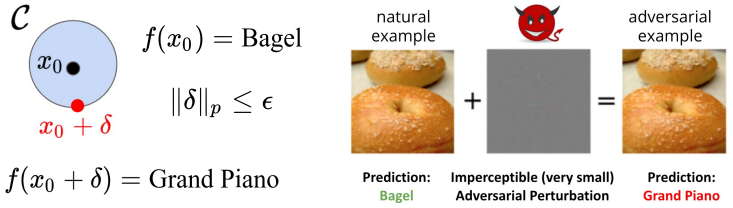
Fast and Complete: Enabling Complete Neural Network Verification with Rapid and Massively Parallel Incomplete Verifiers

Kaidi Xu* **Huan Zhang*** **Shiqi Wang** **Yihan Wang** **Xue Lin** **Suman Jana** **Cho-Jui Hsieh**

*Equal contribution

Paper: arxiv.org/abs/2011.13824
Code: PaperCode.cc/FastAndComplete

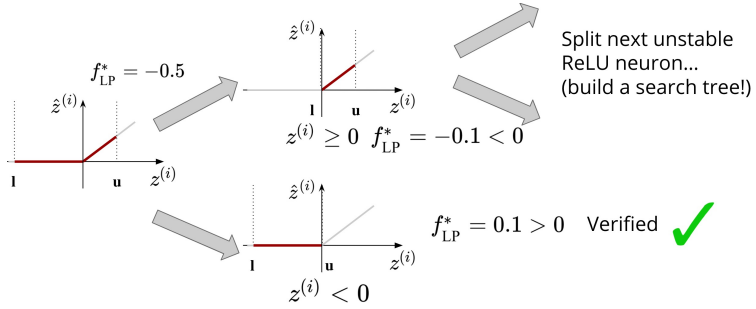
Adversarial Examples and Robustness



Property to Prove: $f(x) > 0, \forall x \in \mathcal{C}$

Provable Robustness Guarantee with BaB

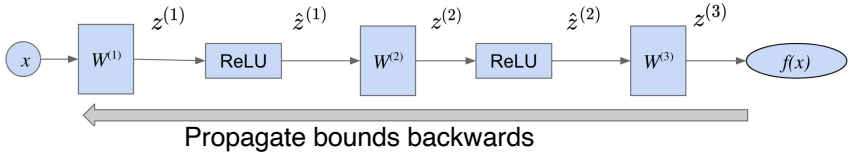
Branch & Bound (BaB) for ReLU Neural networks



Split next unstable ReLU neuron... (build a search tree!)

BaB requires an efficient solver for each subproblem.
Typically, linear programming (LP) verifiers are used

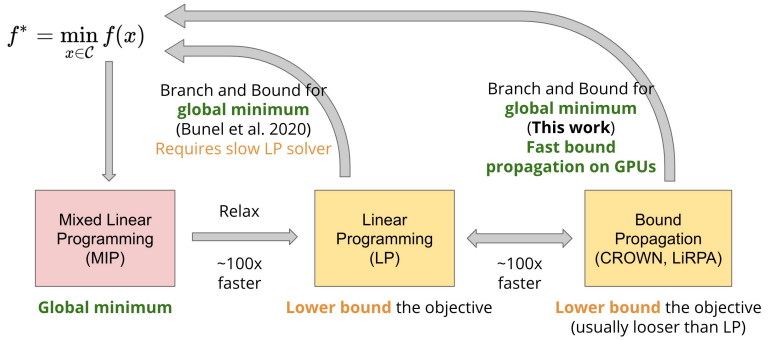
Bound Propagation Based Verifiers



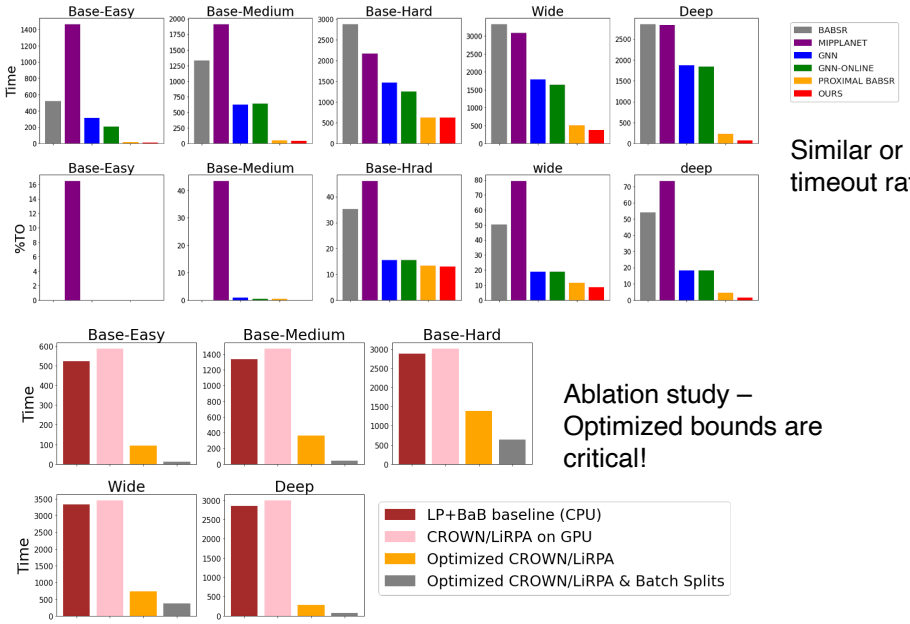
CROWN Bound: $\min_{x \in \mathcal{C}} f(x) \geq \min_{x \in \mathcal{C}} \boldsymbol{a}^\top_{\text{CROWN}} x + c_{\text{CROWN}} := \min_{x \in \mathcal{C}} f_{\text{CROWN}}(x; \boldsymbol{\alpha})$

Very fast, but tend to produce looser bounds than LP

Our Main Contribution: Combining **rapid** bound propagation based incomplete verifiers such as CROWN/LiRPA with branch and bound to **scale up neural network verification** on GPUs



Results: 1-2 Magnitudes faster than LP based verifiers



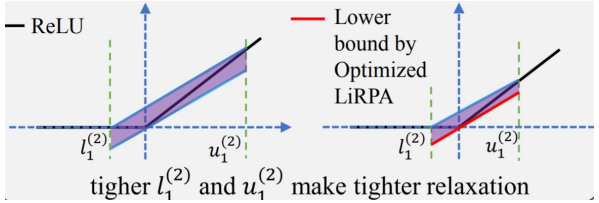
Similar or lower timeout rates

Ablation study – Optimized bounds are critical!

Keys to success:

1. CROWN/LiRPA is ~100x faster than LP verifier; Massively parallelizable on GPUs;

2. Optimize bounds together with intermediate layer bounds for each layer, allowing tightening the relaxation and even be tighter than regular LP verifiers



CROWN cannot detect infeasible split constraints => occasional usage of LP in BaB to guarantee completeness

NEW!

See our **new paper** **β-CROWN** which **eliminates LP** and uses bound propagation to achieve SOTA in both **complete** and **incomplete** verification

https://arxiv.org/pdf/**2103.06624**

[1] "Efficient Neural Network Robustness Certification with General Activation Functions", Zhang, Huan et al. NeurIPS 2018
[2] "Automatic Perturbation Analysis for Scalable Certified Robustness and Beyond.", Xu, Kaidi et al., NeurIPS 2020
[3] "A unified view of piecewise linear neural network verification. Bunel, Rudy, et al. NeurIPS. 2018.