Generalization Guarantees for Learning Score-Based Branch-and-Cut Policies

From Generalization Theory to Algorithm Design in Branch-and-Cut

Hongyu Cheng and Amitabh Basu

November 6, 2025

Department of Applied Mathematics and Statistics
Johns Hopkins University

Motivation and Background

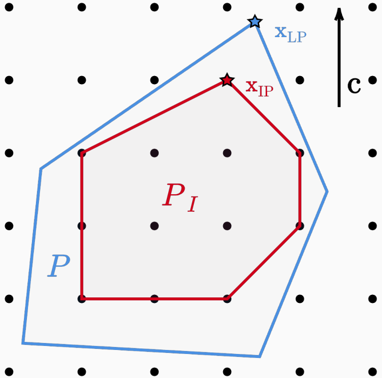
Mixed-Integer Linear Programs: General Formulation

- **General Form:**

$$\max_x \ c^\top x, \quad \text{s.t.} \ Ax \le b, \quad x \in \mathbb{Z}^n \times \mathbb{R}^d$$

  - $\mathcal{P} = \{x : Ax \le b\}$
  - $S = \mathcal{P} \cap (\mathbb{Z}^n \times \mathbb{R}^d)$
  - $\mathcal{P}_I = \text{conv}(S)$

- **Solution Approach:**
  - Branch-and-Bound
  - Cutting Planes



Branch-and-Bound

- If $x^\star_{\mathrm{LP}}$ has fractional coordinates, select a fractional variable (e.g., $x_i^\star$) and partition:

$$\mathcal{P} \to \mathcal{P}_1 \cup \mathcal{P}_2$$

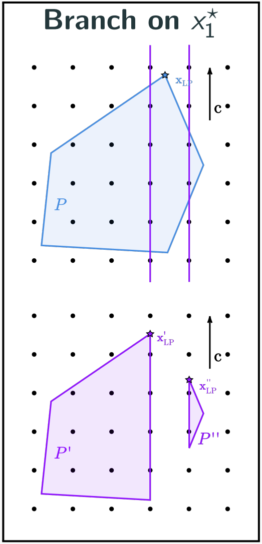
$$\mathcal{P}_1 = \mathcal{P} \cap \{x : x_i \leq \lfloor x_i^\star \rfloor\}$$

$$\mathcal{P}_2 = \mathcal{P} \cap \{x : x_i \geq \lceil x_i^\star \rceil\}$$

- **Critical Decision:** Which fractional variable to branch on?

$$\min_i \ \max\{z_{\mathrm{LP}}(\mathcal{P}_1), z_{\mathrm{LP}}(\mathcal{P}_2)\}$$

(e.g., strong branching)



Branch-and-Bound

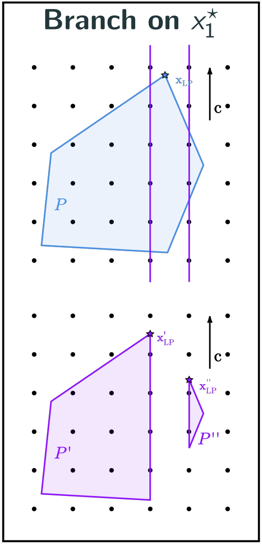
- If $x^\star_{\mathrm{LP}}$ has fractional coordinates, select a fractional variable (e.g., $x_i^\star$) and partition:

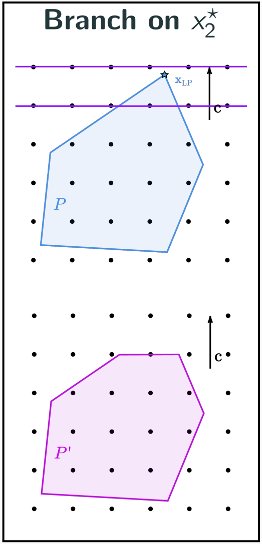
  $$\mathcal{P} \to \mathcal{P}_1 \cup \mathcal{P}_2$$
  $$\mathcal{P}_1 = \mathcal{P} \cap \{x : x_i \leq \lfloor x_i^\star \rfloor\}$$
  $$\mathcal{P}_2 = \mathcal{P} \cap \{x : x_i \geq \lceil x_i^\star \rceil\}$$

- **Critical Decision:** Which fractional variable to branch on?

  $$\min_i \ \max\{z_{\mathrm{LP}}(\mathcal{P}_1), z_{\mathrm{LP}}(\mathcal{P}_2)\}$$

  (e.g., strong branching)





Cutting Planes

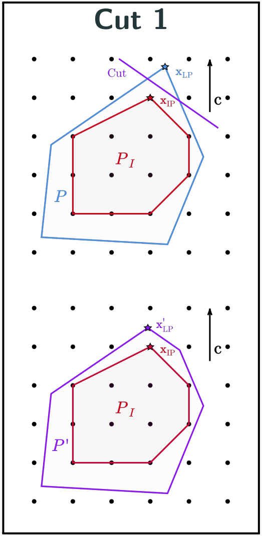
- A cut $\pi^\top x \geq 1$ is valid for $\mathcal{P}_I$ if it satisfies all integer points:

$$\pi^\top x \geq 1 \quad \forall x \in \mathcal{P}_I$$

- And ideally, cuts off the current LP solution:

$$\pi^\top x^\star_{\mathrm{LP}} < 1$$

- **Critical Decision:** From numerous candidate cuts in pool $\mathcal{C}$, select $S \subseteq \mathcal{C}$ to add.



Cutting Planes

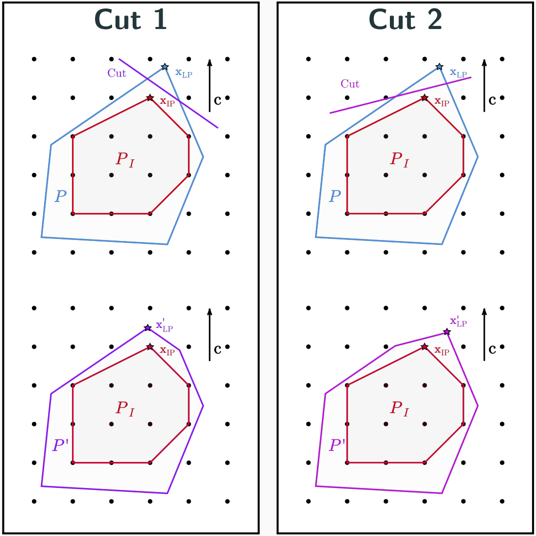
- A cut $\pi^\top x \geq 1$ is valid for $\mathcal{P}_I$ if it satisfies all integer points:

$$\pi^\top x \geq 1 \quad \forall x \in \mathcal{P}_I$$

- And ideally, cuts off the current LP solution:

$$\pi^\top x^\star_{\mathrm{LP}} < 1$$

- **Critical Decision:** From numerous candidate cuts in pool $\mathcal{C}$, select $S \subseteq \mathcal{C}$ to add.

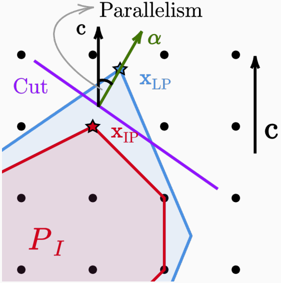


Traditional Approach: Hand-Crafted Scoring Functions

- State $s \in \mathcal{S}$: current B&C state
- Action $a \in \mathcal{A}$: candidate cut $\boldsymbol{\alpha}^\top \mathbf{x} \leq \beta$
- Hand-crafted score vector: $\phi : \mathcal{S} \times \mathcal{A} \to \mathbb{R}^\ell$.

Traditional Approach: Hand-Crafted Scoring Functions

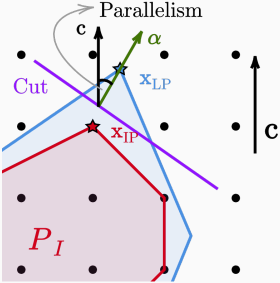
- State $s \in \mathcal{S}$: current B&C state
- Action $a \in \mathcal{A}$: candidate cut $\boldsymbol{\alpha}^\top \mathbf{x} \leq \beta$
- Hand-crafted score vector: $\phi : \mathcal{S} \times \mathcal{A} \to \mathbb{R}^\ell$.

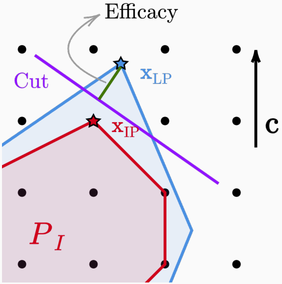


$$\phi_{\text{obj}}(s, a) = \frac{|\boldsymbol{\alpha}^T \mathbf{c}|}{\|\boldsymbol{\alpha}\|_2 \|\mathbf{c}\|_2}$$

Traditional Approach: Hand-Crafted Scoring Functions

- State $s \in \mathcal{S}$: current B&C state
- Action $a \in \mathcal{A}$: candidate cut $\boldsymbol{\alpha}^\top \mathbf{x} \leq \beta$
- Hand-crafted score vector: $\phi : \mathcal{S} \times \mathcal{A} \to \mathbb{R}^\ell$.



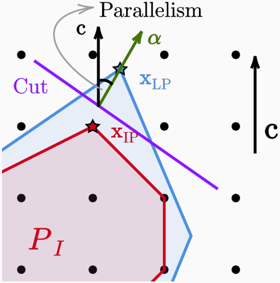


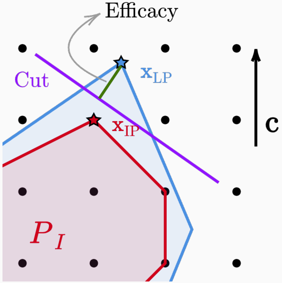
$$\phi_{\text{obj}}(s, a) = \frac{|\boldsymbol{\alpha}^T \mathbf{c}|}{\|\boldsymbol{\alpha}\|_2 \|\mathbf{c}\|_2}$$

$$\phi_{\text{eff}}(s, a) = \frac{\boldsymbol{\alpha}^T \mathbf{x}_{\text{LP}} - \beta}{\|\boldsymbol{\alpha}\|_2}$$

Traditional Approach: Hand-Crafted Scoring Functions

- State $s \in \mathcal{S}$: current B&C state
- Action $a \in \mathcal{A}$: candidate cut $\boldsymbol{\alpha}^\top \mathbf{x} \leq \beta$
- Hand-crafted score vector: $\phi : \mathcal{S} \times \mathcal{A} \to \mathbb{R}^\ell$.





$$\phi_{\text{obj}}(s, a) = \frac{|\boldsymbol{\alpha}^T \mathbf{c}|}{\|\boldsymbol{\alpha}\|_2 \|\mathbf{c}\|_2} \qquad \phi_{\text{eff}}(s, a) = \frac{\boldsymbol{\alpha}^T \mathbf{x}_{\text{LP}} - \beta}{\|\boldsymbol{\alpha}\|_2}$$

- SCIP: $\phi(s, a) = \begin{bmatrix} \phi_{\text{obj}}(s, a) & \phi_{\text{eff}}(s, a) & \phi_{\text{dcd}}(s, a) & \phi_{\text{int}}(s, a) \end{bmatrix}^\top$

Cut Selection in SCIP

- Score function, based on hand-tuned weights $\mathbf{w} \in \mathbb{R}^4_+$:

$$\begin{aligned} f_{\text{SCIP}}(s, a, \mathbf{w}) &= \mathbf{w}_1 \phi_{\text{obj}}(s, a) + \mathbf{w}_2 \phi_{\text{eff}}(s, a) + \mathbf{w}_3 \phi_{\text{dcd}}(s, a) + \mathbf{w}_4 \phi_{\text{int}}(s, a) \\ &= \mathbf{w}^\top \phi(s, a) \end{aligned}$$

- At state s, select cut $a = (\boldsymbol{\alpha}, \beta)$ with the highest score:

$$a^* \in \arg\max_{a \in \mathcal{A}^s} f_{\text{SCIP}}(s, a, \mathbf{w})$$

Cut Selection in SCIP

- Score function, based on hand-tuned weights $\mathbf{w} \in \mathbb{R}^4_+$:

$$
\begin{aligned}
f_{\text{SCIP}}(s, a, \mathbf{w}) &= \mathbf{w}_1 \phi_{\text{obj}}(s, a) + \mathbf{w}_2 \phi_{\text{eff}}(s, a) + \mathbf{w}_3 \phi_{\text{dcd}}(s, a) + \mathbf{w}_4 \phi_{\text{int}}(s, a) \\
&= \mathbf{w}^\top \phi(s, a)
\end{aligned}
$$

- At state s, select cut $a = (\boldsymbol{\alpha}, \beta)$ with the highest score:

$$
a^* \in \underset{a \in \mathcal{A}^s}{\arg\max}\, f_{\text{SCIP}}(s, a, \mathbf{w})
$$

Problems with this approach:

- $\mathbf{w}$ are manually tuned (good for some, not for others)
- Linear model $f_{\text{SCIP}}(s, a, \mathbf{w}) = \mathbf{w}^\top \phi(s, a)$ is less powerful

An Example: Limitations of SCIP's Cut Scoring Function

- LP objective value improvement, $\Delta(s, a)$, is the most popular performance metric for cut selection ($s = (A, b, c)$, $a = (\boldsymbol{\alpha}, \beta)$).

An Example: Limitations of SCIP's Cut Scoring Function

- LP objective value improvement, $\Delta(s, a)$, is the most popular performance metric for cut selection ($s = (A, b, c)$, $a = (\alpha, \beta)$).
- Consider $\max \left\{ x_n : \sum_{i=1}^{n} x_i = r, x \in \mathbb{R}^n_+ \right\}$, for some $r > 0$.

  It's not hard to verify that

$$\Delta(s, a) = \frac{\phi_{\text{eff}}(s, a)}{\phi_{\text{par}}(s, a) - \sqrt{1 - \phi_{\text{par}}(s, a)^2 / (n - 1)}}$$

An Example: Limitations of SCIP's Cut Scoring Function

- LP objective value improvement, $\Delta(s, a)$, is the most popular performance metric for cut selection ($s = (A, b, c)$, $a = (\alpha, \beta)$).
- Consider $\max \left\{ x_n : \sum_{i=1}^{n} x_i = r, x \in \mathbb{R}_+^n \right\}$, for some $r > 0$.

  It's not hard to verify that

$$\Delta(s, a) = \frac{\phi_{\text{eff}}(s, a)}{\phi_{\text{par}}(s, a) - \sqrt{1 - \phi_{\text{par}}(s, a)^2 / (n - 1)}}$$

- There is no $\mathbf{w}$ such that $f_{\text{SCIP}}(s, a, \mathbf{w}) = \mathbf{w}^\top \phi(s, a)$ approximates $\Delta(s, a)$ well for all (s, a) in this family.

  Also, there exists s, a^1, a^2 such that

  $\Delta(s, a^1) > \Delta(s, a^2)$ but $f_{\text{SCIP}}(s, a^1, \mathbf{w}) < f_{\text{SCIP}}(s, a^2, \mathbf{w})$ for all $\mathbf{w}$.

An Example: Limitations of SCIP's Cut Scoring Function

- LP objective value improvement, $\Delta(s, a)$, is the most popular performance metric for cut selection ($s = (A, b, c)$, $a = (\alpha, \beta)$).
- Consider $\max \left\{ x_n : \sum_{i=1}^{n} x_i = r, x \in \mathbb{R}_+^n \right\}$, for some $r > 0$.
  It's not hard to verify that

$$\Delta(s, a) = \frac{\phi_{\text{eff}}(s, a)}{\phi_{\text{par}}(s, a) - \sqrt{1 - \phi_{\text{par}}(s, a)^2}/(n-1)}$$

- There is no $\mathbf{w}$ such that $f_{\text{SCIP}}(s, a, \mathbf{w}) = \mathbf{w}^\top \phi(s, a)$ approximates $\Delta(s, a)$ well for all (s, a) in this family.
  Also, there exists s, a^1, a^2 such that
  $\Delta(s, a^1) > \Delta(s, a^2)$ but $f_{\text{SCIP}}(s, a^1, \mathbf{w}) < f_{\text{SCIP}}(s, a^2, \mathbf{w})$ for all $\mathbf{w}$.

$f_{\text{SCIP}}(s, a, \mathbf{w}) = \mathbf{w}^\top \phi(s, a) \longrightarrow f(s, a, \mathbf{w}) = h(\phi(s, a), \mathbf{w})$,
where $h(\cdot, \mathbf{w}) : \mathbb{R}^\ell \to \mathbb{R}$ is an expressive, nonlinear function.

Recent ML Approaches under The Unified Framework

Cut Selection: $s =$ state, $a =$ cut

- **Huang+ 2022:** $f_{\text{cut}}(s, a, \mathbf{w}) = \text{MLP}(\phi_{\text{cut}}(s, a), \mathbf{w})$

  $\phi_{\text{cut}} \in \mathbb{R}^{14}$: efficacy, support, dynamism, etc.

Recent ML Approaches under The Unified Framework

Cut Selection: s = state, a = cut

- **Huang+ 2022:** $f_{\text{cut}}(s, a, \mathbf{w}) = \text{MLP}(\phi_{\text{cut}}(s, a), \mathbf{w})$

  $\phi_{\text{cut}} \in \mathbb{R}^{14}$: efficacy, support, dynamism, etc.

- **Paulus+ 2022:** $f_{\text{cut}}(s, a, \mathbf{w}) = \text{GNN+Attn}_{\mathbf{w}'}(\phi_{\text{cut}}(s, a), \mathbf{w})$

  ϕ_{cut}: tripartite graph encoding

Node Selection: s = B&C state, a = node

- **Yilmaz+ 2021:** $f_{\text{node}}(s, a, \mathbf{w}) = \text{MLP}(\phi_{\text{node}}(s, a), \mathbf{w})$

  $\phi_{\text{node}} \in \mathbb{R}^{29}$: basis status, bounds, depth

Branching: s = state, a = variable

- **Khalil+ 2016:** $f_{\text{branch}}(s, a, \mathbf{w}) = \mathbf{w}^T \phi_{\text{branch}}(s, a)$ (linear)

  $\phi_{\text{branch}} \in \mathbb{R}^{72}$: variable features
- **Gasse+ 2019:** $f_{\text{branch}}(s, a, \mathbf{w}) = \text{GCNN}(\phi_{\text{branch}}(s, a), \mathbf{w})$

  ϕ_{branch}: bipartite LP graph features

Theoretical Challenges in Learning for Branch-and-Cut

- Empirical success $\implies$ learning works in practice!
- Learning theory problems:

Theoretical Challenges in Learning for Branch-and-Cut

- Empirical success $\implies$ learning works in practice!
- Learning theory problems:
  - Can learned parameters $\mathbf{w}$ generalize from finite training data?

Theoretical Challenges in Learning for Branch-and-Cut

- Empirical success $\implies$ learning works in practice!
- Learning theory problems:
  - Can learned parameters $\mathbf{w}$ generalize from finite training data?
    Generalization Error $\implies$ this talk!

Theoretical Challenges in Learning for Branch-and-Cut

- Empirical success $\implies$ learning works in practice!
- Learning theory problems:
  - Can learned parameters $\mathbf{w}$ generalize from finite training data?
    Generalization Error $\implies$ this talk!
  - How to design a hypothesis class $\{h(\cdot, \mathbf{w}) : \mathbb{R}^\ell \mid \mathbf{w} \in \mathbb{R}^W\}$ such that some $h(\cdot, \mathbf{w})$ imitates the "expert" scoring function well?

Theoretical Challenges in Learning for Branch-and-Cut

- Empirical success $\implies$ learning works in practice!
- Learning theory problems:
  - Can learned parameters $\mathbf{w}$ generalize from finite training data?
    Generalization Error $\implies$ this talk!
  - How to design a hypothesis class $\{h(\cdot, \mathbf{w}) : \mathbb{R}^\ell \mid \mathbf{w} \in \mathbb{R}^W\}$ such that some $h(\cdot, \mathbf{w})$ imitates the "expert" scoring function well?
    Approximation Error $\implies$ working paper!

Theoretical Challenges in Learning for Branch-and-Cut

- Empirical success $\implies$ learning works in practice!
- Learning theory problems:
  - Can learned parameters $\mathbf{w}$ generalize from finite training data?
    Generalization Error $\implies$ this talk!
  - How to design a hypothesis class $\{h(\cdot, \mathbf{w}) : \mathbb{R}^\ell \mid \mathbf{w} \in \mathbb{R}^W\}$ such that some $h(\cdot, \mathbf{w})$ imitates the "expert" scoring function well?
    Approximation Error $\implies$ working paper!
  - How to efficiently find $\mathbf{w}$ minimizing training error?

Theoretical Challenges in Learning for Branch-and-Cut

- Empirical success $\implies$ learning works in practice!
- Learning theory problems:
  - Can learned parameters $\mathbf{w}$ generalize from finite training data?
    Generalization Error $\implies$ this talk!
  - How to design a hypothesis class $\{h(\cdot, \mathbf{w}) : \mathbb{R}^\ell \mid \mathbf{w} \in \mathbb{R}^W\}$ such that some $h(\cdot, \mathbf{w})$ imitates the "expert" scoring function well?
    Approximation Error $\implies$ working paper!
  - How to efficiently find $\mathbf{w}$ minimizing training error?
    Optimization Error $\implies$ no clue and no hope!!!

Generalization Error: Formulation

- Let $V(I, \mathbf{w})$ be the B&C tree size (e.g., number of nodes) on instance I using the scoring function parameterized by $\mathbf{w}$.
- Assume instances $I_1, \ldots, I_N$ are i.i.d. samples from an unknown distribution $\mathcal{D}$.
- We are interested in the uniform convergence:

$$\sup_{\mathbf{w} \in \mathcal{W}} \left| \frac{1}{N} \sum_{i=1}^{N} V(I_i, \mathbf{w}) - \mathbb{E}_{I \sim \mathcal{D}}[V(I, \mathbf{w})] \right|$$

Generalization Error: Formulation

Why uniform convergence?

$$\sup_{\mathbf{w}\in\mathcal{W}} \left| \frac{1}{N}\sum_{i=1}^{N} V(I_i,\mathbf{w}) - \mathbb{E}_{I\sim\mathcal{D}}[V(I,\mathbf{w})] \right| = \varepsilon_N$$

$$\implies \left| \frac{1}{N}\sum_{i=1}^{N} V(I_i^{\text{train}},\mathbf{w}) - \frac{1}{N}\sum_{j=1}^{N} V(I_j^{\text{test}},\mathbf{w}) \right| \le 2\varepsilon_N, \quad \forall \mathbf{w}\in\mathcal{W}$$

Generalization Error: Formulation

Why uniform convergence?

$$\sup_{\mathbf{w}\in\mathcal{W}} \left| \frac{1}{N} \sum_{i=1}^{N} V(I_i, \mathbf{w}) - \mathbb{E}_{I\sim\mathcal{D}}[V(I, \mathbf{w})] \right| = \varepsilon_N$$

$$\implies \left| \frac{1}{N} \sum_{i=1}^{N} V(I_i^{\text{train}}, \mathbf{w}) - \frac{1}{N} \sum_{j=1}^{N} V(I_j^{\text{test}}, \mathbf{w}) \right| \leq 2\varepsilon_N, \quad \forall \mathbf{w} \in \mathcal{W}$$

Let $\hat{\mathbf{w}} \in \arg\min_{\mathbf{w}\in\mathcal{W}} \frac{1}{N} \sum_{i=1}^{N} V(I_i^{\text{train}}, \mathbf{w})$:

$$\implies \mathbb{E}_{I\sim\mathcal{D}}[V(I, \hat{\mathbf{w}})] - \min_{\mathbf{w}\in\mathcal{W}} \mathbb{E}_{I\sim\mathcal{D}}[V(I, \mathbf{w})] \leq 2\varepsilon_N,$$

Derivation of Generalization Bounds

Piecewise structure

Definition

A function class $\mathcal{G}$ has a (Γ, γ, β)-structure if $\forall g_1, ..., g_N \in \mathcal{G}$:

- Domain can be decomposed into $\leq N^\gamma \Gamma$ regions.
- In each region: every g_i is a polynomial of degree $\leq \beta$.

Piecewise structure

Definition

A function class $\mathcal{G}$ has a (Γ, γ, β)-structure if $\forall g_1, \ldots, g_N \in \mathcal{G}$:

- Domain can be decomposed into $\leq N^\gamma \Gamma$ regions.
- In each region: every g_i is a polynomial of degree $\leq \beta$.

- Linear: $\left\{g(\mathbf{x}) = \boldsymbol{\alpha}^\top \mathbf{x} + \beta \mid (\boldsymbol{\alpha}, \beta) \in \mathbb{R}^{d+1}\right\}$
  $\implies$ has a $(1, 0, 1)$-structure.

Piecewise structure

Definition

A function class $\mathcal{G}$ has a (Γ, γ, β)-structure if $\forall g_1, ..., g_N \in \mathcal{G}$:

- Domain can be decomposed into $\leq N^\gamma \Gamma$ regions.
- In each region: every g_i is a polynomial of degree $\leq \beta$.

- Linear: $\left\{g(\mathbf{x}) = \boldsymbol{\alpha}^\top \mathbf{x} + \beta \mid (\boldsymbol{\alpha}, \beta) \in \mathbb{R}^{d+1}\right\}$
  $\implies$ has a $(1, 0, 1)$-structure.
- Piecewise linear: $\left\{g(\mathbf{x}) = \max\{0, \boldsymbol{\alpha}^\top \mathbf{x} + \beta\} \mid (\boldsymbol{\alpha}, \beta) \in \mathbb{R}^{d+1}\right\}$
  $\implies$ has a $(1, d, 1)$-structure.

Piecewise structure

Definition

A function class $\mathcal{G}$ has a (Γ, γ, β)-structure if $\forall g_1, ..., g_N \in \mathcal{G}$:

- Domain can be decomposed into $\leq N^\gamma \Gamma$ regions.
- In each region: every g_i is a polynomial of degree $\leq \beta$.

- Linear: $\left\{g(\mathbf{x}) = \boldsymbol{\alpha}^\top \mathbf{x} + \beta \mid (\boldsymbol{\alpha}, \beta) \in \mathbb{R}^{d+1}\right\}$
  $\implies$ has a $(1, 0, 1)$-structure.
- Piecewise linear: $\left\{g(\mathbf{x}) = \max\{0, \boldsymbol{\alpha}^\top \mathbf{x} + \beta\} \mid (\boldsymbol{\alpha}, \beta) \in \mathbb{R}^{d+1}\right\}$
  $\implies$ has a $(1, d, 1)$-structure.

Lemma

ReLU MLP: $\left\{g(\mathbf{w}) = \mathrm{MLP}(\mathbf{x}, \mathbf{w}) \mid \mathbf{x} \in \mathbb{R}^\ell\right\}$

$$\implies \text{has a } \left(2^L \left(\frac{4eU}{W}\right)^{LW}, LW, L\right)\text{-structure.}$$

Generalization Bound via Dual Piecewise Structure

Consider $V : \mathcal{I} \times \mathcal{W} \to \mathbb{R}$.

- Primal class: $\mathcal{V} = \{V(\cdot, \mathbf{w}) : \mathcal{I} \to \mathbb{R} \mid \mathbf{w} \in \mathcal{W}\}$
- Dual class: $\mathcal{V}^* = \{V(I, \cdot) : \mathcal{W} \to \mathbb{R} \mid I \in \mathcal{I}\}$

Generalization Bound via Dual Piecewise Structure

Consider $V : \mathcal{I} \times \mathcal{W} \to \mathbb{R}$.

- Primal class: $\mathcal{V} = \{V(\cdot, \mathbf{w}) : \mathcal{I} \to \mathbb{R} \mid \mathbf{w} \in \mathcal{W}\}$
- Dual class: $\mathcal{V}^* = \{V(I, \cdot) : \mathcal{W} \to \mathbb{R} \mid I \in \mathcal{I}\}$

Lemma

If $\mathcal{V}^*$ has a (Γ, γ, β)-structure and $V(I, \mathbf{w}) \leq H$, then

$$\sup_{\mathbf{w} \in \mathcal{W}} \left| \frac{1}{N} \sum_{i=1}^{N} V(I_i, \mathbf{w}) - \mathbb{E}_{I \sim \mathcal{D}}[V(I, \mathbf{w})] \right| = \mathcal{O}\left(H \sqrt{\frac{\gamma \log \gamma + W \log \beta + \log \Gamma + \log(1/\delta)}{N}} \right)$$

Generalization Bound via Dual Piecewise Structure

Lemma

If $\mathcal{V}^*$ has a (Γ, γ, β)-structure and $V(I, \mathbf{w}) \leq H$, then

$$\sup_{\mathbf{w} \in \mathcal{W}} \left| \frac{1}{N} \sum_{i=1}^{N} V(I_i, \mathbf{w}) - \mathbb{E}_{I \sim \mathcal{D}}[V(I, \mathbf{w})] \right| = \mathcal{O}\left(H \sqrt{\frac{\gamma \log \gamma + W \log \beta + \log \Gamma + \log(1/\delta)}{N}} \right)$$

- Recall at each state s, we select action $a^* \in \arg\max_{a \in \mathcal{A}^s} f(s, a, \mathbf{w})$.
- Assume $|\mathcal{A}^s| \leq \rho$ for all s.
- Assume that we will terminate after M steps.

Generalization Bound via Dual Piecewise Structure

Lemma

If $\mathcal{V}^*$ has a (Γ, γ, β)-structure and $V(I, \mathbf{w}) \leq H$, then

$$\sup_{\mathbf{w} \in \mathcal{W}} \left| \frac{1}{N} \sum_{i=1}^{N} V(I_i, \mathbf{w}) - \mathbb{E}_{I \sim \mathcal{D}}[V(I, \mathbf{w})] \right| = \mathcal{O}\left(H \sqrt{\frac{\gamma \log \gamma + W \log \beta + \log \Gamma + \log(1/\delta)}{N}} \right)$$

- Recall at each state s, we select action $a^* \in \arg\max_{a \in \mathcal{A}^s} f(s, a, \mathbf{w})$.
- Assume $|\mathcal{A}^s| \leq \rho$ for all s.
- Assume that we will terminate after M steps.

Theorem

If $\{f(\mathbf{x}, \cdot) : \mathcal{W} \to \mathbb{R} \mid \mathbf{x} \in \mathbb{R}^\ell\}$ has a (Γ, γ, β)-structure, then $\mathcal{V}^*$ has a $\left(2\rho^{(\gamma+W)(M+1)} \Gamma \left(e \frac{\rho^2 \beta}{W} \right)^W, \gamma + W, 0 \right)$-structure.

Generalization Bound via Dual Piecewise Structure

Theorem

If $\{f(\mathbf{x}, \cdot) : \mathcal{W} \to \mathbb{R} \mid \mathbf{x} \in \mathbb{R}^\ell\}$ has a (Γ, γ, β)-structure, then $\mathcal{V}^*$ has a $\left(2\rho^{(\gamma+W)(M+1)}\Gamma\left(e\frac{\rho^2\beta}{W}\right)^W, \gamma + W, 0\right)$-structure.

Recall that

- Linear scoring class ($f(s, a, \mathbf{w}) = \mathbf{w}^\top \phi(s, a)$): $(1, 0, 1)$-structure
- ReLU MLP scoring class: $\left(2^L\left(\frac{4eU}{W}\right)^{LW}, LW, L\right)$-structure

Interpretation of Generalization Bound

Let $\sup_{\mathbf{w}\in\mathcal{W}} \left| \frac{1}{N} \sum_{i=1}^{N} V(I_i, \mathbf{w}) - \mathbb{E}_{I\sim\mathcal{D}}[V(I, \mathbf{w})] \right| = \varepsilon_N.$

Theorem

$$\varepsilon_N = \mathcal{O}\left(H \sqrt{\frac{LW \log(\cdots) + \log(1/\delta)}{N}} \right)$$

Interpretation of Generalization Bound

Let $\sup_{\mathbf{w}\in\mathcal{W}} \left| \frac{1}{N} \sum_{i=1}^{N} V(I_i, \mathbf{w}) - \mathbb{E}_{I\sim\mathcal{D}}[V(I, \mathbf{w})] \right| = \varepsilon_N.$

Theorem

$$\varepsilon_N = \mathcal{O}\left(H \sqrt{\frac{LW \log(\cdots) + \log(1/\delta)}{N}} \right)$$

- **Scaling Law for ReLU Policies:** $N \lesssim LW$.
- L: number of layers; W: number of parameters.
- Scale the number of training samples N linearly with model size LW to maintain generalization performance.

Experimental Validation: Generalization Guarantees

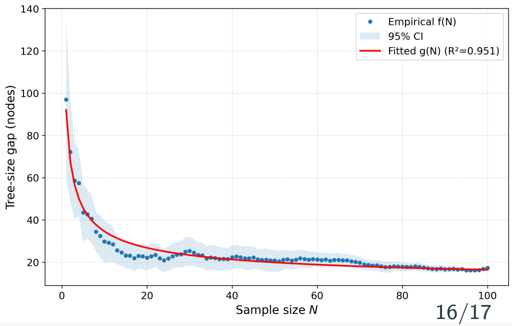
- MLP: 2 hidden ReLU layers (width 10)
- Input: Hand-crafted features
- Label: LP objective improvement

Empirical Generalization Metric:

$$f(N) = \left| \frac{1}{N} \sum_{i=1}^{N} V(I_i^{\text{train}}, \mathbf{w}) - \frac{1}{N} \sum_{j=1}^{N} V(I_j^{\text{test}}, \mathbf{w}) \right|$$

Results:

- Theory suggests $\mathcal{O}(1/\sqrt{N})$ decay
- Fit empirical data to $g(N; a, b) = a/\sqrt{N} + b$, $R^2 > 0.93$



Thank you.