Multi-path Bound Propagation for Neural Network Verification

Ye Zheng

Safety of Neural Networks

- Neural networks are sensible to natural or adversarial attack
- Safety need to be guaranteed in safety-critical scenarios
- Testing based methods can not provide safety guarantee (infinite images)

Autonomous-driving accident

Adversarial traffic sign

Image source: https://en.wikipedia.org/wiki/Smart_system; images from Google Image

Neural Network Verification

- Verifies whether a region input results in unsafe outputs
- Input: **a region** (determined by an original image and a perturbation size)
- Output: **safe or unsafe** (need to compute the reachable region)

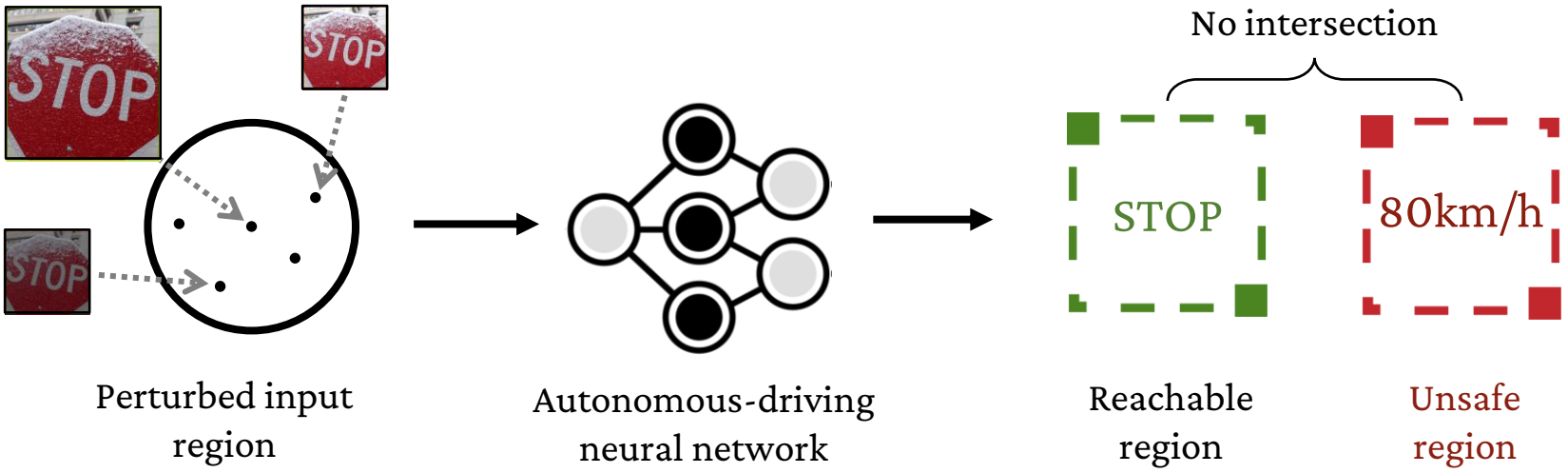


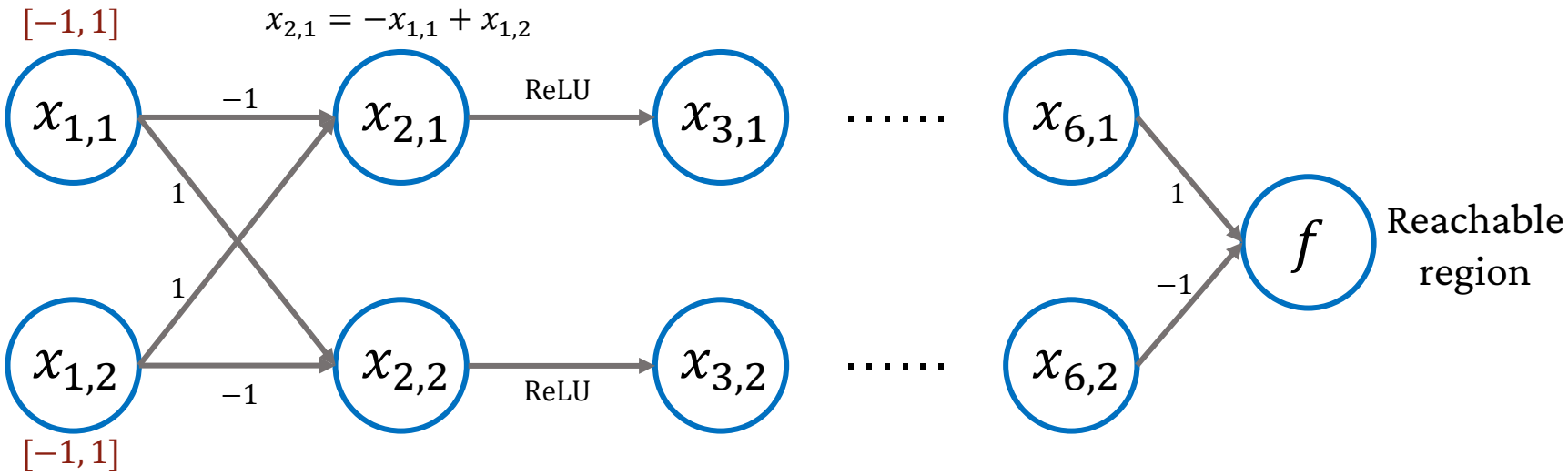
Image source: https://www.businessinsider.com/why-are-stop-signs-red

Neural Network Verification

- Difficulty: the composition of non-linear activations (e.g. ReLU) (NP-hard)
- Methods
  - Constraint solving
    - Encode the network as constraints and check the satisfiability with...
    - NP-hard
  - Bound propagation
    - Easier, but more efficient

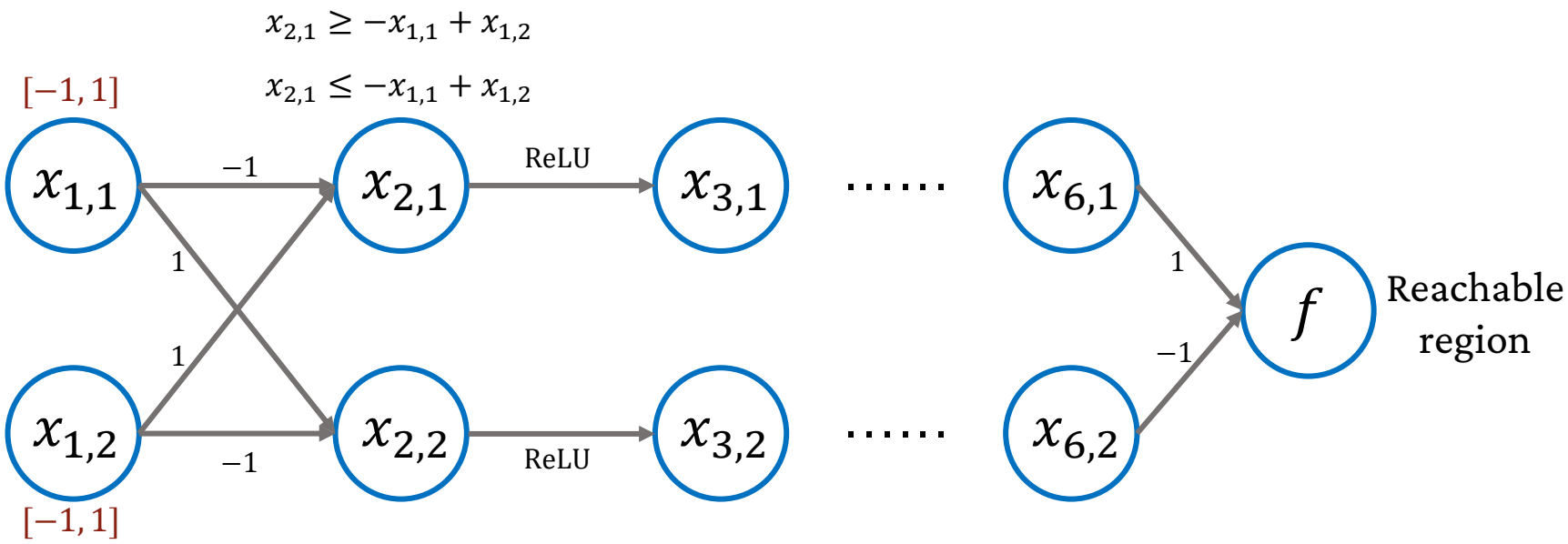
Bound Propagation

- Propagates bound functions along the neural network
- Bound function is a pair of linear bounds



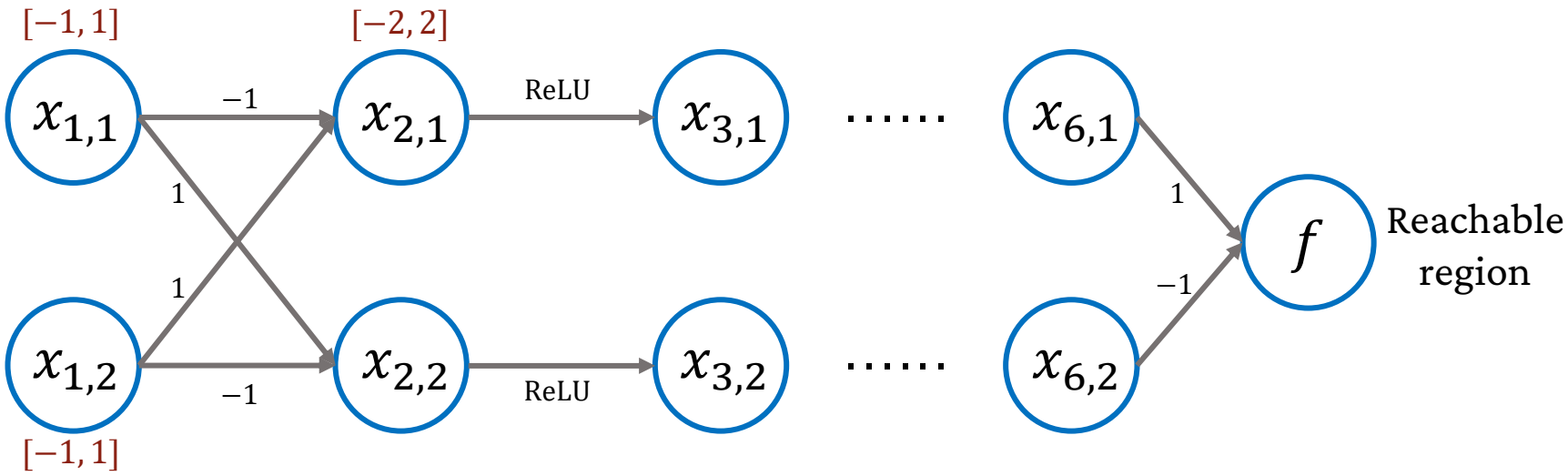
Bound Propagation

- Propagates bound functions along the neural network
- Bound function is a pair of linear bounds



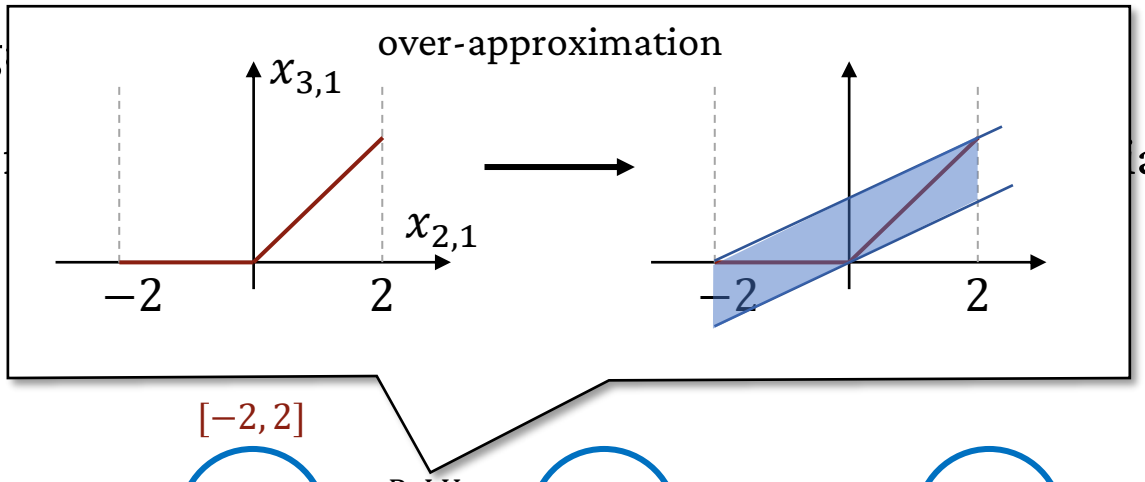
Bound Propagation

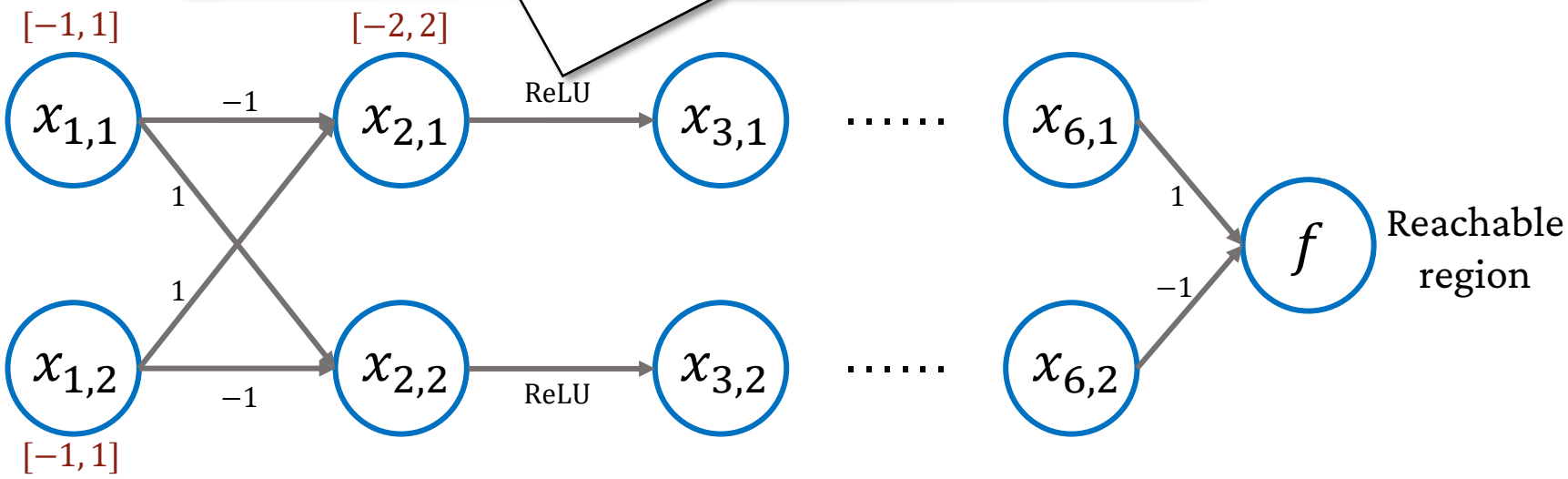
- Propagates bound functions along the neural network
- Bound function is a pair of linear bounds



Bound Propagation

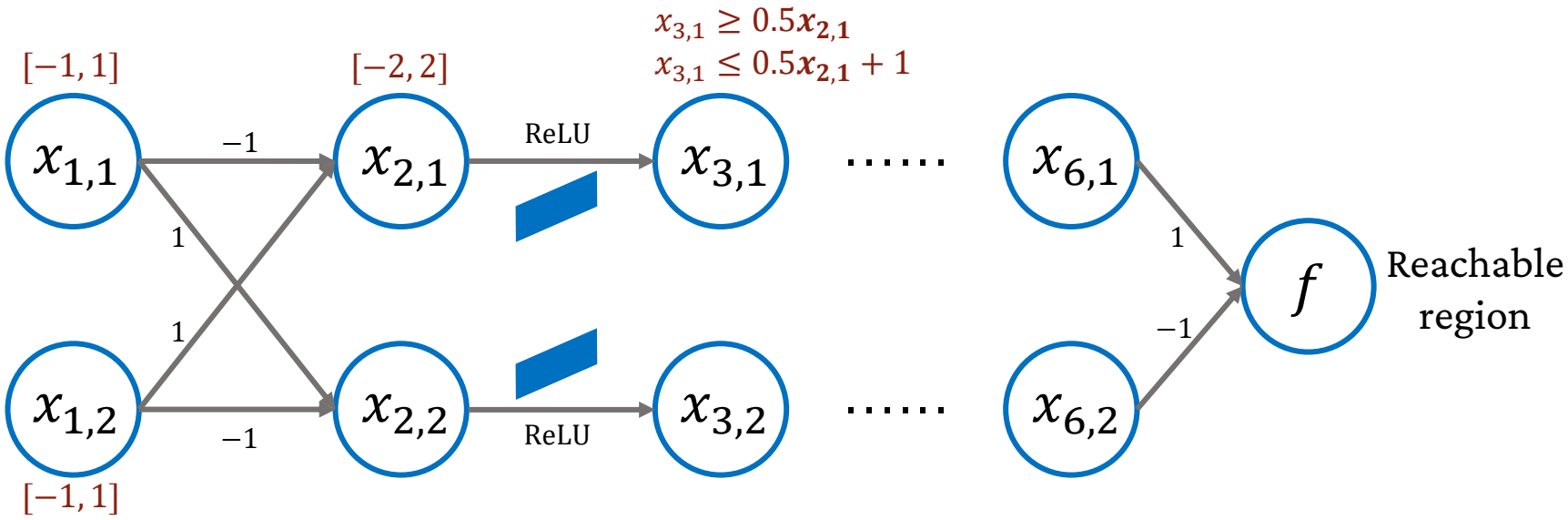
- Propag
- Bound ables





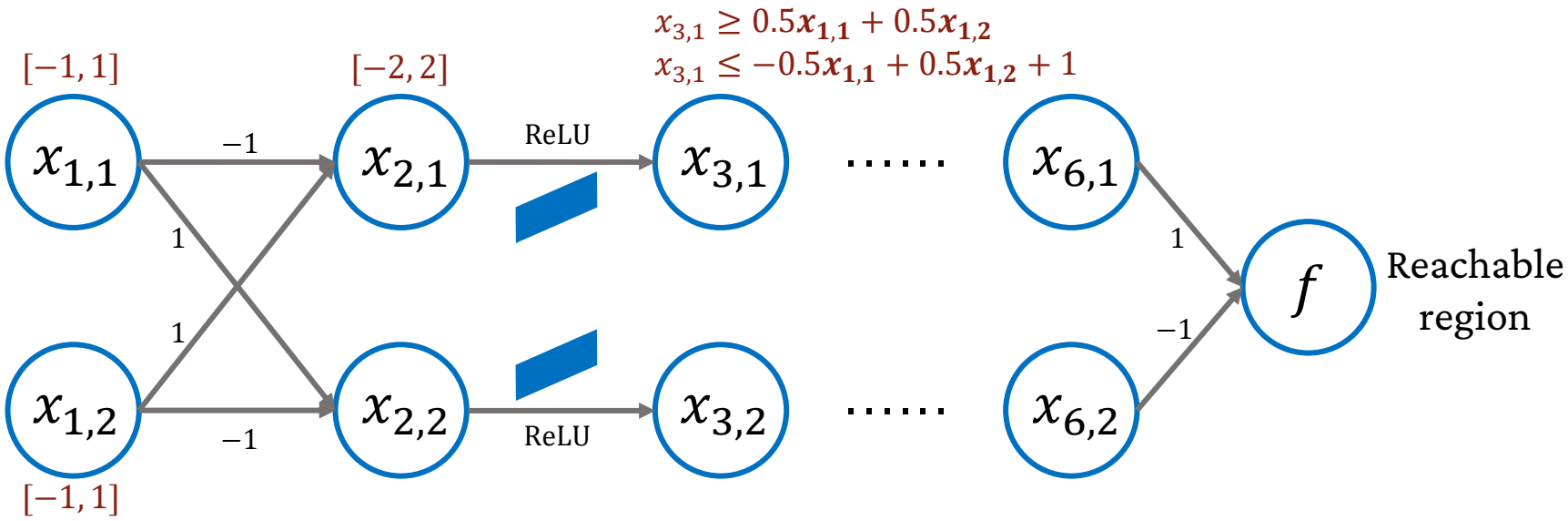
Bound Propagation

- Propagates bound functions along the neural network
- Bound function is a pair of linear bounds



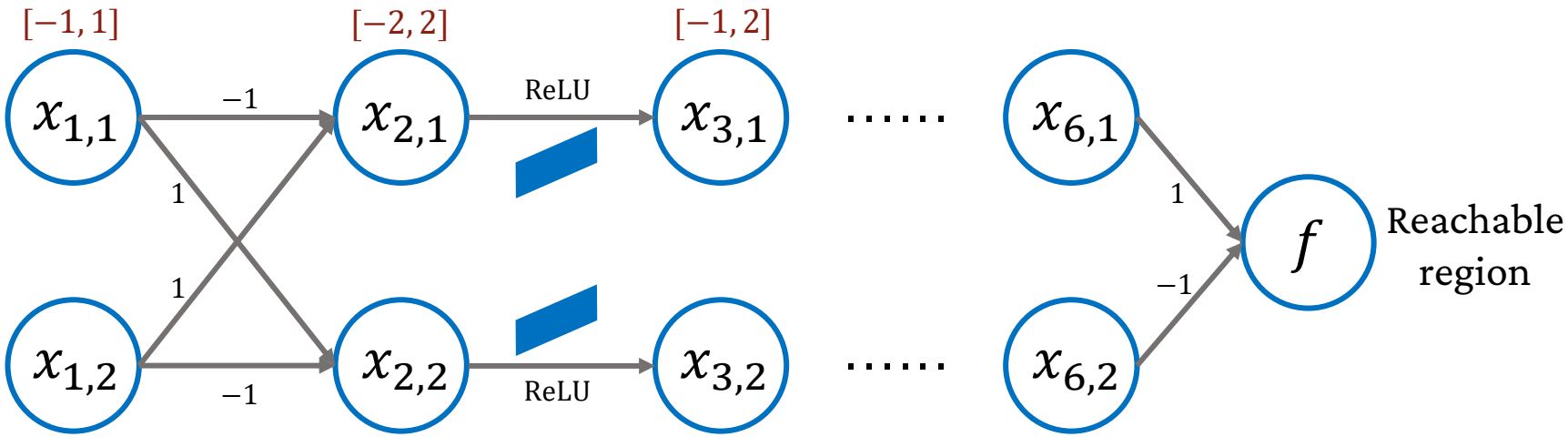
Bound Propagation

- Propagates bound functions along the neural network
- Bound function is a pair of linear bounds



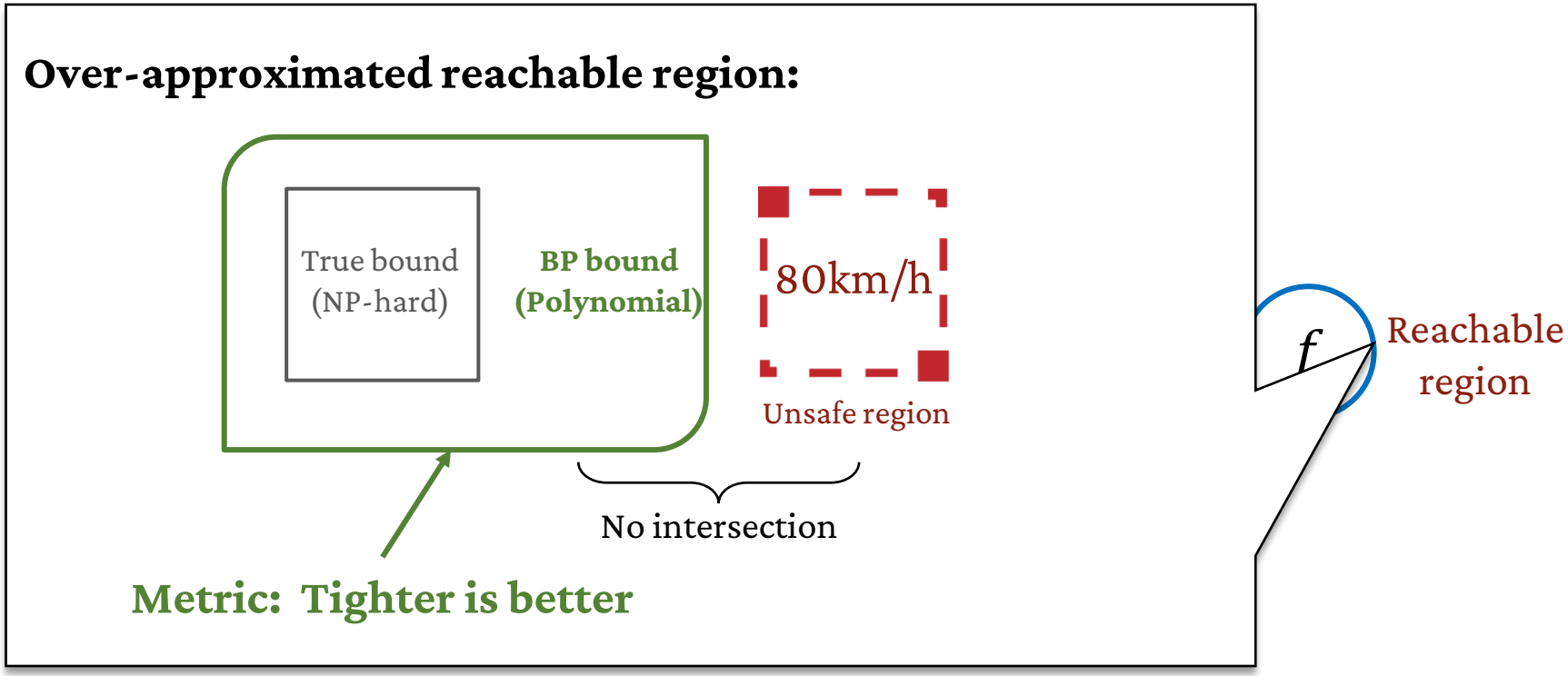
Bound Propagation

- Propagates bound functions along the neural network
- Bound function is a pair of linear bounds

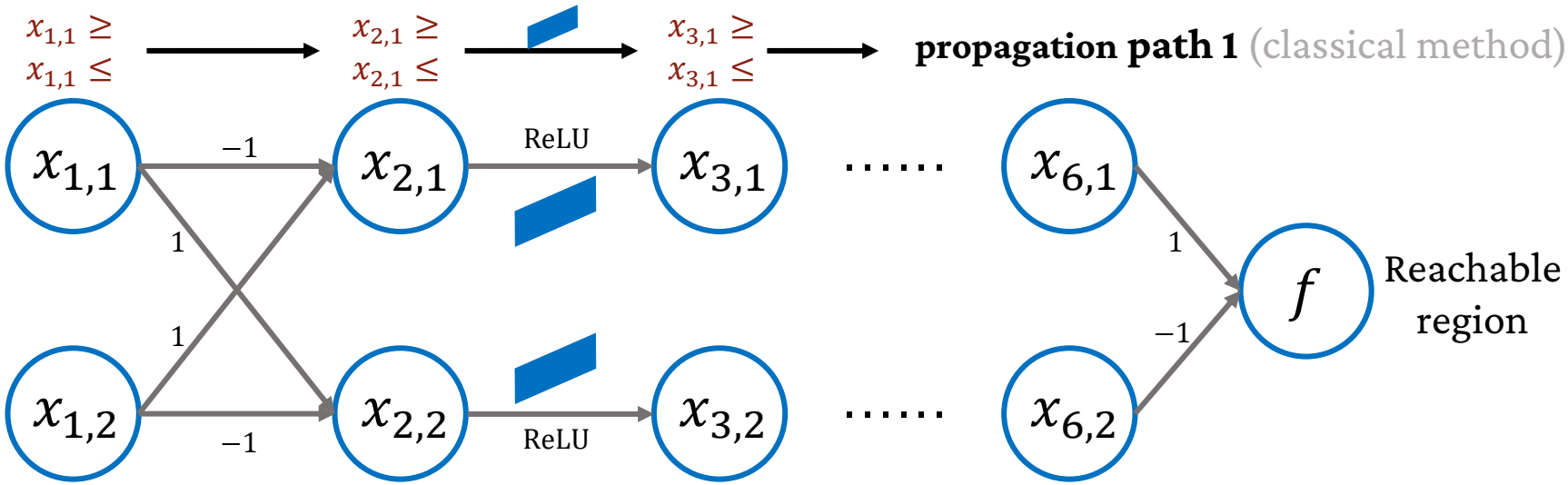


Bound Propagation

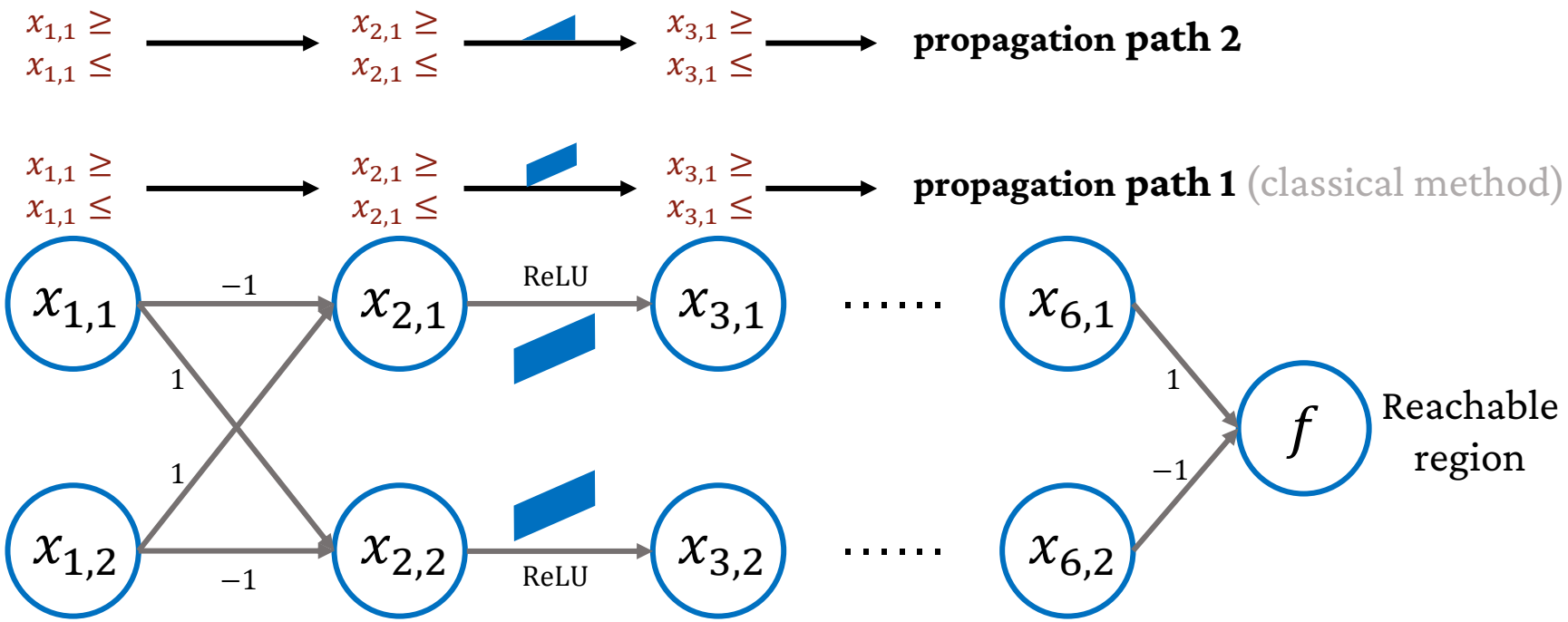
- Propagates bound functions along the neural network
- Bound function is a pair of linear bounds



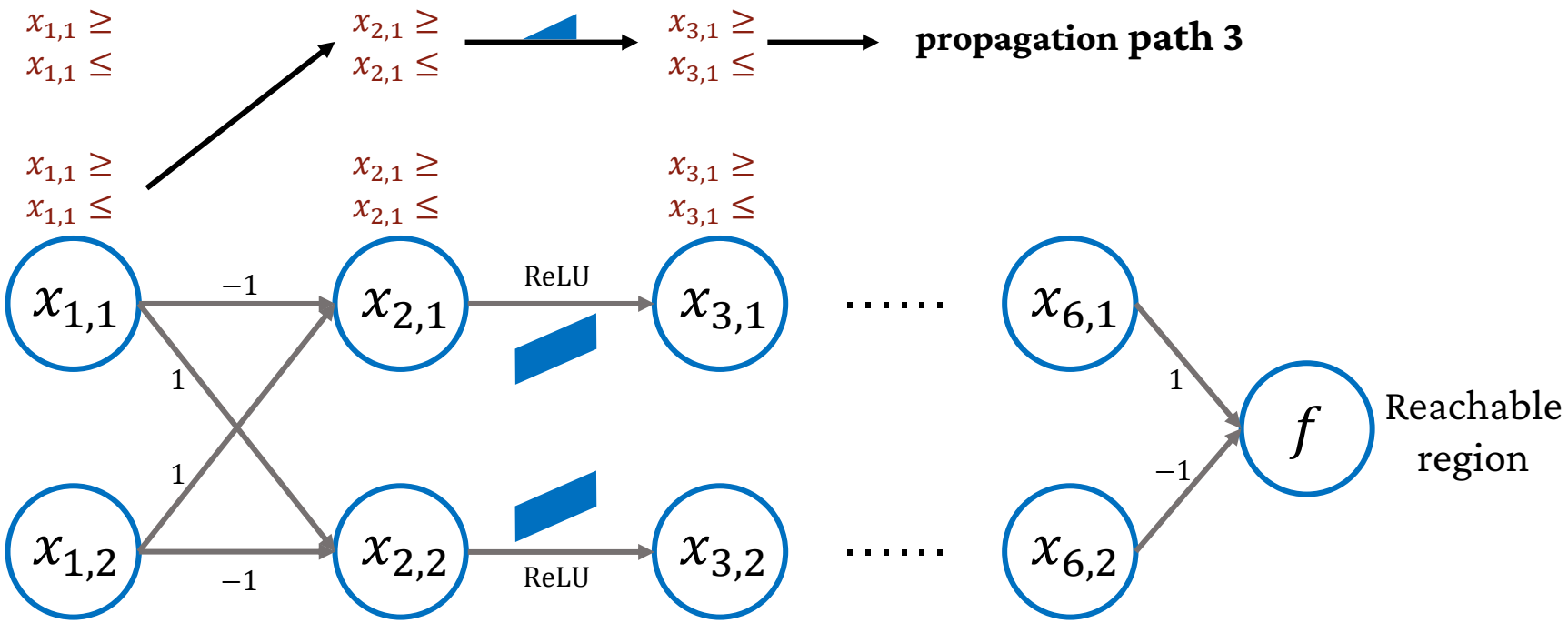
Our: Bound Propagation Path



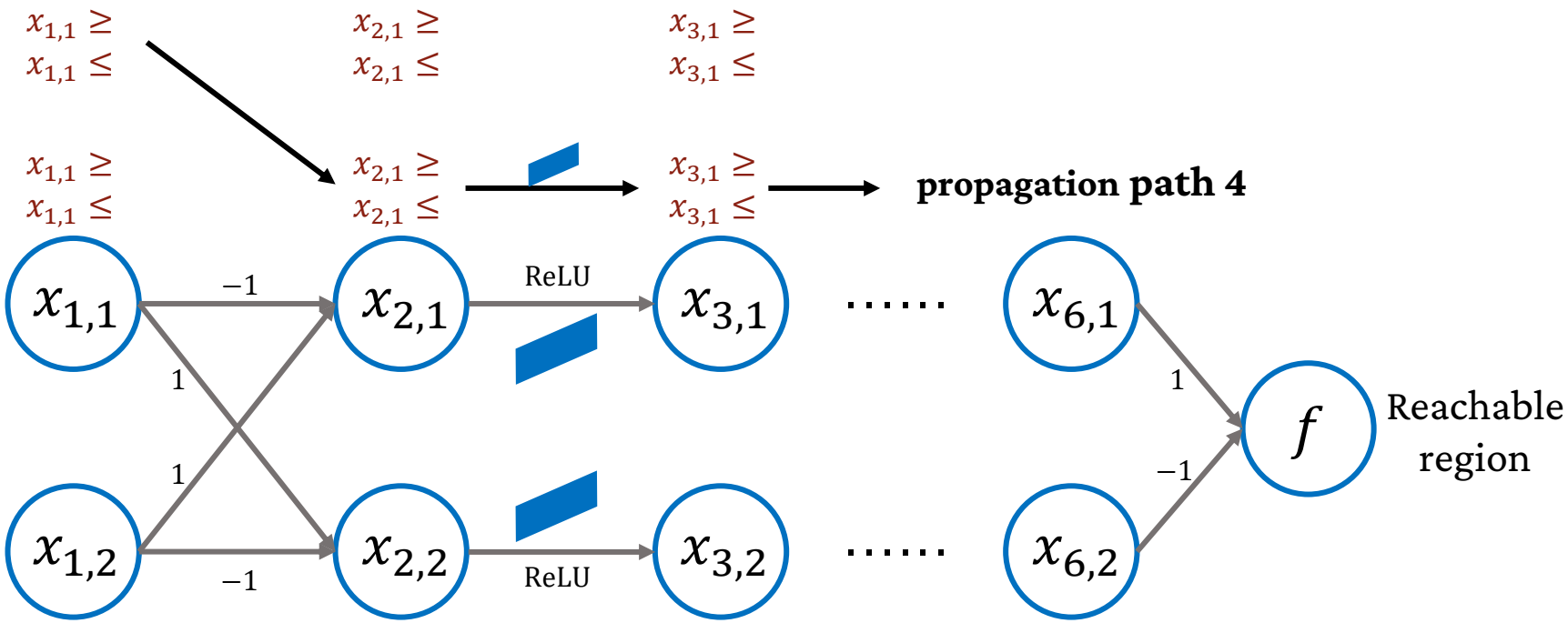
Our: Two-path Bound Propagation



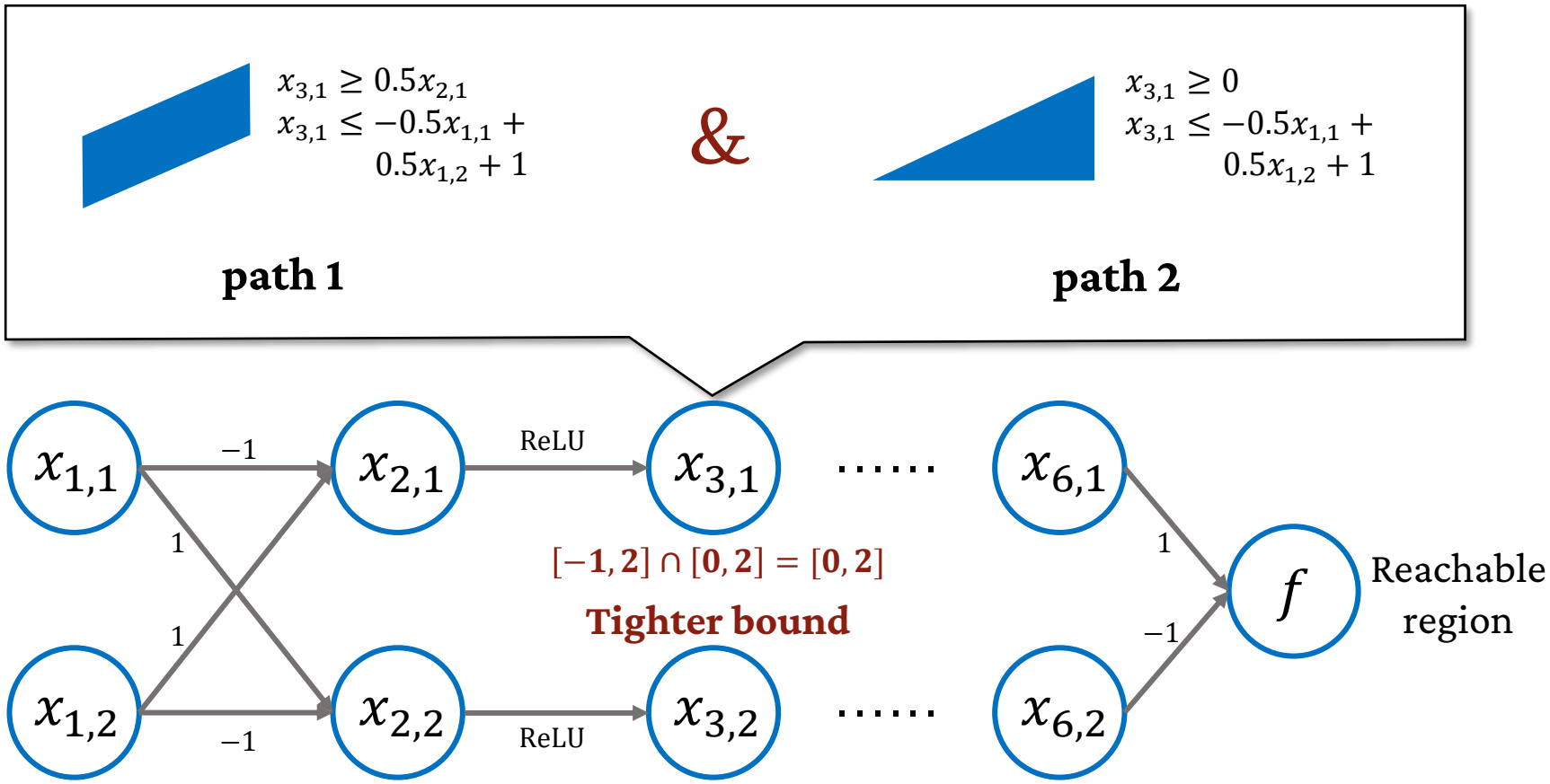
Our: Two-path Bound Propagation



Our: Two-path Bound Propagation



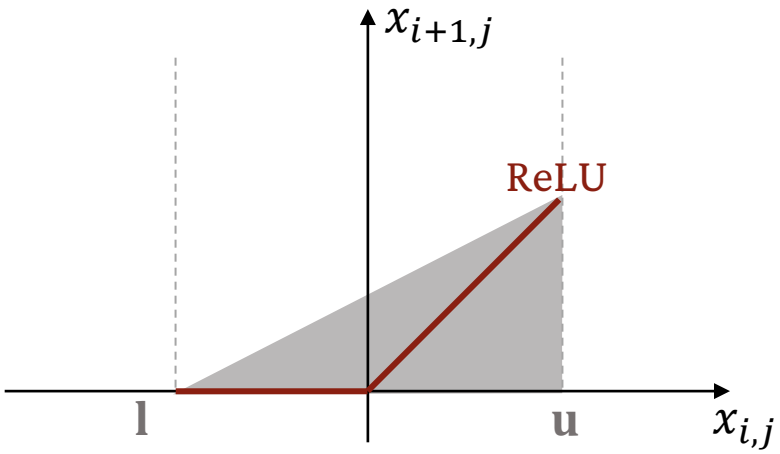
Our: Two-path Bound Propagation



Advantage: Accuracy Accumulation

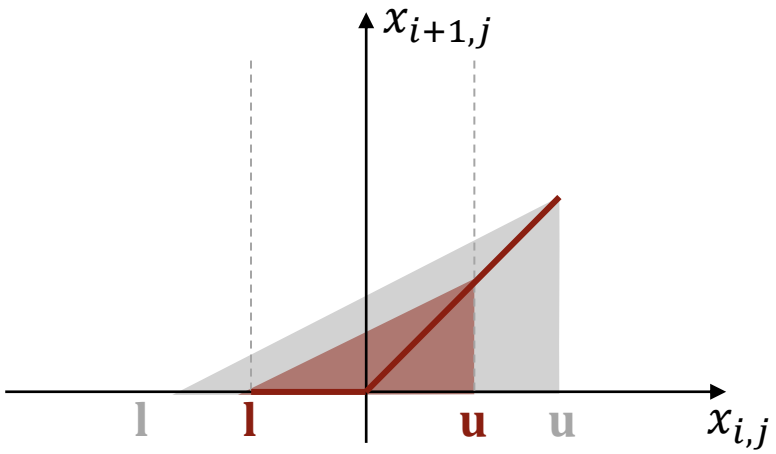
- For each neuron
  - More accurate bound and over-approximation

Classical: One bound



Loose approximation

Two bounds

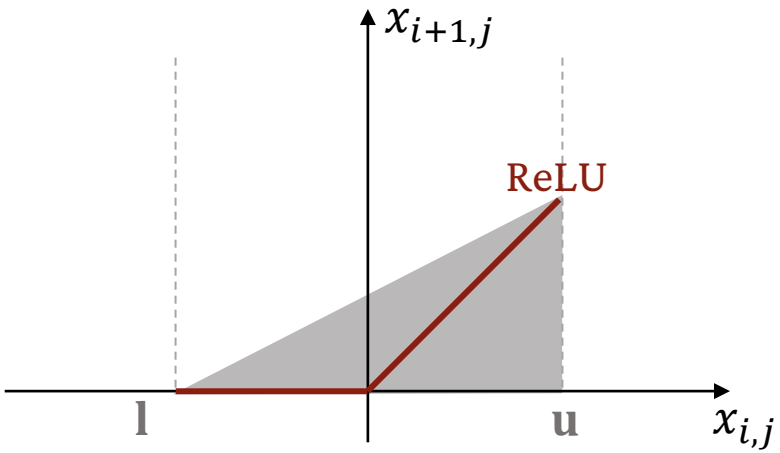


Tighter approximation

Advantage: Accuracy Accumulation

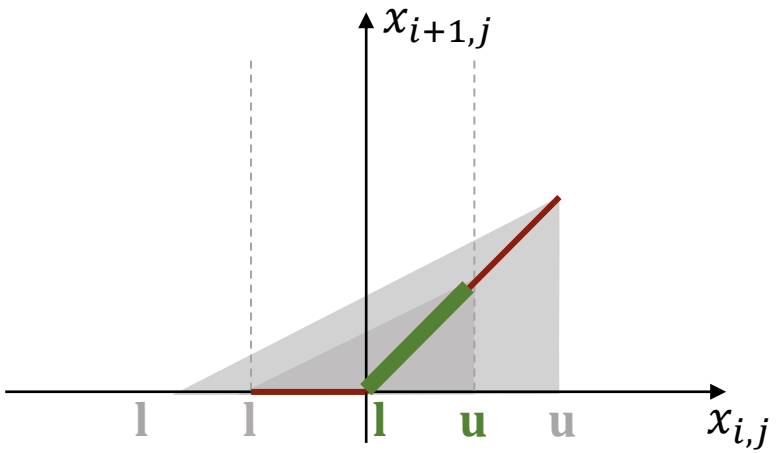
- For each neuron
  - More accurate bound and over-approximation

Classical: One bound



Loose approximation

More bounds ...



No approximation

Multi-path Bound Propagation for Neural Network Verification

Safety of Neural Networks

- Neural networks are sensible to natural or adversarial attack
- Safety need to be guaranteed in safety-critical scenarios
- Testing based methods can not provide safety guarantee

Autonomous-driving accident

Adversarial traffic sign

Image source: https://en.wikipedia.org/wiki/Smart_system; images from Google Image

Neural Network Verification

- Verifies whether a region input results in unsafe outputs
- Difficulty: the composition of non-linear activations (e.g. ReLU)

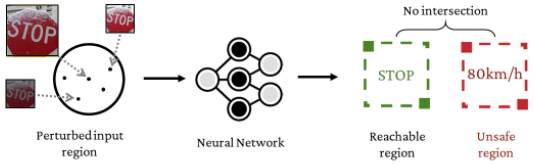
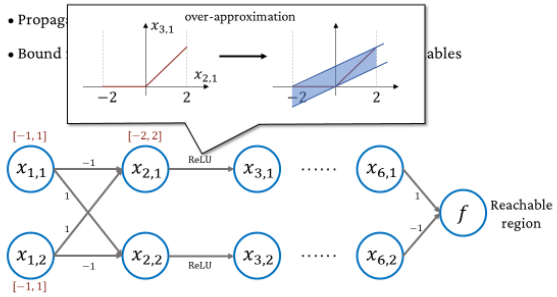


Image source: https://www.businessinsider.com/why-are-stop-signs-red

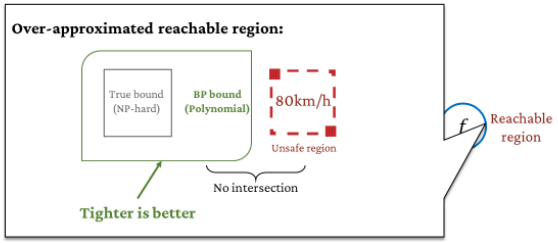
Bound Propagation

- Propag
- Bound ables

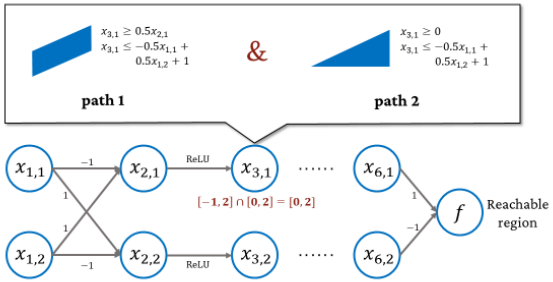


Bound Propagation

- Propagates bound functions along the neural network
- Bound function is a pair of linear bounds

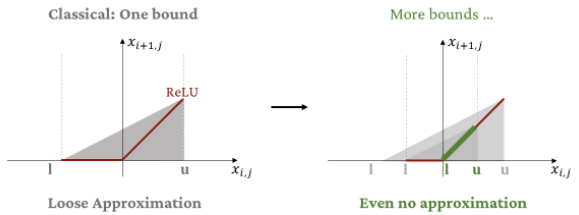


Our: Two-path Bound Propagation



Advantage: Accuracy Accumulation

- For each neuron
  - More accurate bound and over-approximation



Thank you!