025)

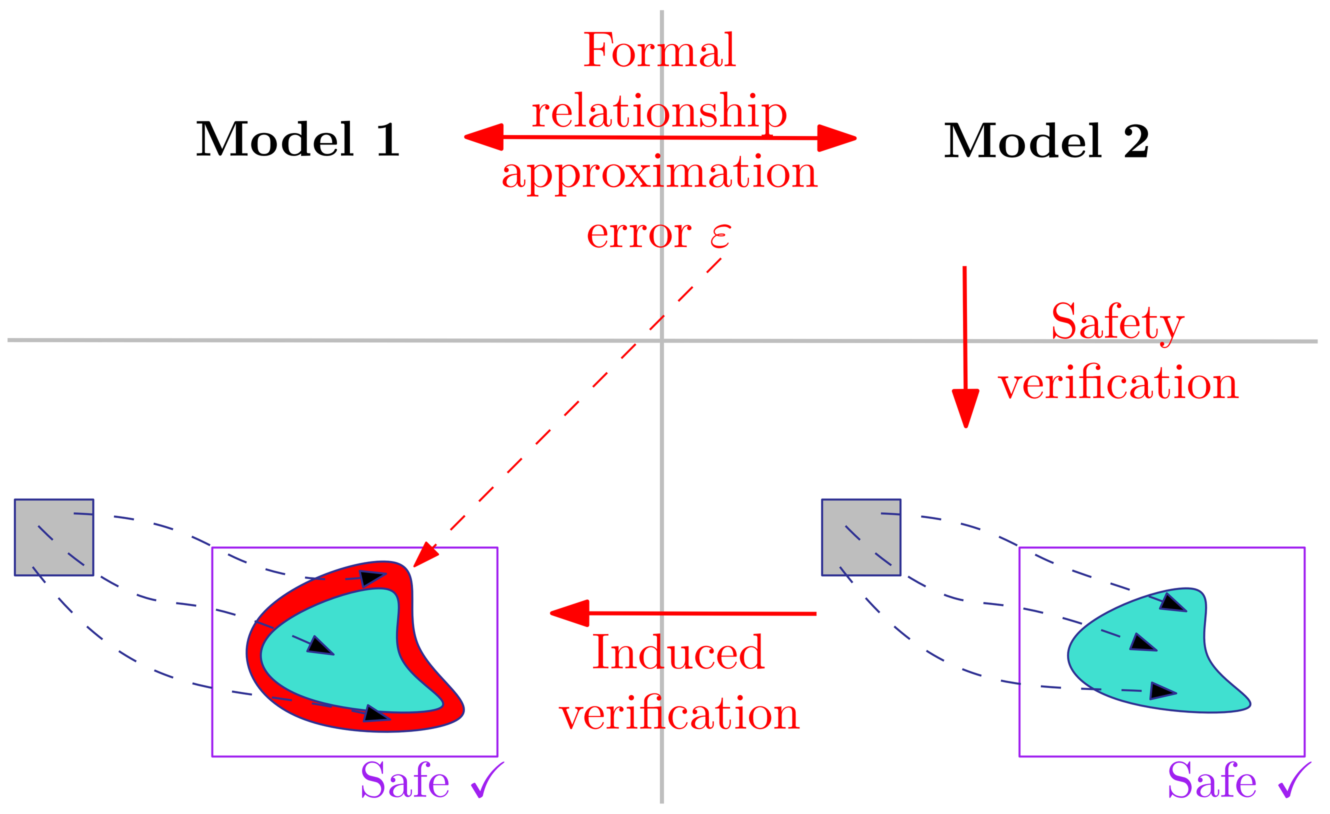
[2] Sayed, A.S., Meyer, P.J., Ghazel, M.: Mixed monotonicity reachability analysis of neural ode: A trade-off between tightness and efficiency. In: NeurIPS 2025 Workshop on Symmetry and Geometry in Neural Representations (2025)

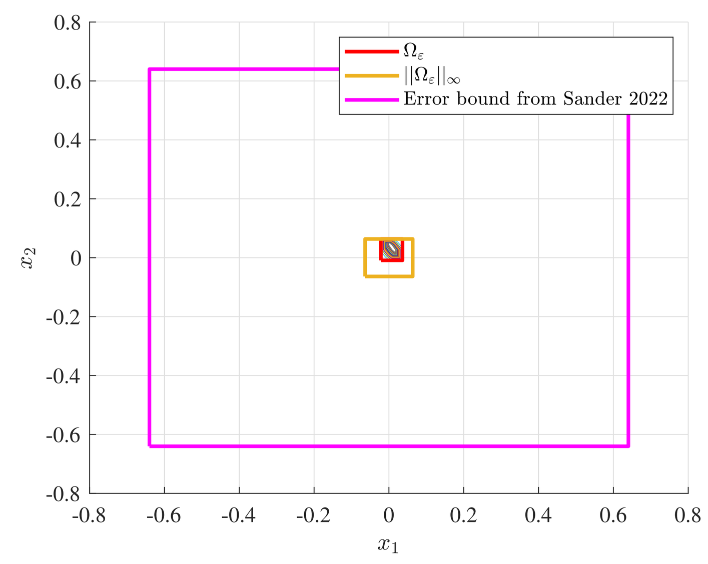
Thesis Contributions

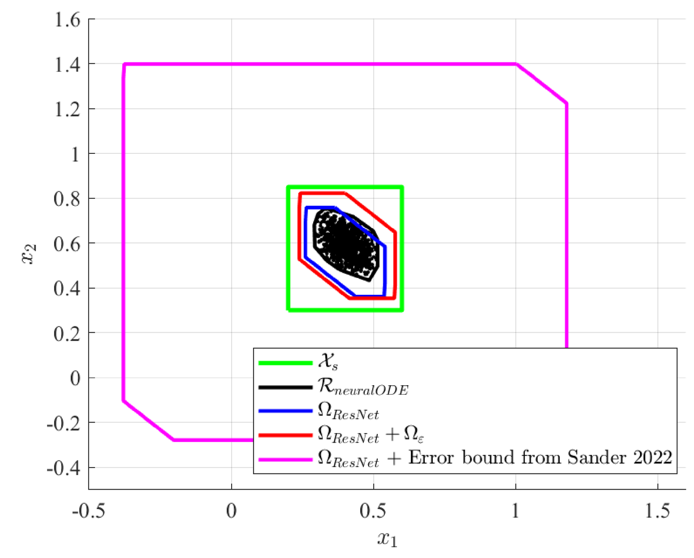
1. R.D 1: Formal relationships between neural networks and neural ODE

(Sayed et al. 2025a)

- Rigorous error bound (ε) on the approximation error between neural ODE and ResNet.
- Tighter Ω_ε over-approximation $<$ **16 million times** $\times$ SOTA *(Sander et al. 2022)*.
- Verification proxy to verify one model based on the reachable set of the other $\pm\,\Omega_\varepsilon$.



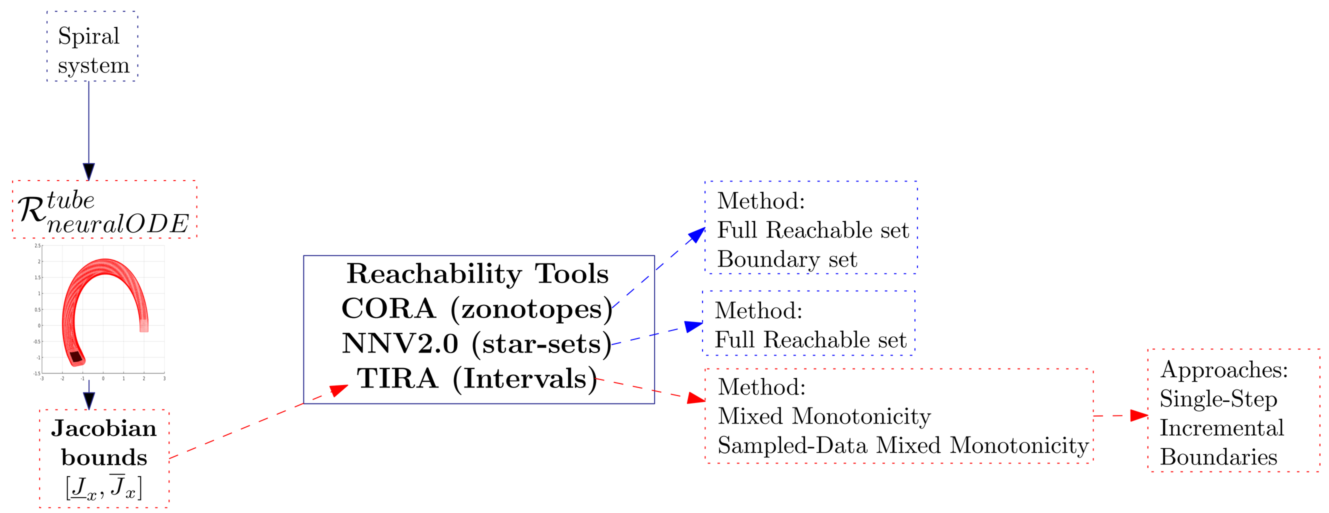


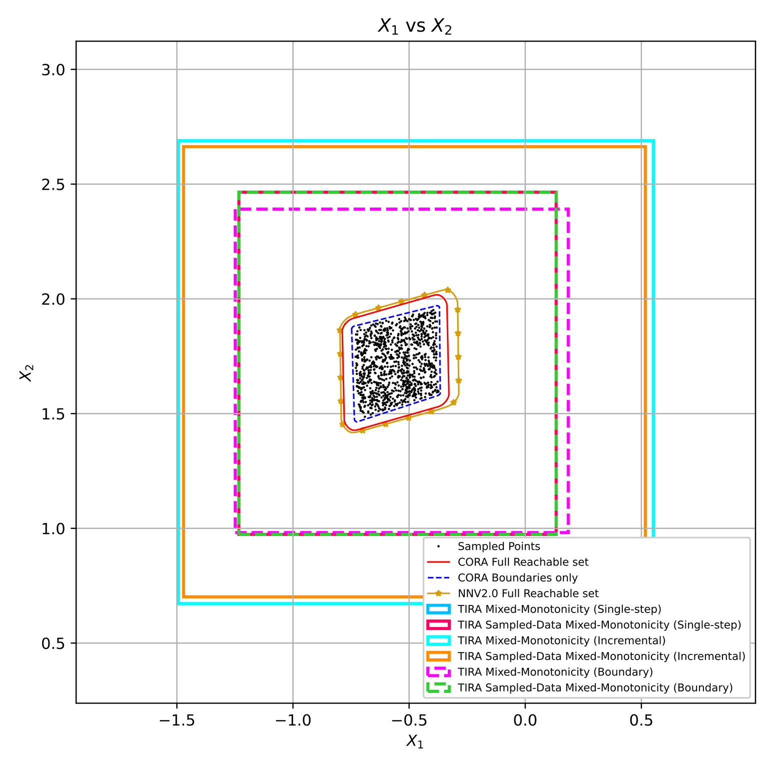


2. R.D 2: Safety verification of neural ODE based on reachability analysis methods

(Sayed et al. 2025b)

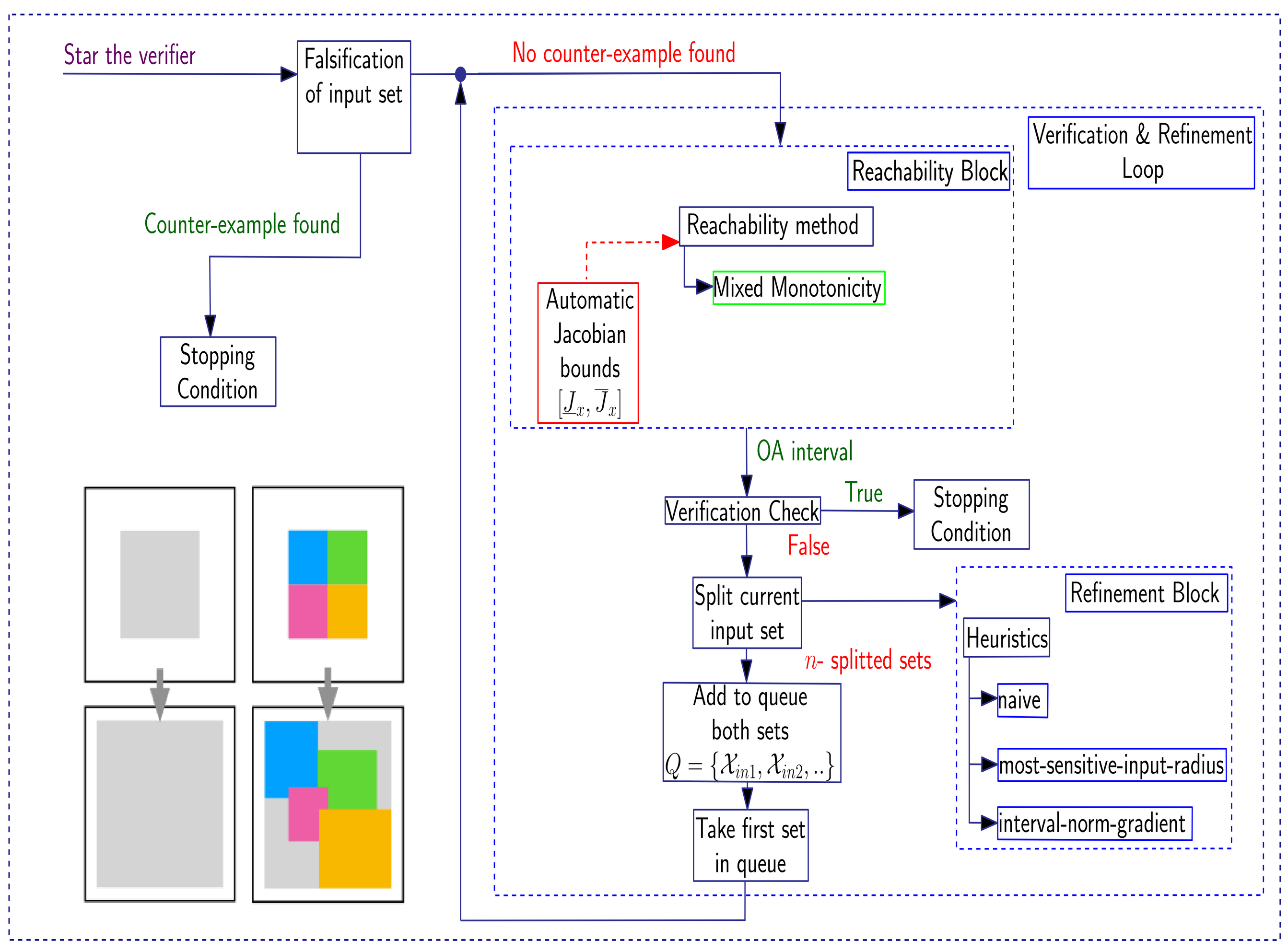
- Reachability analysis of neural ODE based on continuous-time mixed monotonicity.
- Light weight reachability analysis methods for scalable verification.

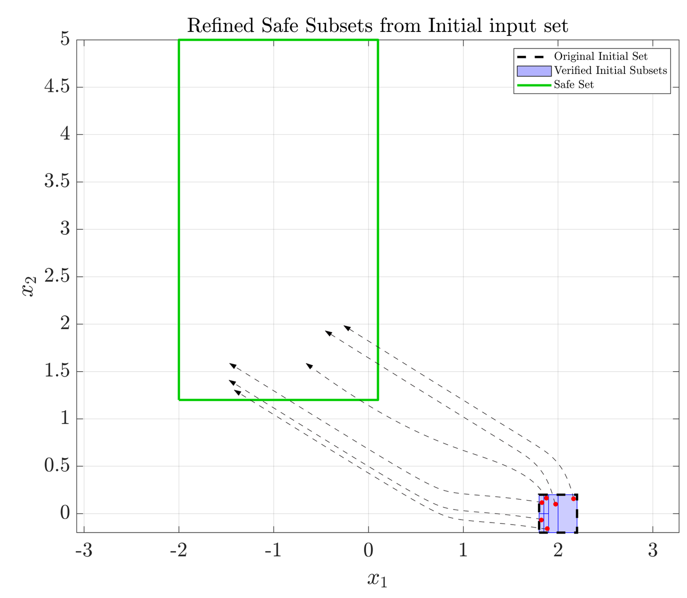


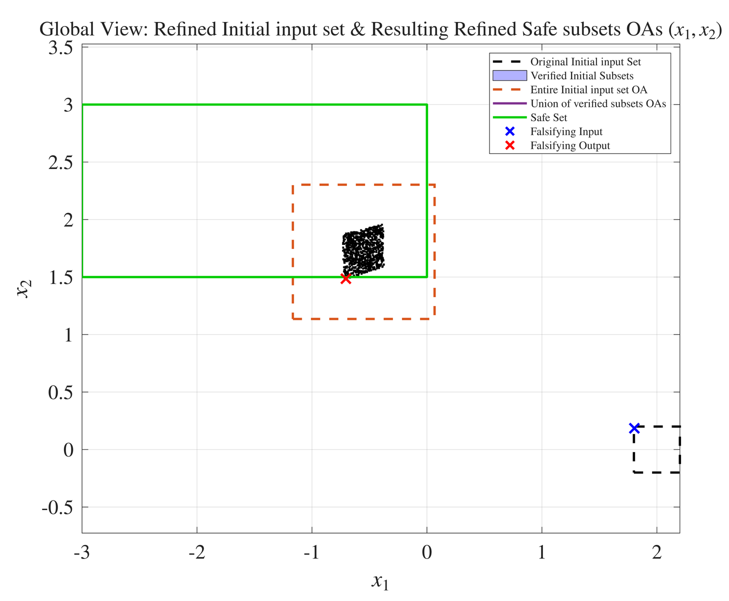


3. R.D 3: Verification toolbox for neural ODE

- Iterative refinement approach for neural ODE input set.
- Benchmarking with other neural ODE verification tools.







4. Verification and testing on autonomous underwater vehicles (AUVs)

- Robustness verification for ROI from in situ images.