Near-optimal Bootstrapping of Hitting Sets

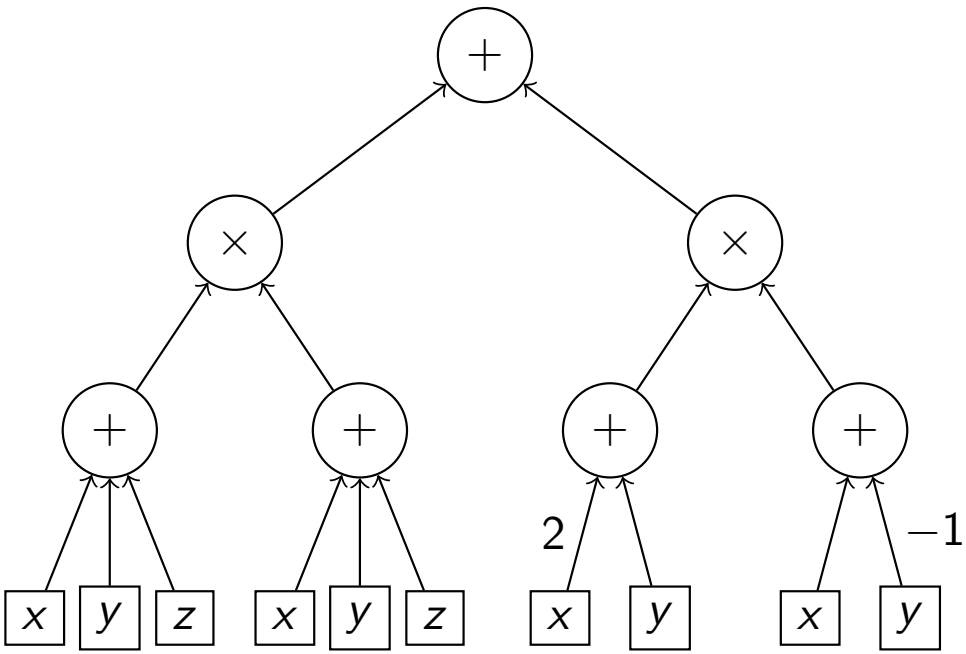
Mrinal Kumar
(Simons Institute)

Ramprasad Saptharishi
(TIFR)

Anamay Tengse[1]
(TIFR)

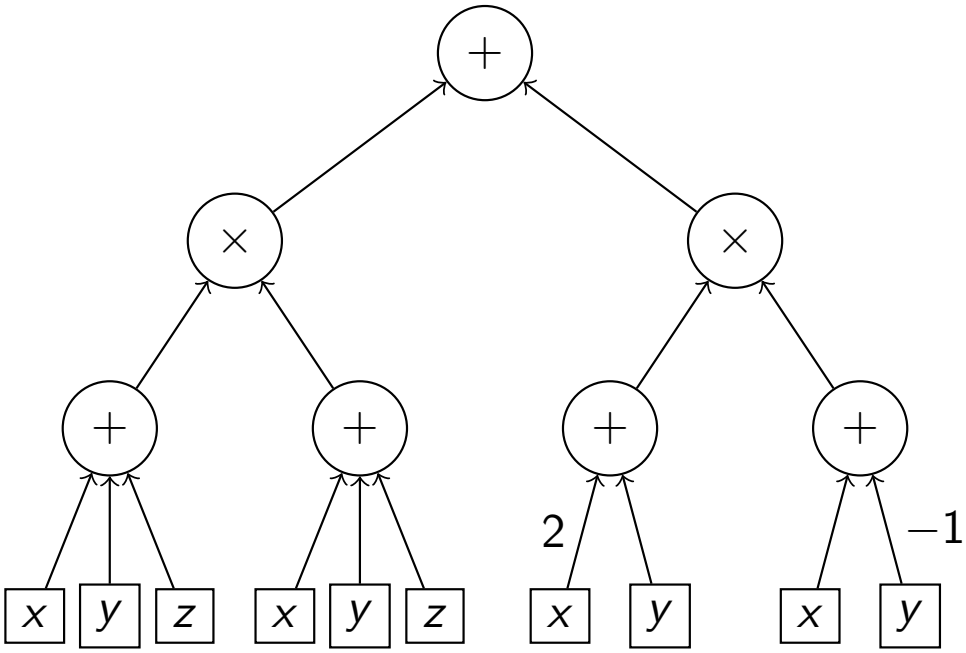
ICTS WACT 2019

Algebraic Models

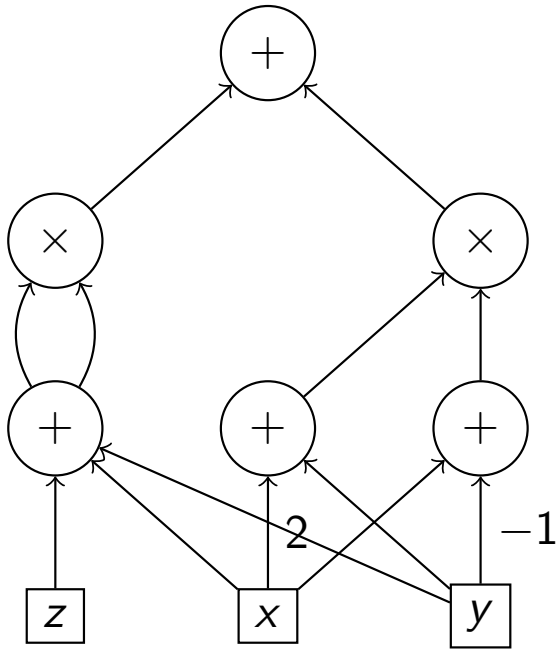


Algebraic Formula

Algebraic Models



Algebraic Formula

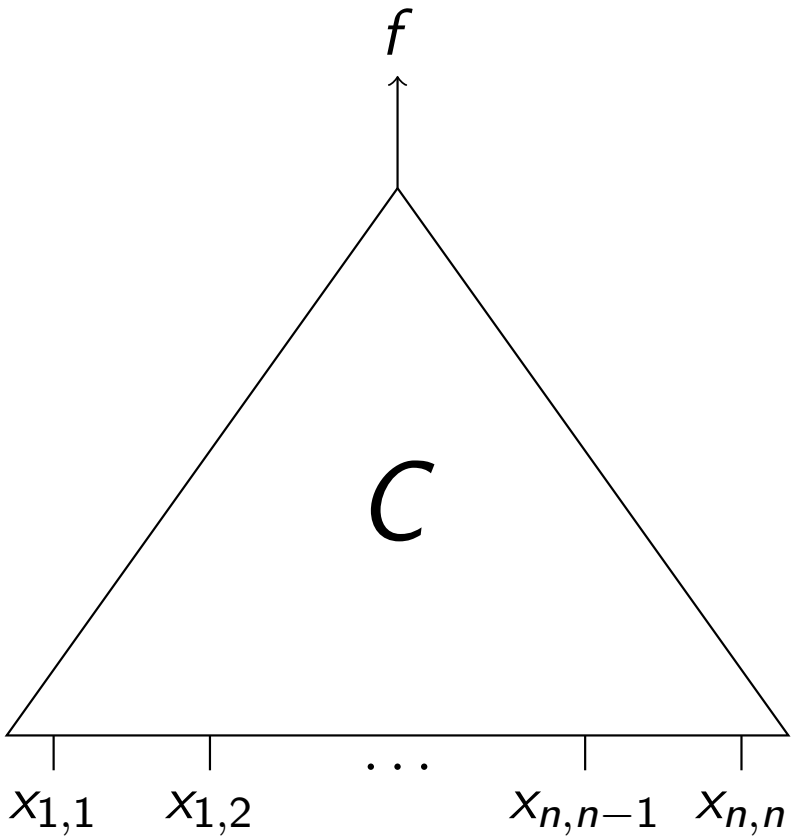


Algebraic Circuit

The Hardness Question

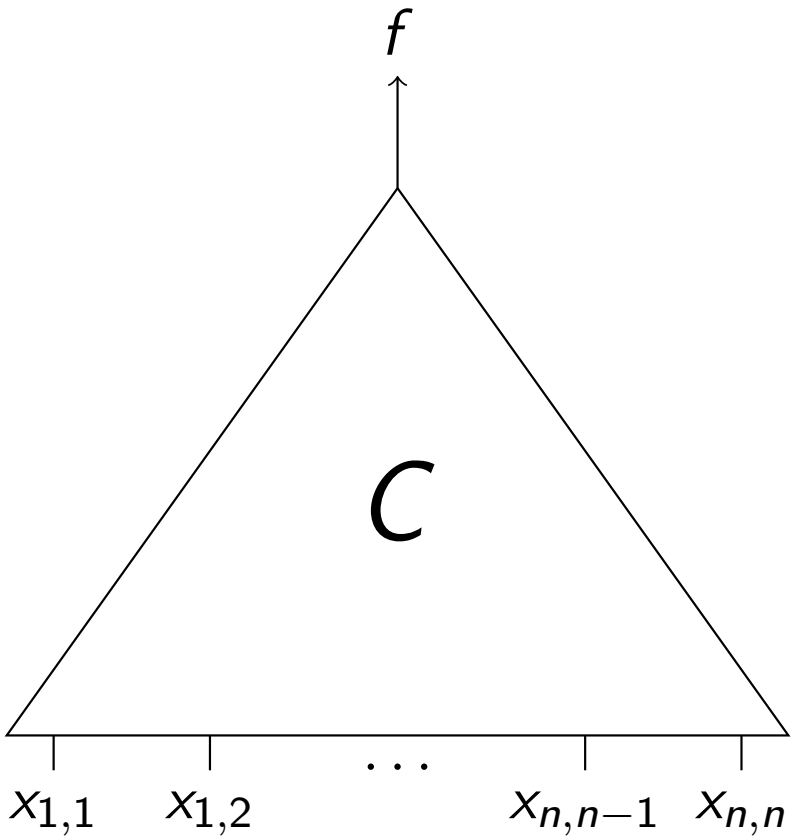
$$f = \mathrm{Det}\left(\begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{bmatrix}\right)$$

The Hardness Question



$$f = \mathrm{Det}\left(\begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{bmatrix}\right)$$

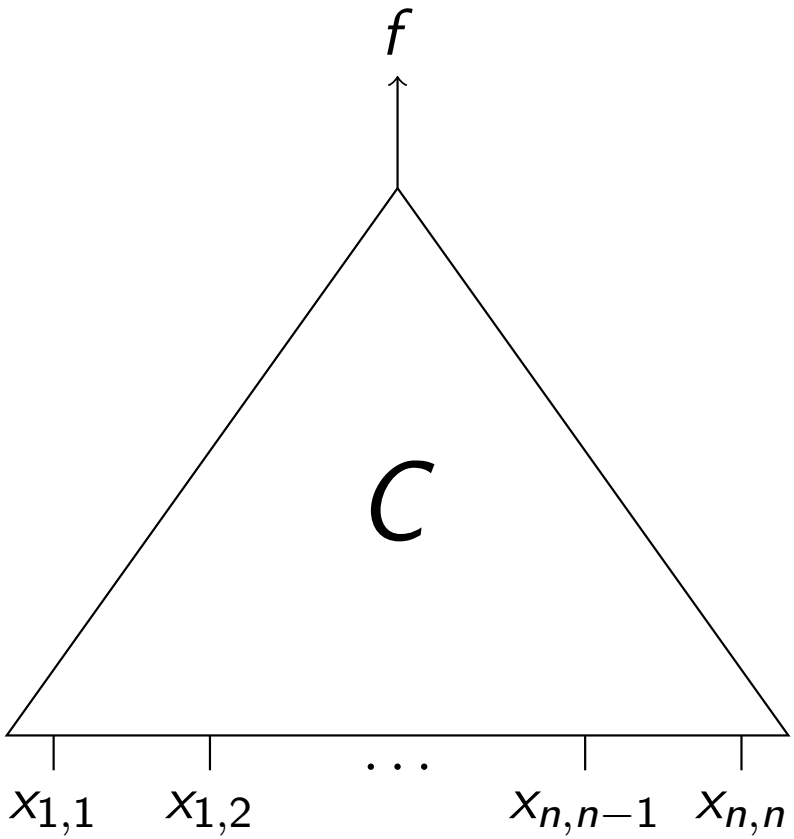
The Hardness Question



$$f = \mathrm{Det}\left(\begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{bmatrix}\right)$$

Does C require size $> n^3$?

The Hardness Question

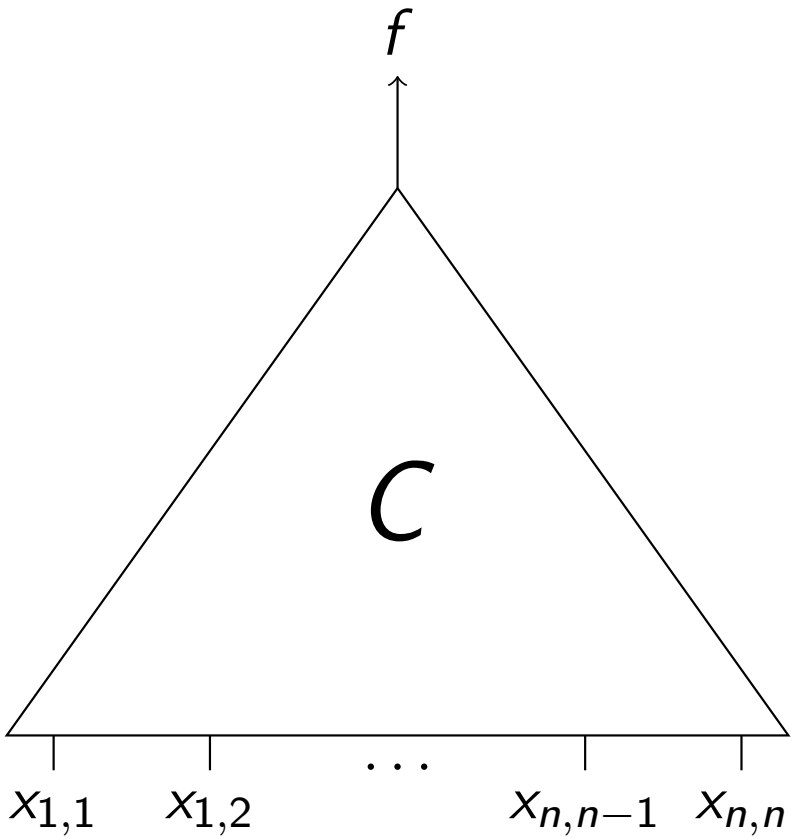


$$f = \mathrm{Det}\left(\begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{bmatrix}\right)$$

Does C require size $> n^3$?

Find an "explicit" n-variate $f(\mathbf{x})$ that requires $n^{\Omega(1)}$ sized circuits?

The Hardness Question

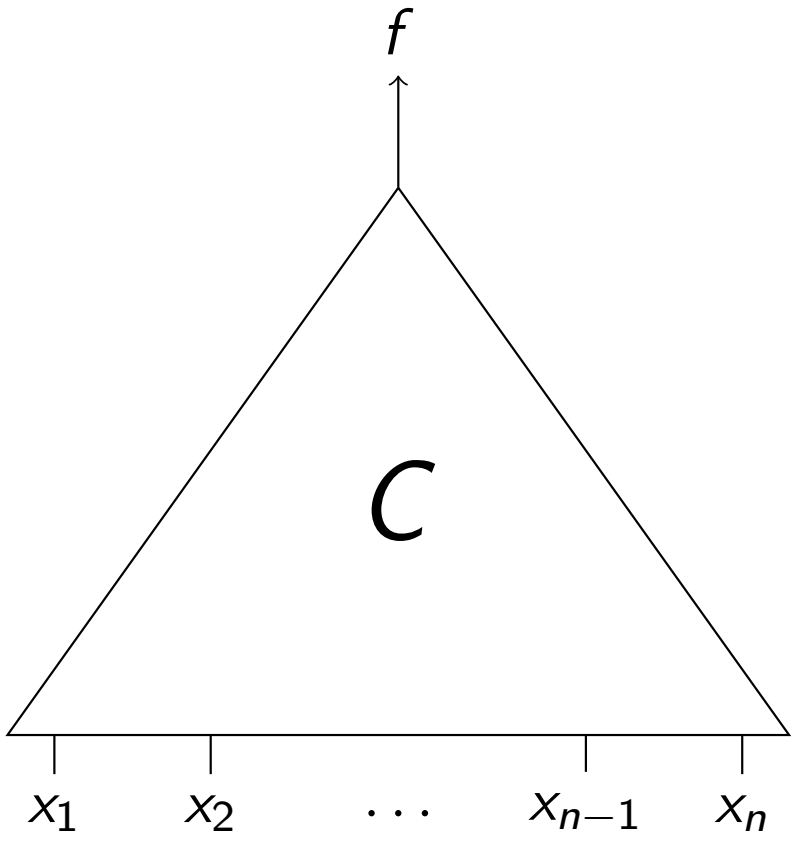


$$f = \mathrm{Det}\left(\begin{bmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{bmatrix}\right)$$

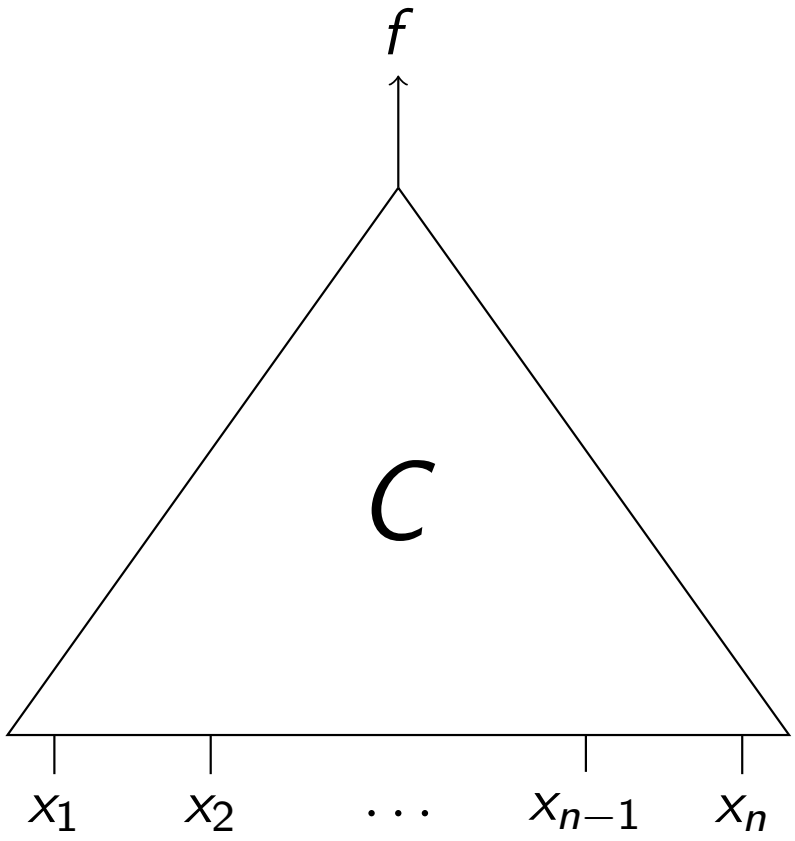
Does C require size $> n^3$?

Find an "explicit" n-variate $f(\mathbf{x})$ that is hard for circuits? **OPEN**

Identity Testing

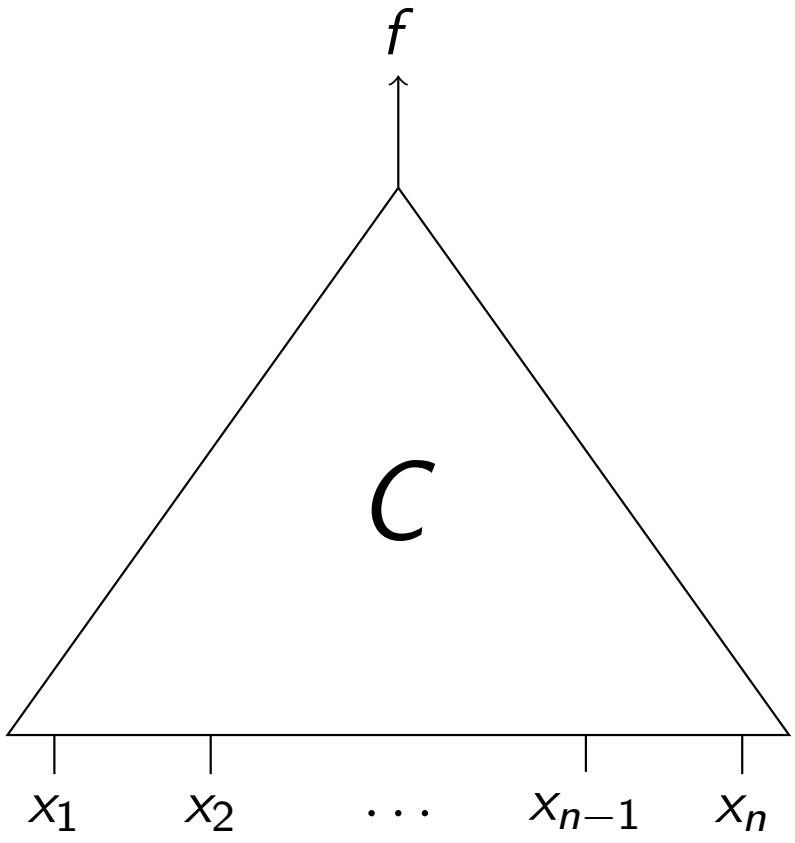


Identity Testing



Can we say something about f?

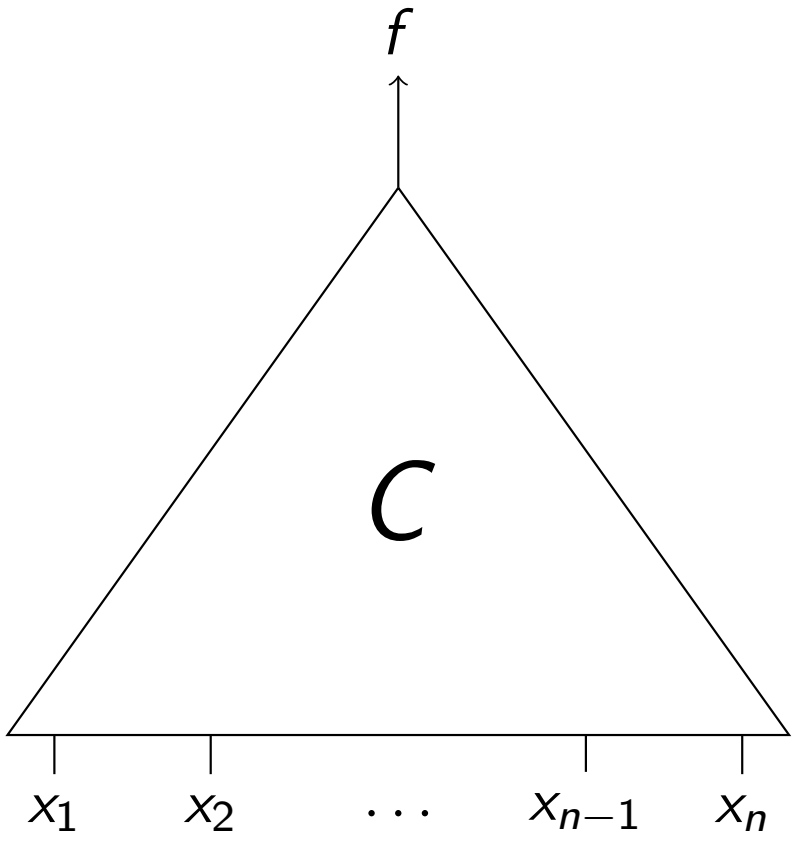
Identity Testing



Can we say something about f?

Is $f = 0$?

Identity Testing

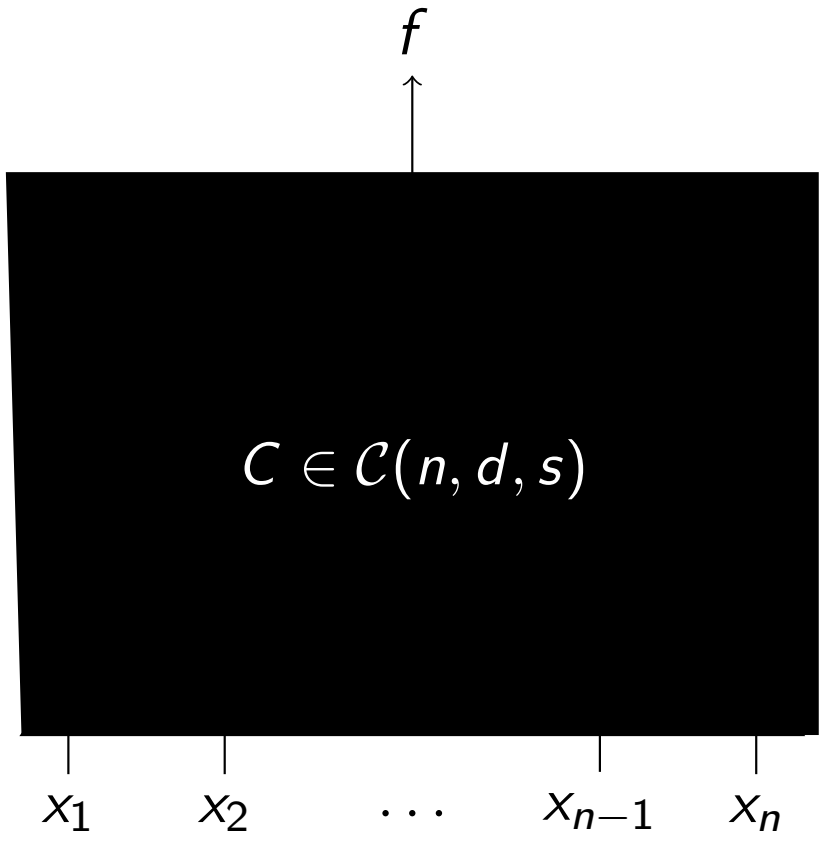


Can we say something about f?

Is $f = 0$?

Whitebox: Does the given circuit compute 0?

Identity Testing

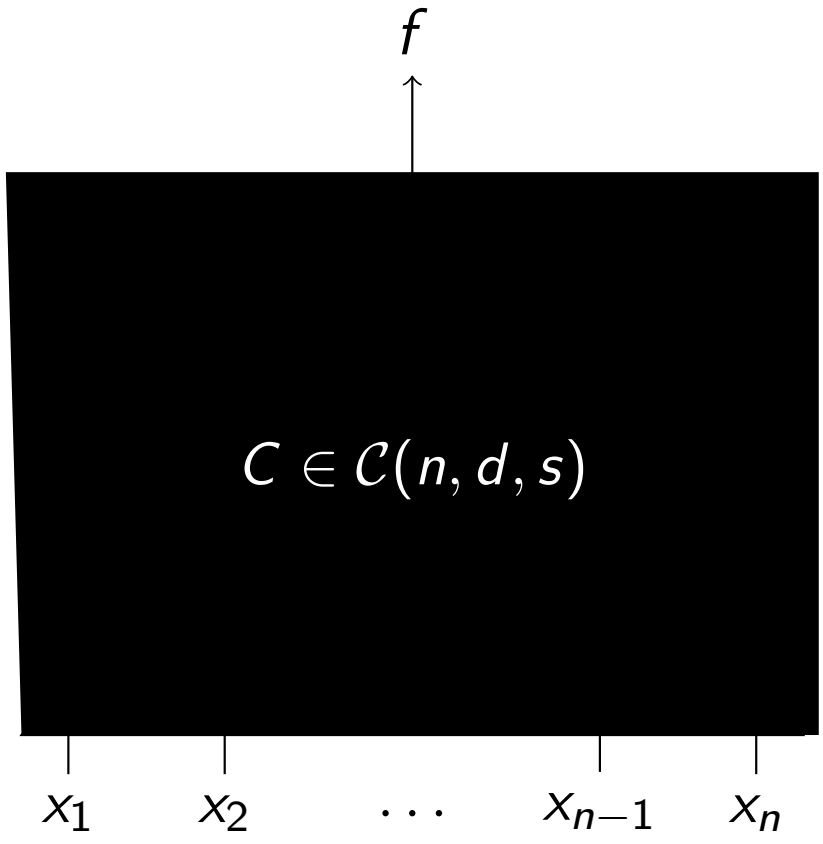


Can we say something about f?

Is $f = 0$?

Blackbox: Evaluate C on some points to tell if $C = 0$.

Identity Testing



Can we say something about f?

Is $f = 0$?

Hitting Set: Find $H_{\mathcal{C}}$ such that $C = 0$ **iff** it is 0 on every $h \in H_{\mathcal{C}}$.

Hitting Sets

Hitting Sets

Counting Argument: There is a non-explicit poly(n, d, s) sized hitting set for the class of all n-variate, degree-d circuits of size s, $\mathcal{C}(n, d, s)$.

Hitting Sets

Counting Argument: There is a non-explicit poly(n, d, s) sized hitting set for the class of all n-variate, degree-d circuits of size s, $\mathcal{C}(n, d, s)$.

Lemma [Ore, DeMillo-Lipton, Schwartz, Zippel]: Any nonzero polynomial of degree d on n variables evaluates to a nonzero value on some point in $[d+1]^n$.

Hitting Sets

Counting Argument: There is a non-explicit poly(n, d, s) sized hitting set for the class of all n-variate, degree-d circuits of size s, $\mathcal{C}(n, d, s)$.

Lemma [Ore, DeMillo-Lipton, Schwartz, Zippel]: Any nonzero polynomial of degree d on n variables evaluates to a nonzero value on some point in $[d+1]^n$.

Corollary: Explicit hitting set of size $d^{O(n)}$ for $\mathcal{C}(n, d, s)$.

Hitting Sets

Counting Argument: There is a non-explicit poly(n, d, s) sized hitting set for the class of all n-variate, degree-d circuits of size s, $\mathcal{C}(n, d, s)$.

Lemma [Ore, DeMillo-Lipton, Schwartz, Zippel]: Any nonzero polynomial of degree d on n variables evaluates to a nonzero value on some point in $[d+1]^n$.

Corollary: Explicit hitting set of size $d^{O(n)}$ for $\mathcal{C}(n, d, s)$.

OPEN: Find an explicit hitting set of size $d^{o(n)}$ for $\mathcal{C}(n, d, s)$.

Improving slightly non-trivial Hitting Sets

Theorem [Agrawal, Ghosh, Saxena 2018]
Suppose for a large constant n and all $s \geq n$, there is an explicit hitting set of size

$$s^{n^{0.49}} \quad \textit{for} \quad \mathcal{C}(n, s, s).$$

Improving slightly non-trivial Hitting Sets

Theorem [Agrawal, Ghosh, Saxena 2018]
Suppose for a large constant n and all $s \geq n$, there is an explicit hitting set of size

$$s^{n^{0.49}} \quad \text{for} \quad \mathcal{C}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \text{for} \quad \mathcal{C}(s, s, s).$$

Improving slightly non-trivial Hitting Sets

Theorem [Agrawal, Ghosh, Saxena 2018]
Suppose for a large constant n and all $s \geq n$, there is an explicit hitting set of size

$$s^{n^{0.49}} \quad \text{for} \quad \mathcal{C}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \text{for} \quad \mathcal{C}(s, s, s).$$

$\text{tiny}(s) = \exp(\exp(O(\log^* s)))$

Improving slightly non-trivial Hitting Sets

Theorem [Agrawal, Ghosh, Saxena 2018]
Suppose for a large constant n and all $s \geq n$, there is an explicit hitting set of size

$$s^{n^{0.49}} \quad \text{for} \quad \mathcal{C}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \text{for} \quad \mathcal{C}(s, s, s).$$

Improving barely non-trivial Hitting Sets

Theorem* [Kumar, Saptharishi, T (arXiv)]
Suppose for a large constant n and all $s \geq n$, there is an explicit hitting set of size

$$s^{n^{0.49}} \quad \text{for} \quad \mathcal{C}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \text{for} \quad \mathcal{C}(s, s, s).$$

Improving barely non-trivial Hitting Sets

Theorem* [Kumar, Saptharishi, T (arXiv)]
Suppose for a constant $n \geq 2$ and all $s \geq n$, there is an explicit hitting set of size

$$s^{n^{0.49}} \quad \text{for} \quad \mathcal{C}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \text{for} \quad \mathcal{C}(s, s, s).$$

Improving barely non-trivial Hitting Sets

Theorem* [Kumar, Saptharishi, T (arXiv)]
Suppose for a constant $n \geq 2$, some $\epsilon > 0$ and all $s \geq n$, there is an explicit hitting set of size

$$s^{n-\epsilon} \quad \text{for} \quad \mathcal{C}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \text{for} \quad \mathcal{C}(s, s, s).$$

Improving barely non-trivial Hitting Sets

Theorem* [Kumar, Saptharishi, T (arXiv)]
Suppose for a constant $n \geq 2$, some $\epsilon > 0$ and all $s \geq n$, there is an explicit hitting set of size

$$s^{n-\epsilon} \quad \text{for} \quad \text{Formula}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

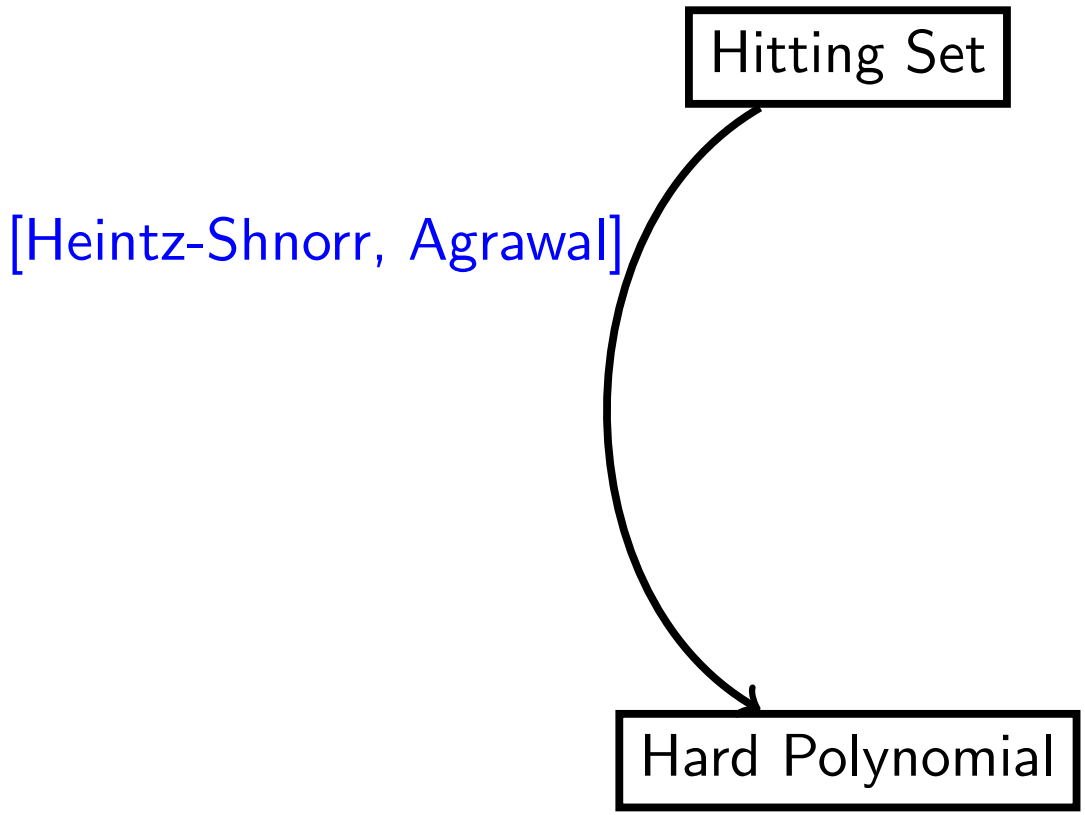
$$s^{\text{tiny}(s)} \quad \text{for} \quad \text{Formula}(s, s, s).$$

What is Bootstrapping?

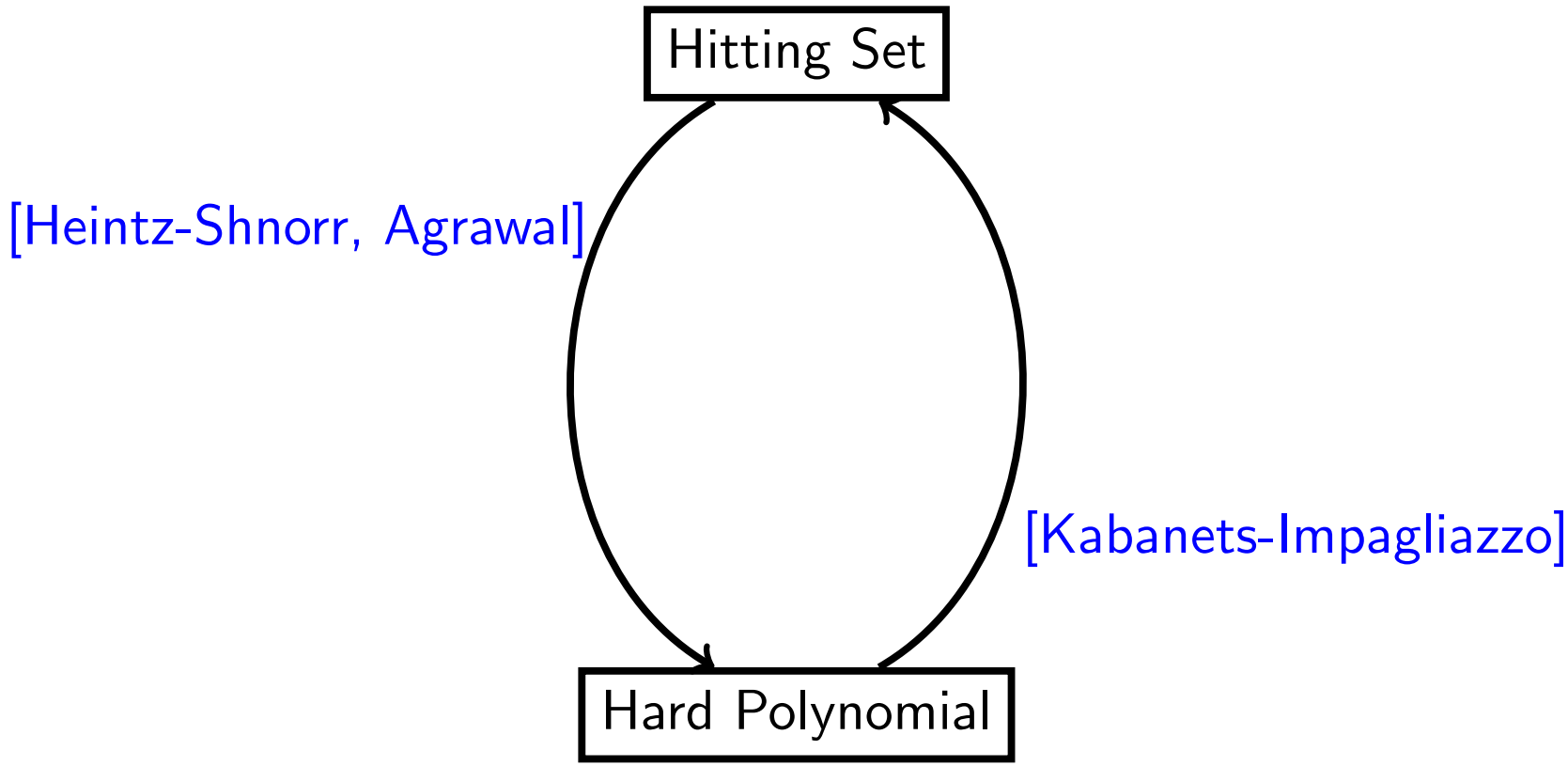
Hitting Set

Hard Polynomial

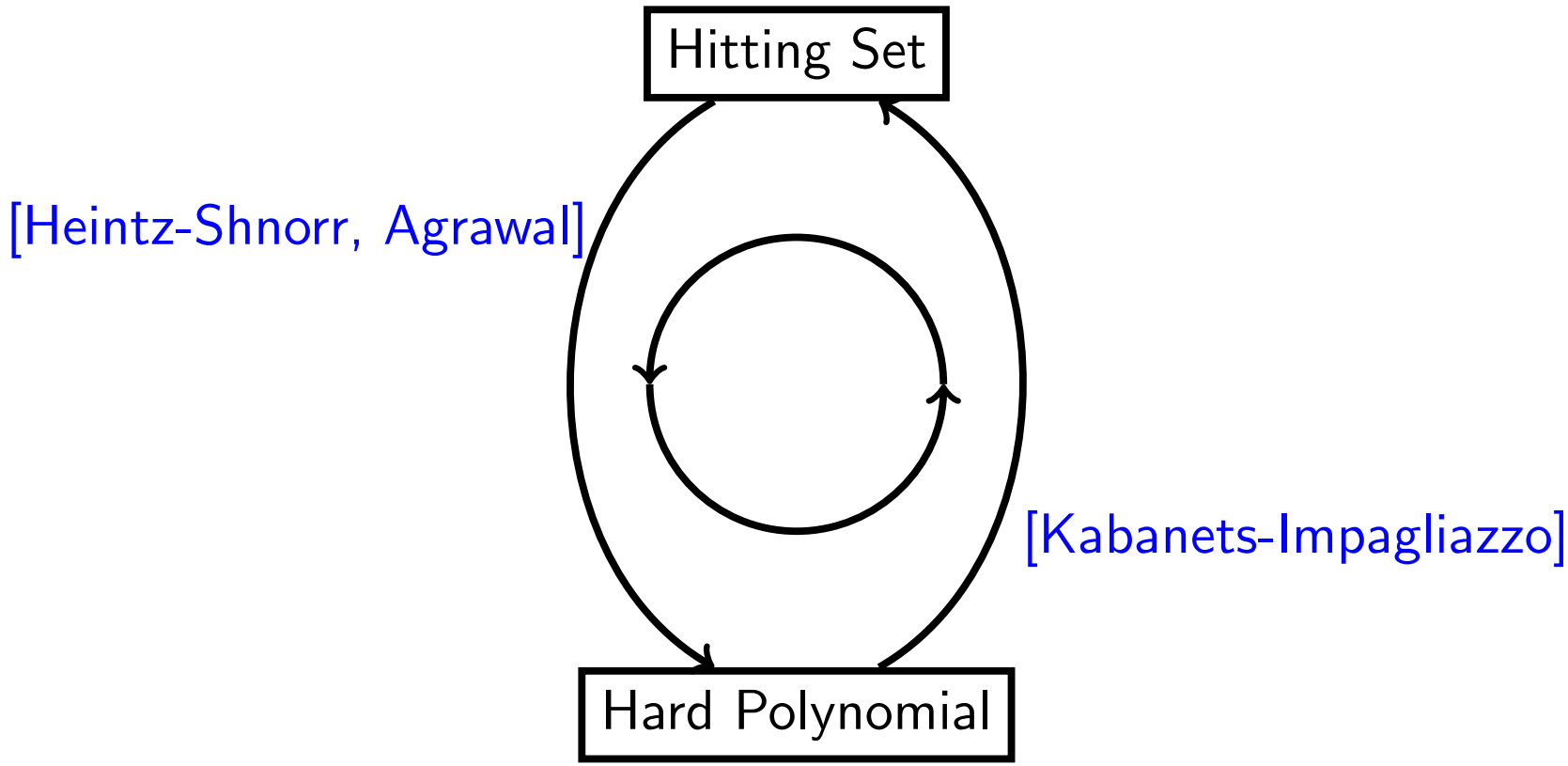
What is Bootstrapping?



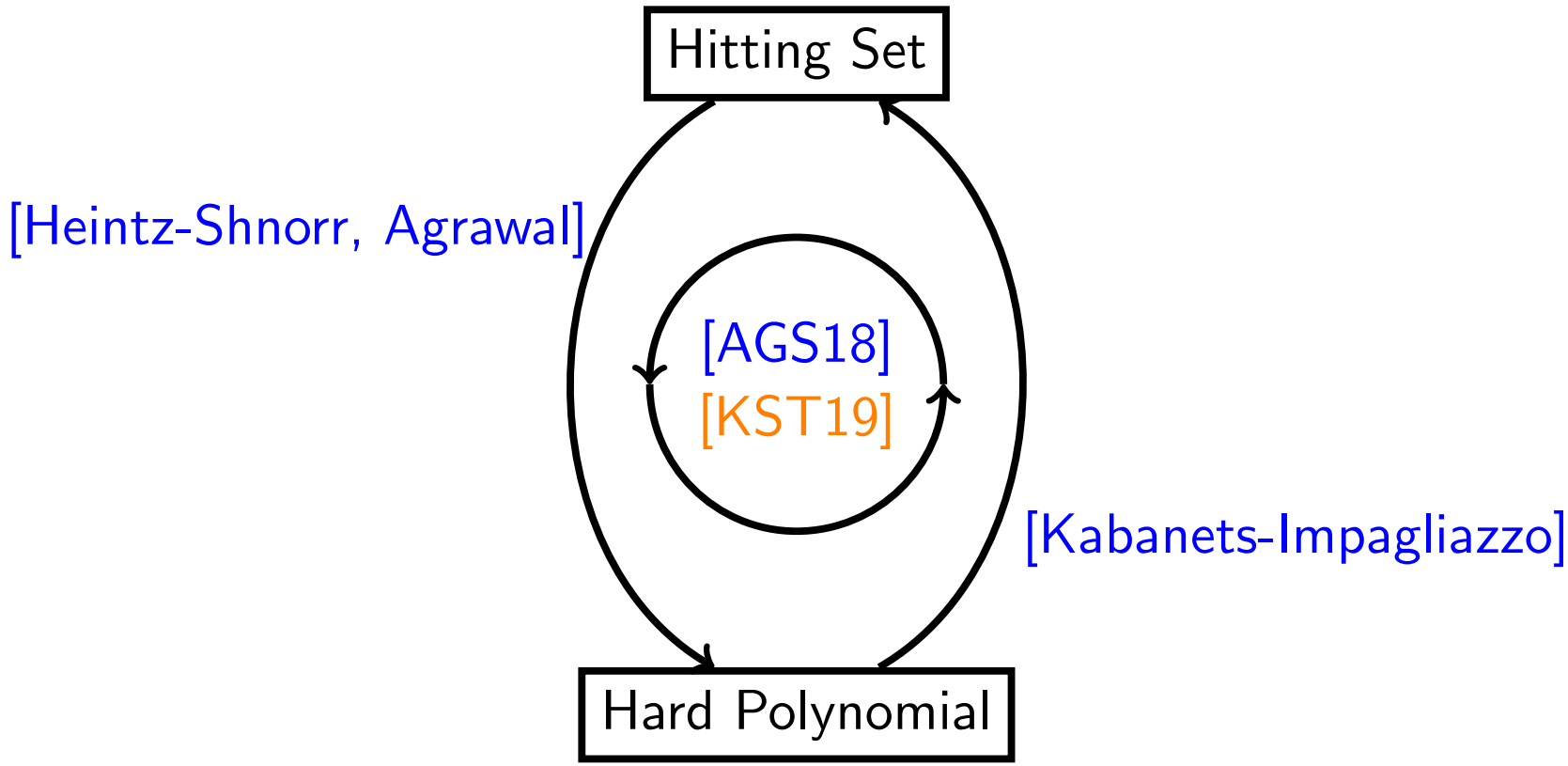
What is Bootstrapping?



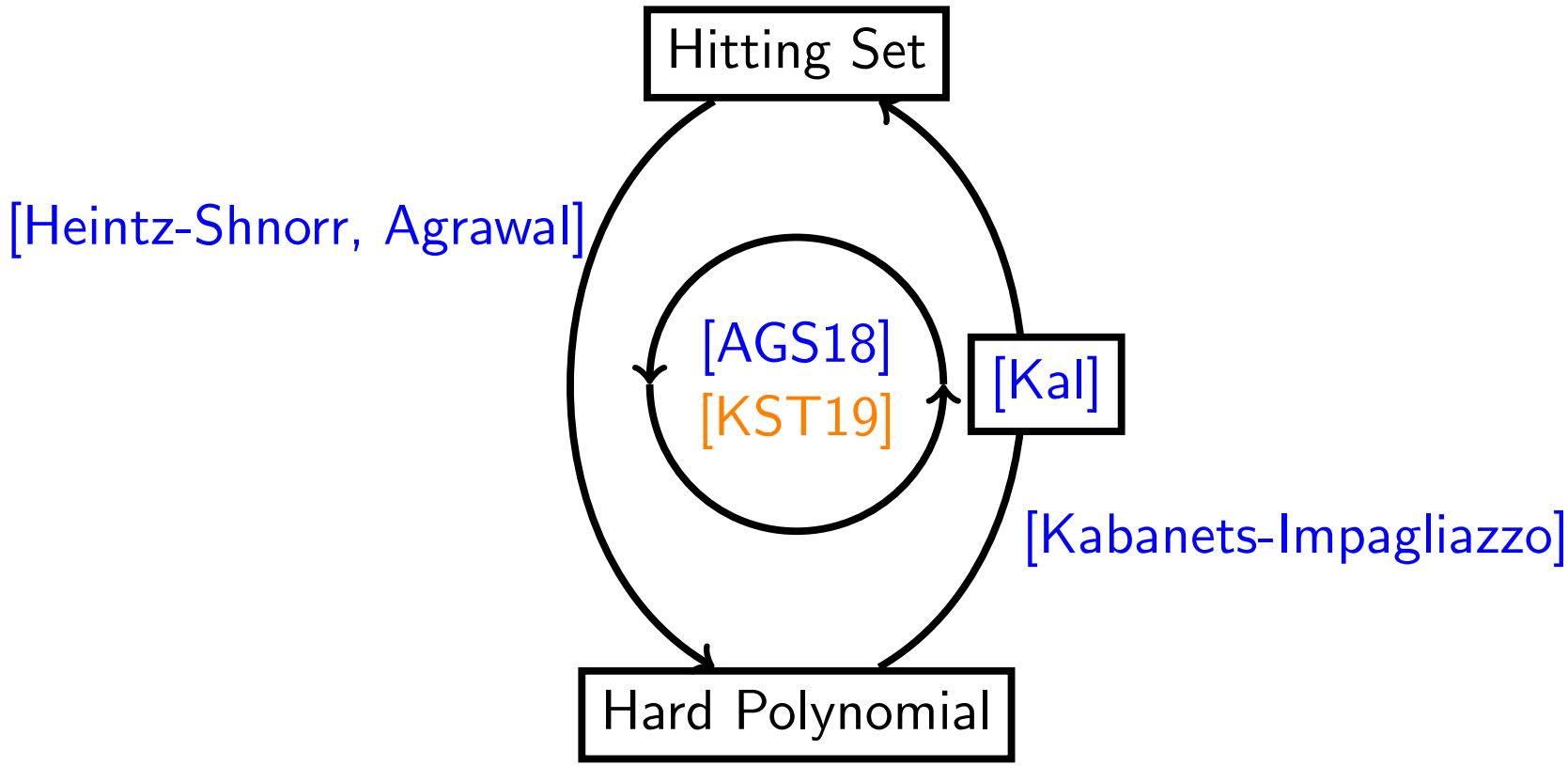
What is Bootstrapping?



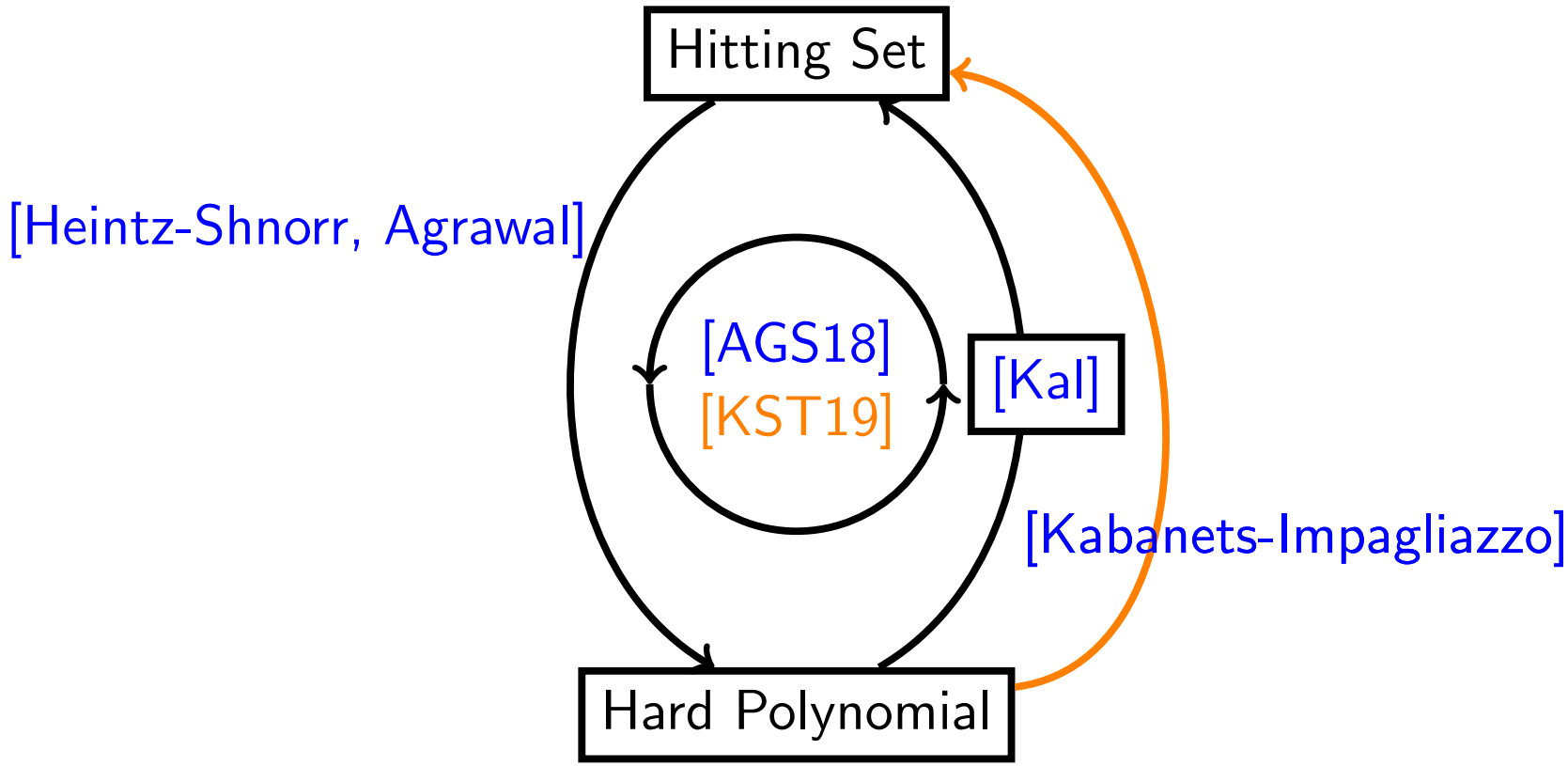
What is Bootstrapping?



What is Bootstrapping?



What is Bootstrapping?



Hardness from Hitting Sets

Theorem [Heitnz-Schnorr, Agrawal] (Informal)

Suppose H *is a hitting set for* $\mathcal{C}(n, d, s)$.

Hardness from Hitting Sets

Theorem [Heitnz-Schnorr, Agrawal] (Informal)

Suppose H *is a hitting set for* $\mathcal{C}(n, d, s)$.

Then for any $k \leq n$ *and* δ *satisfying*

$$\delta^k > |H| \quad \text{and} \quad k\delta \leq d$$

Hardness from Hitting Sets

Theorem [Heitnz-Schnorr, Agrawal] (Informal)

Suppose H *is a hitting set for* $\mathcal{C}(n, d, s)$.

Then for any $k \leq n$ *and* δ *satisfying*

$$\delta^k > |H| \quad \text{and} \quad k\delta \leq d$$

there is a k*-variate polynomial* Q_k *of individual degree* δ*, that is hard for* $\mathcal{C}(n, d, s)$.

Hardness from Hitting Sets

Theorem [Heitnz-Schnorr, Agrawal] (Informal)

Suppose H *is a hitting set for* $\mathcal{C}(n, d, s)$.

Then for any $k \leq n$ *and* δ *satisfying*

$$\delta^k > |H| \quad \text{and} \quad k\delta \leq d$$

there is a k*-variate polynomial* Q_k *of individual degree* δ*, that is hard for* $\mathcal{C}(n, d, s)$.

Proof Idea: Use interpolation to get a Q_k that vanishes on H.

Hitting Sets from Hardness

Theorem [Kabanets-Impagliazzo] (Informal)

Suppose Q_k *has individual degree* δ *and requires large circuits.*

Hitting Sets from Hardness

Theorem [Kabanets-Impagliazzo] (Informal)

Suppose Q_k *has individual degree* δ *and requires large circuits.*

Then for any nonzero $P \in \mathcal{C}(m, d, s)$,

$$P\left(Q_k(\mathbf{y}_1), Q_k(\mathbf{y}_2), \ldots, Q_k(\mathbf{y}_m)\right) \quad \textit{is nonzero}$$

when $\mathbf{y}_1, \ldots, \mathbf{y}_m$ *are nearly disjoint.*

Hitting Sets from Hardness

Theorem [Kabanets-Impagliazzo] (Informal)

Suppose Q_k *has individual degree* δ *and requires large circuits.*

Then for any nonzero $P \in \mathcal{C}(m, d, s)$,

$$P\left(Q_k(\mathbf{y}_1), Q_k(\mathbf{y}_2), \ldots, Q_k(\mathbf{y}_m)\right) \quad \textit{is nonzero}$$

when $\mathbf{y}_1, \ldots, \mathbf{y}_m \subseteq \{y_1, \ldots, y_{k^2}\}$.

Hitting Sets from Hardness

Theorem [Kabanets-Impagliazzo] (Informal)

Suppose Q_k *has individual degree* δ *and requires large circuits.*

Then for any nonzero $P \in \mathcal{C}(m, d, s)$ *with* $m \sim \exp(\sqrt{k})$,

$$P\left(Q_k(\mathbf{y}_1), Q_k(\mathbf{y}_2), \ldots, Q_k(\mathbf{y}_m)\right) \quad \textit{is nonzero}$$

when $\mathbf{y}_1, \ldots, \mathbf{y}_m \subseteq \{y_1, \ldots, y_{k^2}\}$.

Hitting Sets from Hardness

Theorem [Kabanets-Impagliazzo] (Informal)

Suppose Q_k *has individual degree* δ *and requires large circuits.*

Then for any nonzero $P \in \mathcal{C}(m, d, s)$ *with* $m \sim \exp(\sqrt{k})$,

$$P\left(Q_k(\mathbf{y}_1), Q_k(\mathbf{y}_2), \ldots, Q_k(\mathbf{y}_m)\right) \quad \textit{is nonzero}$$

when $\mathbf{y}_1, \ldots, \mathbf{y}_m \subseteq \{y_1, \ldots, y_{k^2}\}$.

Outcome: $\mathrm{PIT}(m, d, s)$ reduces to $\mathrm{PIT}(k^2, d', s')$,

for slightly larger d', s' and $k \sim \mathrm{polylog}(m)$.

Template for Bootstrapping

On the board!

Why can we do this repeatedly?

Bootstrapping Procedure:

Why can we do this repeatedly?

Bootstrapping Procedure:

Reduce $\mathrm{PIT}(s, s, s)$ to $\mathrm{PIT}(\log^c(s), s', s')$.

Why can we do this repeatedly?

Bootstrapping Procedure:

Reduce $\mathrm{PIT}(s, s, s)$ to $\mathrm{PIT}(\log^c(s), s', s')$.

Then reduce that to $\mathrm{PIT}(\mathrm{loglog}^c(s), s'', s'')$.

Why can we do this repeatedly?

Bootstrapping Procedure:

Reduce $\mathrm{PIT}(s, s, s)$ to $\mathrm{PIT}(\log^c(s), s', s')$.

Then reduce that to $\mathrm{PIT}(\mathrm{loglog}^c(s), s'', s'')$.

$\vdots$

Reduce to constant variate PIT for size $s^{\mathrm{tiny}(s)}$.

Why can we do this repeatedly?

Bootstrapping Procedure:

Reduce $\mathrm{PIT}(s, s, s)$ to $\mathrm{PIT}(\log^c(s), s', s')$.

Q over $k = \mathrm{polylog}(s)$ variables, $s^{\Omega(1)} = \exp(k)$ hard.

Then reduce that to $\mathrm{PIT}(\mathrm{loglog}^c(s), s'', s'')$.

$\vdots$

Reduce to constant variate PIT for size $s^{\mathrm{tiny}(s)}$.

Why can we do this repeatedly?

Bootstrapping Procedure:

Reduce $\mathrm{PIT}(s, s, s)$ to $\mathrm{PIT}(\log^c(s), s', s')$.

Q over $k = \mathrm{polylog}(s)$ variables, $s^{\Omega(1)} = \exp(k)$ hard.

Then reduce that to $\mathrm{PIT}(\mathrm{loglog}^c(s), s'', s'')$.

Q over $k = \mathrm{loglog}^{c'}(s)$ variables, $s^{\Omega(1)} = \exp(\exp(k))$ hard.

$\vdots$

Reduce to constant variate PIT for size $s^{\mathrm{tiny}(s)}$.

Why can we do this repeatedly?

Bootstrapping Procedure:

Reduce $\mathrm{PIT}(s, s, s)$ to $\mathrm{PIT}(\log^c(s), s', s')$.

Q over $k = \mathrm{polylog}(s)$ variables, $s^{\Omega(1)} = \exp(k)$ hard.

Then reduce that to $\mathrm{PIT}(\mathrm{loglog}^c(s), s'', s'')$.

Q over $k = \mathrm{loglog}^{c'}(s)$ variables, $s^{\Omega(1)} = \exp(\exp(k))$ hard.

$\vdots$

Reduce to constant variate PIT for size $s^{\mathrm{tiny}(s)}$.

Possible due to freedom in individual degree of Q.

Why can we do this repeatedly?

Bootstrapping Procedure:

Reduce $\mathrm{PIT}(s, s, s)$ to $\mathrm{PIT}(\log^c(s), s', s')$.

Q over $k = \mathrm{polylog}(s)$ variables, $s^{\Omega(1)} = \exp(k)$ hard.

Then reduce that to $\mathrm{PIT}(\mathrm{loglog}^c(s), s'', s'')$.

Q over $k = \mathrm{loglog}^{c'}(s)$ variables, $s^{\Omega(1)} = \exp(\exp(k))$ hard.

$\vdots$

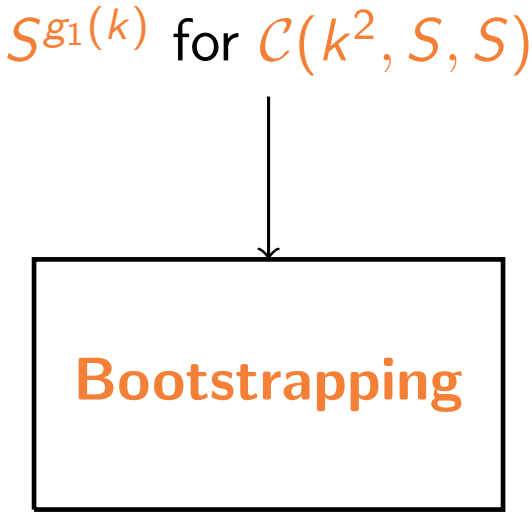
Reduce to constant variate PIT for size $s^{\mathrm{tiny}(s)}$.

Possible due to freedom in individual degree of Q.

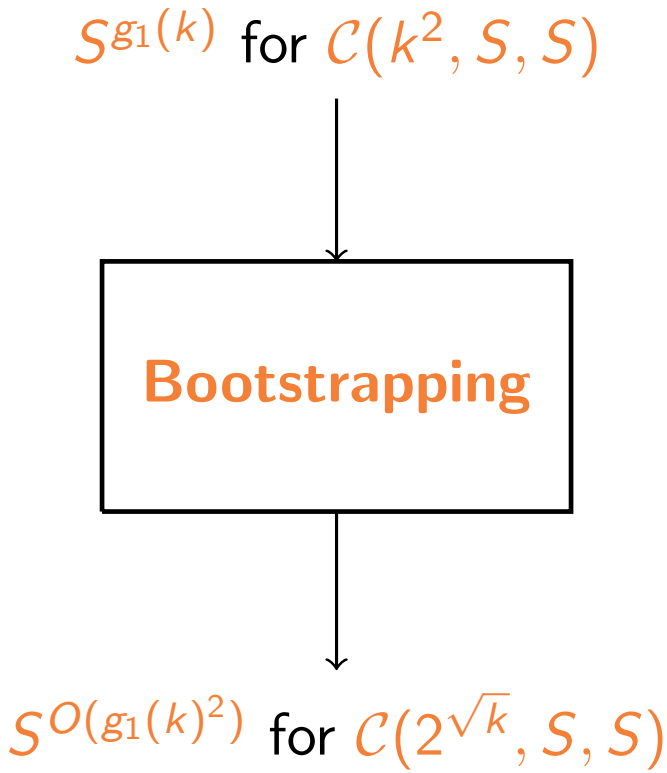
Unlike the boolean case, nothing stops us.

High Level Overview

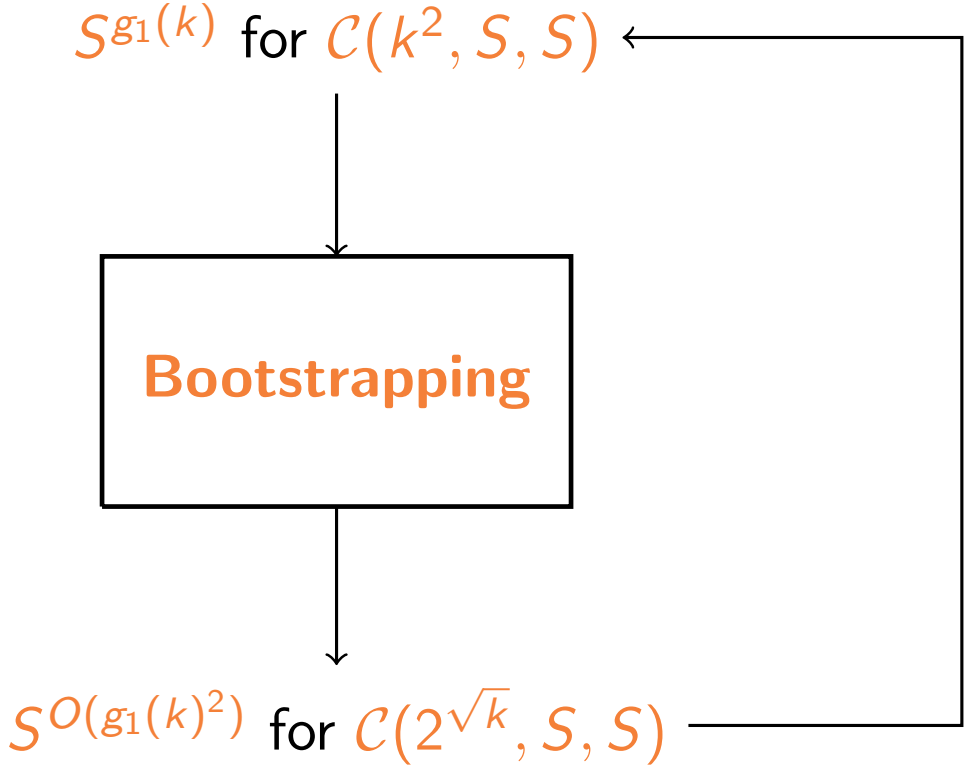
High Level Overview



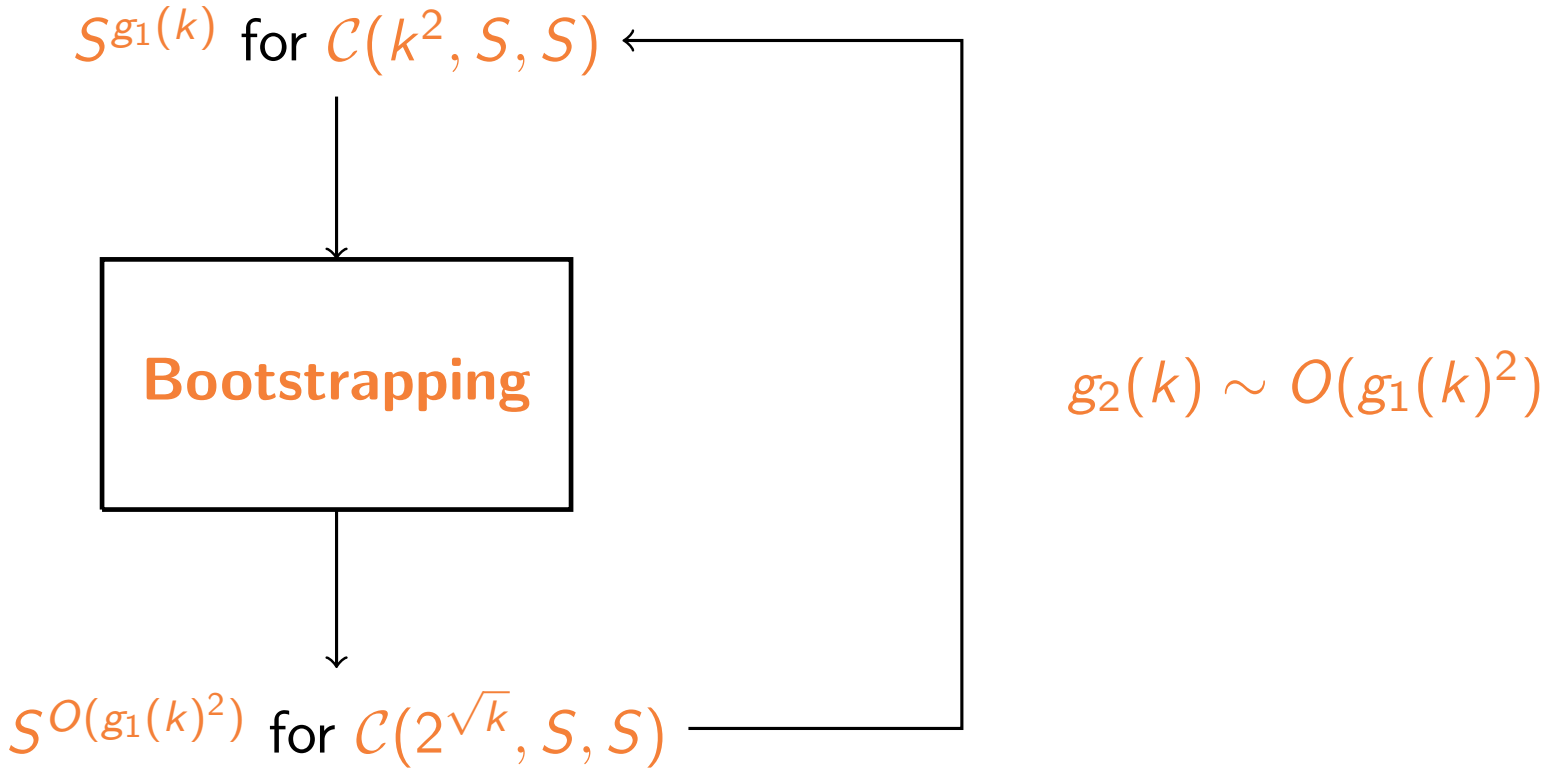
High Level Overview



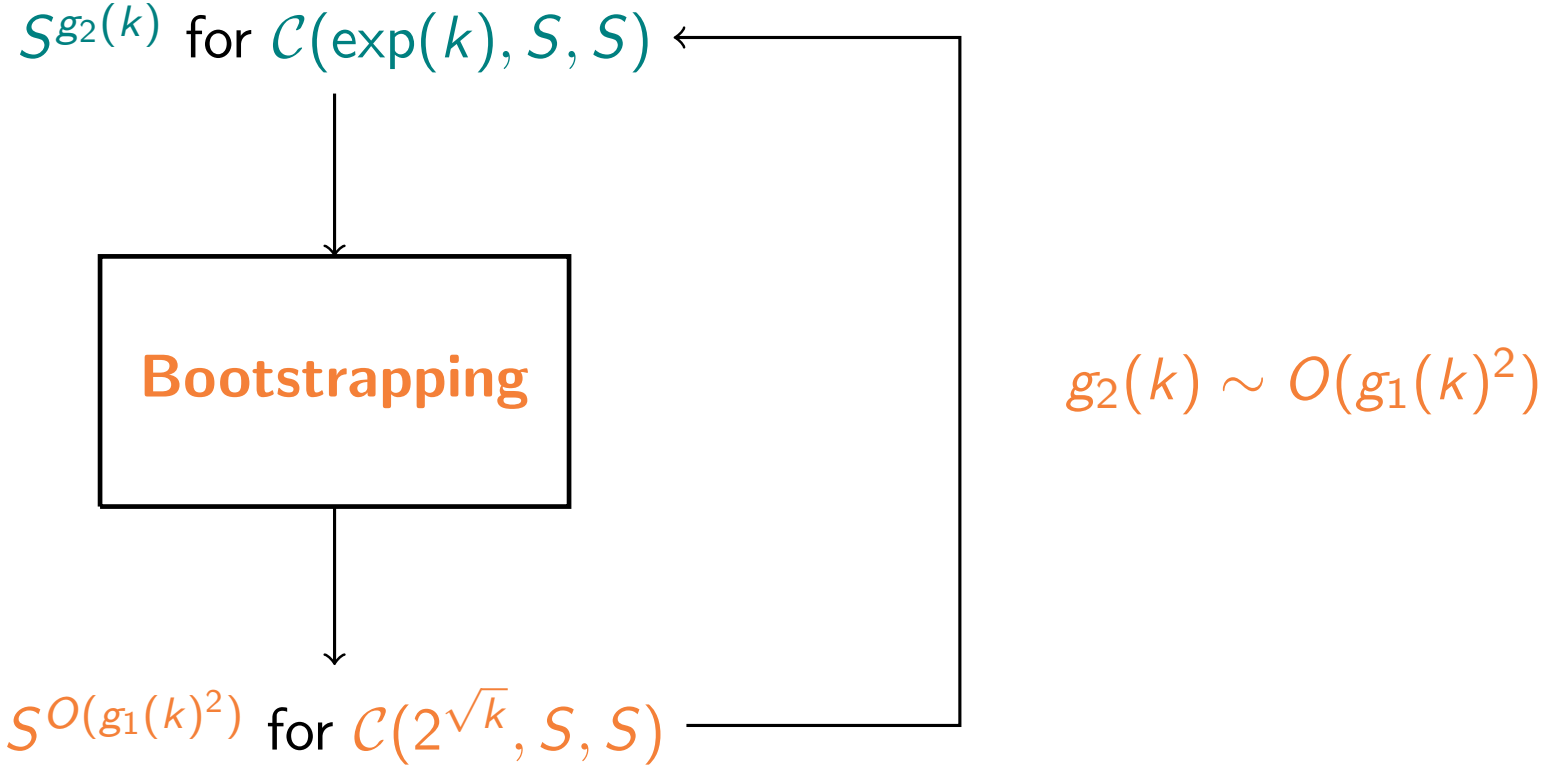
High Level Overview



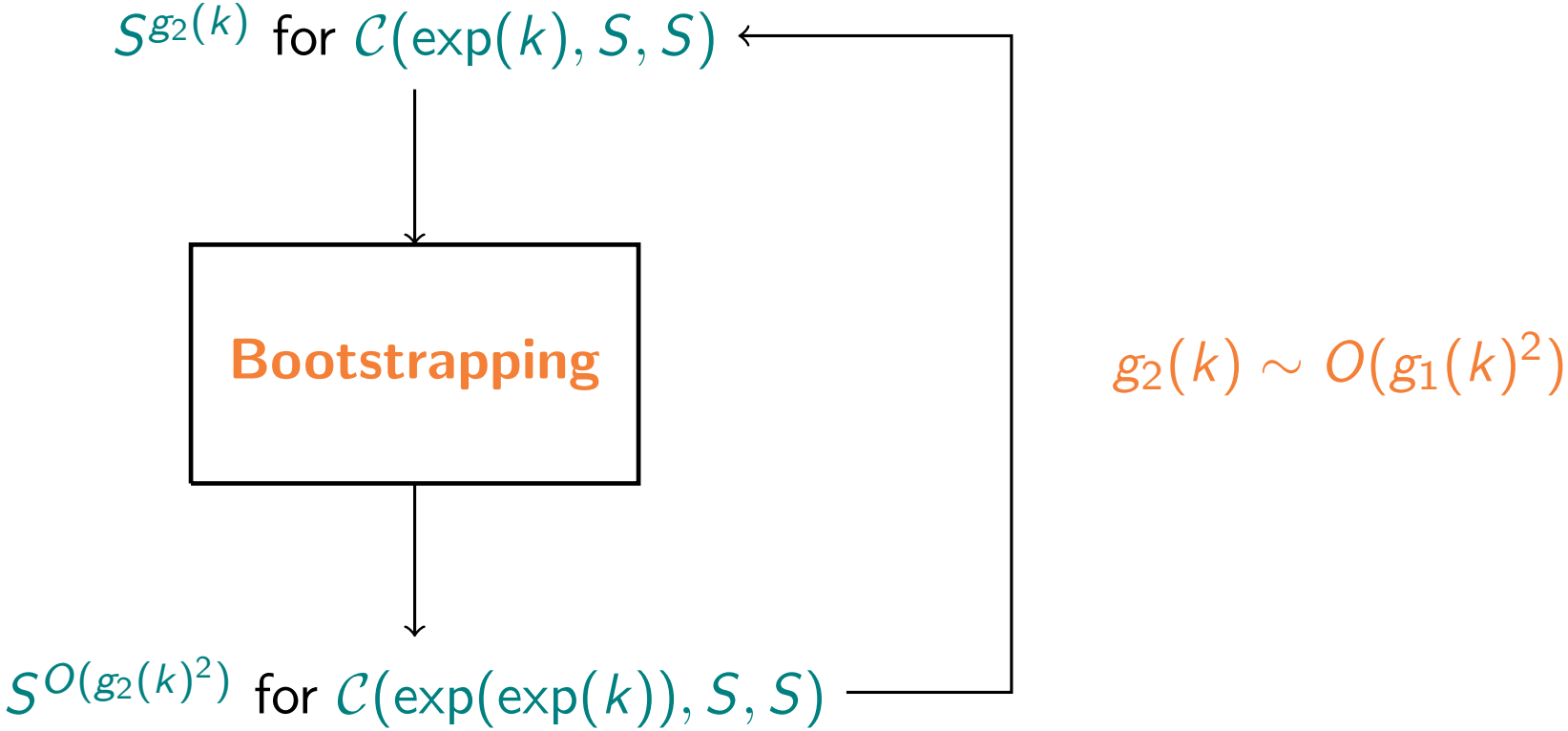
High Level Overview



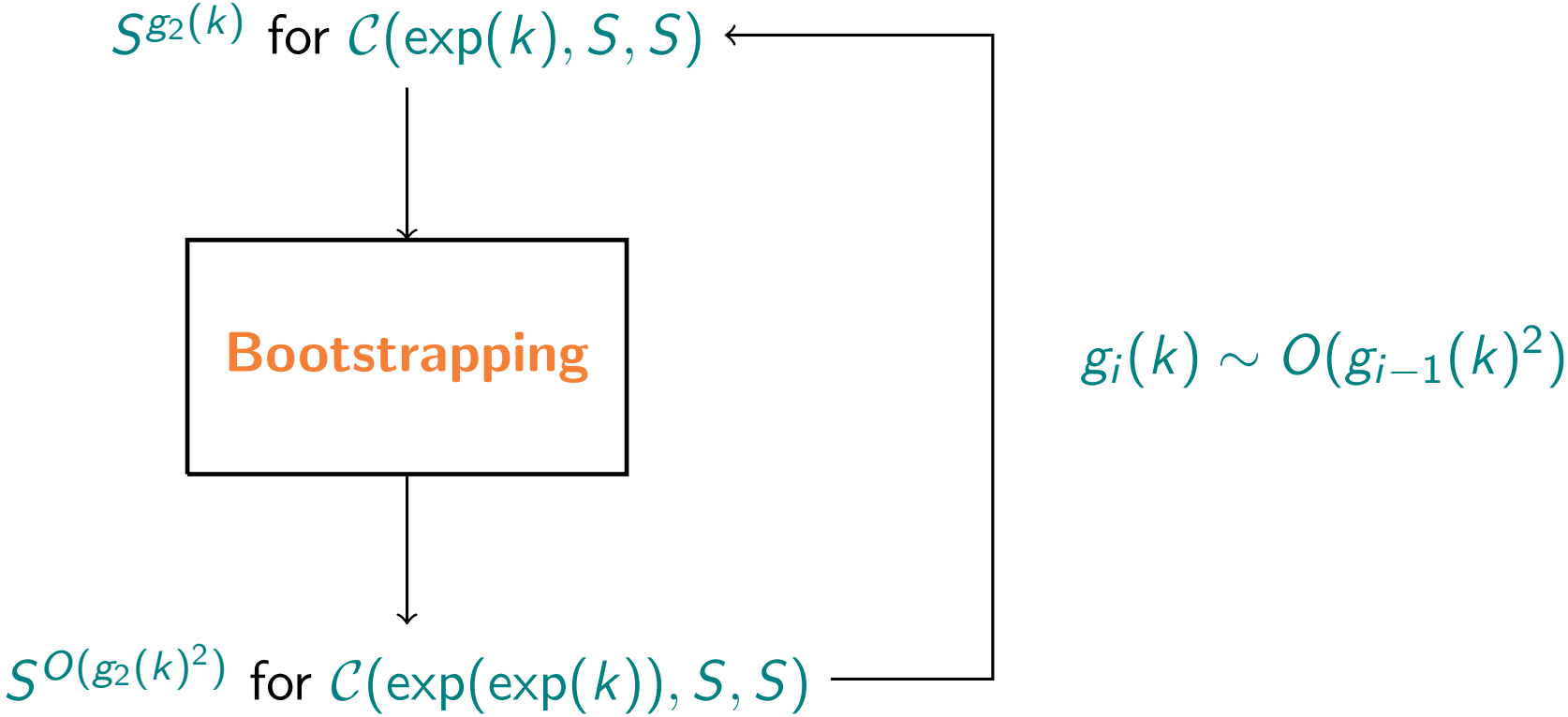
High Level Overview



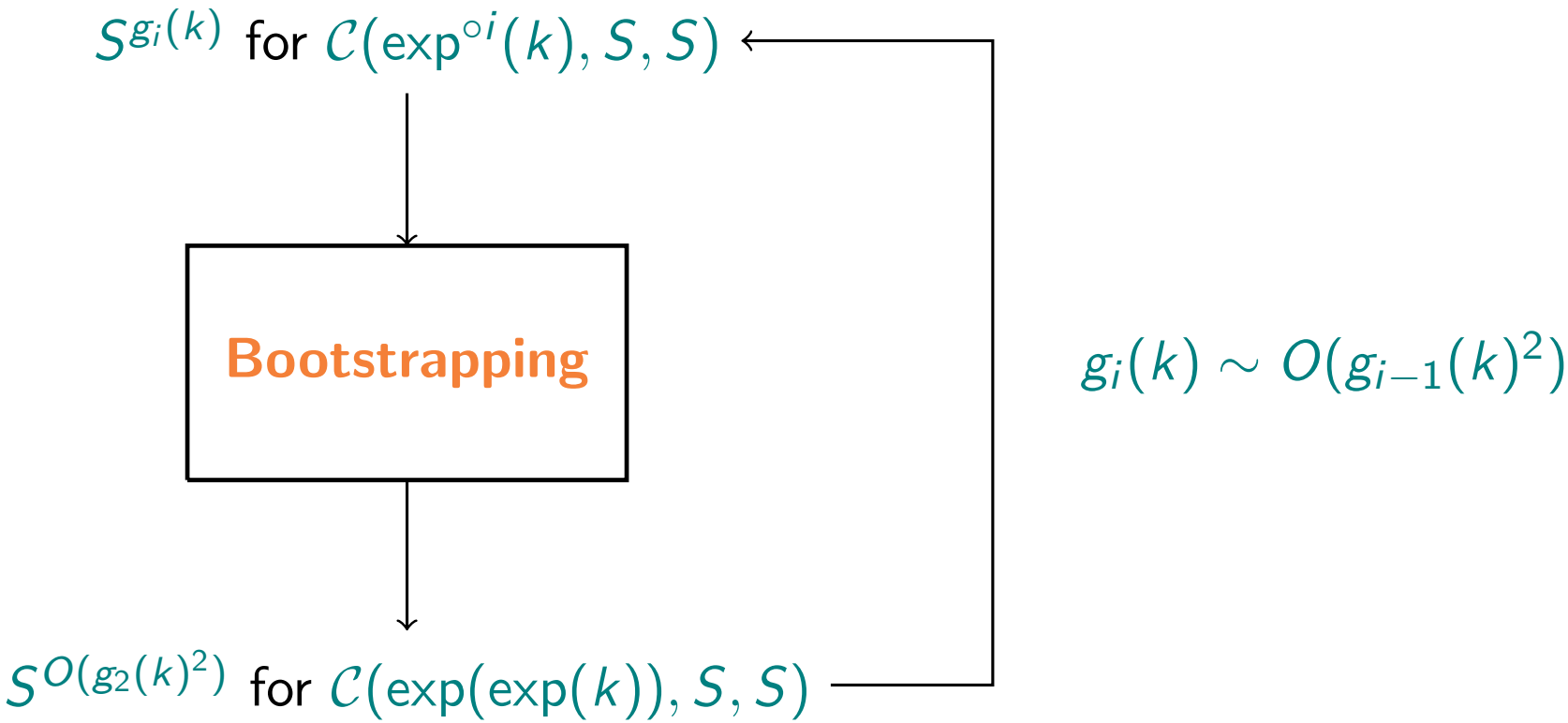
High Level Overview



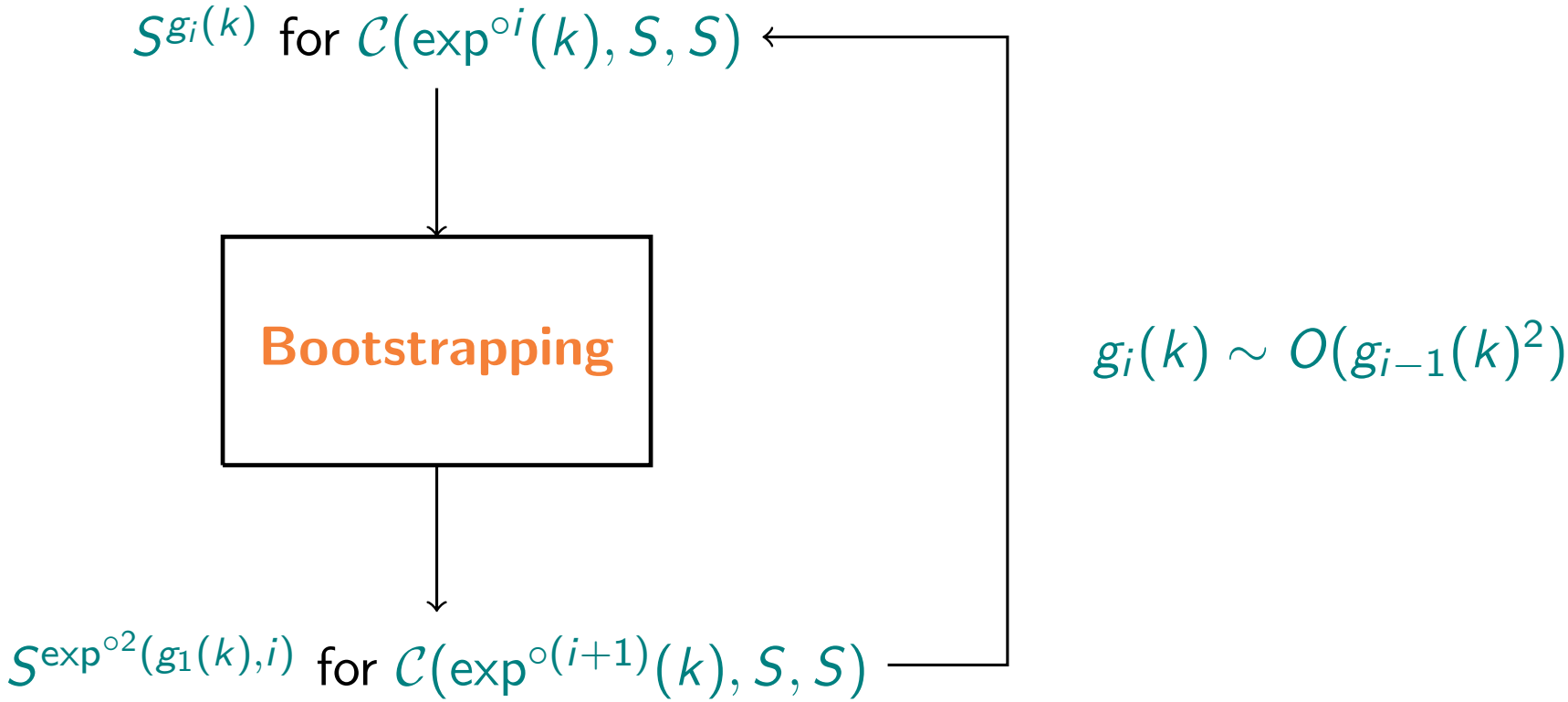
High Level Overview



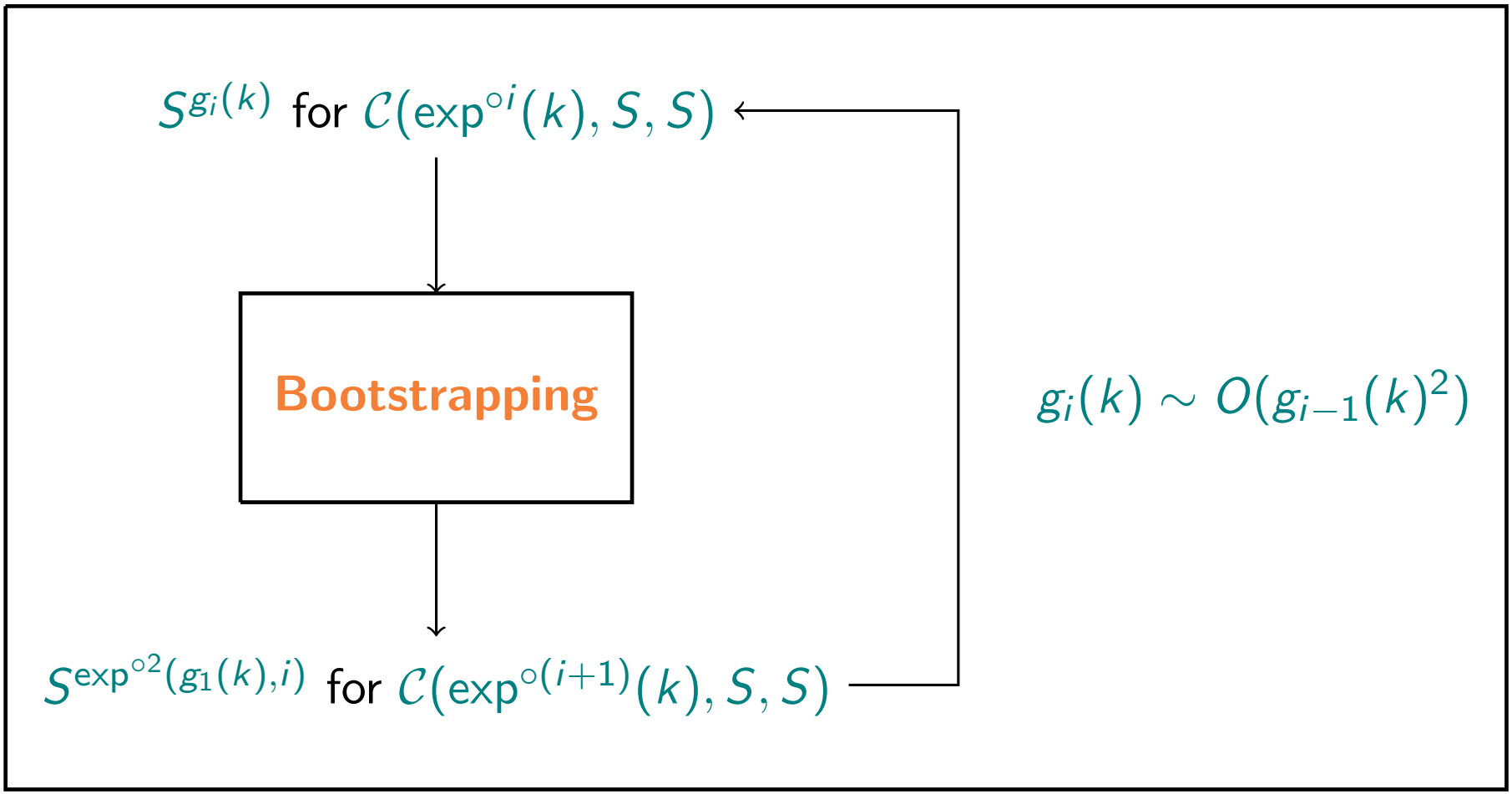
High Level Overview



High Level Overview

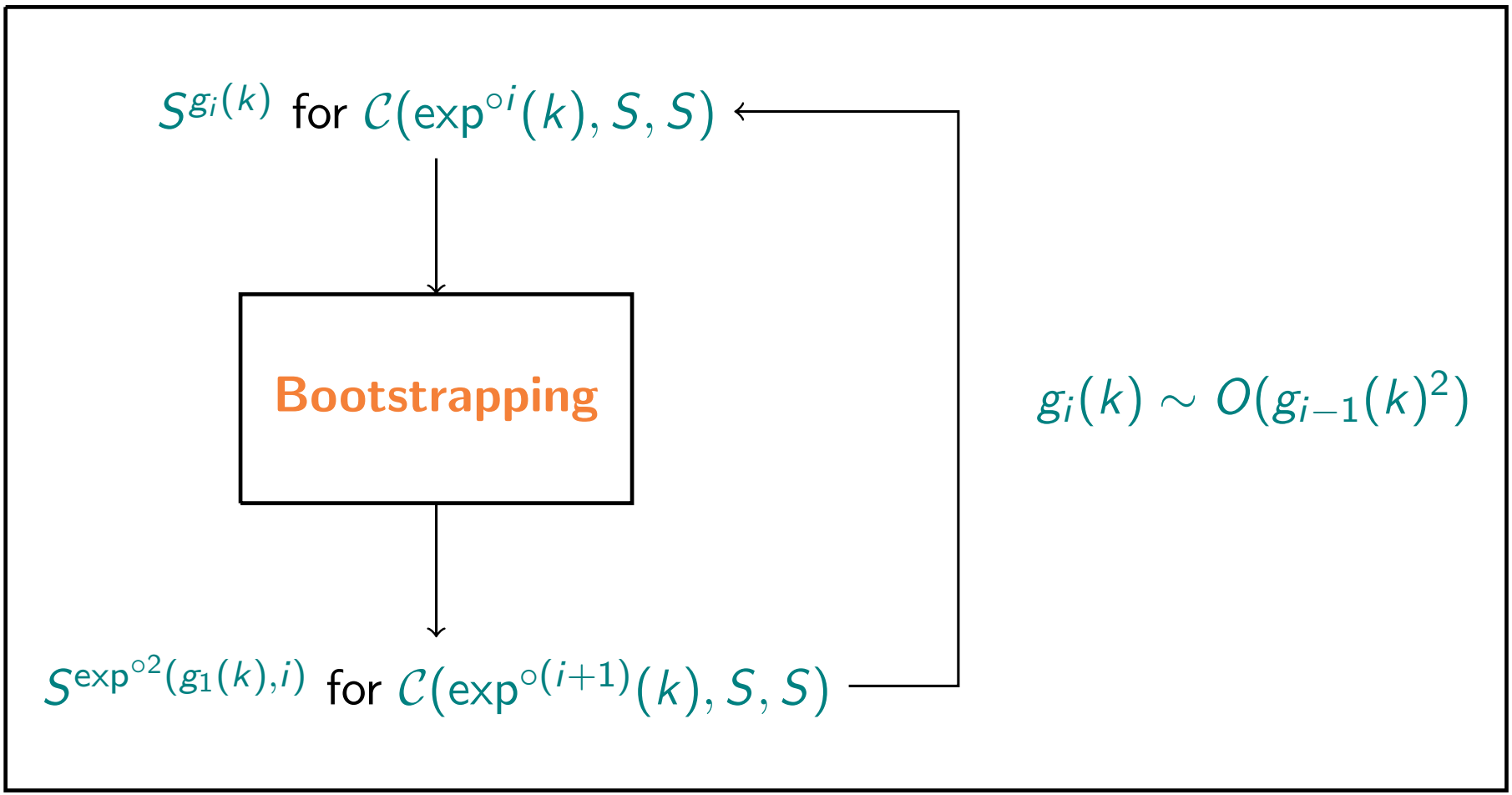


High Level Overview



High Level Overview

$S^{n^{0.2}}$ for $\mathcal{C}(n, S, S) \longrightarrow S^{\exp(\exp(O(\log^* S)))}$ for $\mathcal{C}(S, S, S)$



Summary and Open Questions

Theorem [Kumar, Saptharishi, T]

Suppose for a large constant n and all $s \geq n$, there is an explicit hitting set of size

$$s^{o(n)} \quad \text{for} \quad \mathcal{C}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \text{for} \quad \mathcal{C}(s, s, s).$$

Summary and Open Questions

Theorem* [Kumar, Saptharishi, T (arXiv)]
Suppose for a large constant n and all $s \geq n$, there is an explicit hitting set of size

$$s^{o(n)} \quad \text{for} \quad \mathcal{C}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \text{for} \quad \mathcal{C}(s, s, s).$$

Summary and Open Questions

Theorem* [Kumar, Saptharishi, T (arXiv)]
Suppose for a constant $n \geq 2$ and all $s \geq n$, there is an explicit hitting set of size

$$s^{o(n)} \quad \text{for} \quad \mathcal{C}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \text{for} \quad \mathcal{C}(s, s, s).$$

Summary and Open Questions

Theorem* [Kumar, Saptharishi, T (arXiv)]
Suppose for a constant $n \geq 2$, some $\epsilon > 0$ and all $s \geq n$, there is an explicit hitting set of size

$$s^{n-\epsilon} \quad \text{for} \quad \mathcal{C}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \text{for} \quad \mathcal{C}(s, s, s).$$

Summary and Open Questions

Theorem* [Kumar, Saptharishi, T (arXiv)]
Suppose for a constant $n \geq 2$, some $\epsilon > 0$ and all $s \geq n$, there is an explicit hitting set of size

$$s^{n-\epsilon} \quad \text{for} \quad \text{Formula}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \text{for} \quad \text{Formula}(s, s, s).$$

Summary and Open Questions

Theorem* [Kumar, Saptharishi, T (arXiv)]
Suppose for a constant $n \geq 2$, some $\epsilon > 0$ and all $s \geq n$, there is an explicit hitting set of size

$$s^{n-\epsilon} \quad \textit{for} \quad \text{Formula}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \textit{for} \quad \text{Formula}(s, s, s).$$

Open Questions

Summary and Open Questions

Theorem* [Kumar, Saptharishi, T (arXiv)]
Suppose for a constant $n \geq 2$, some $\epsilon > 0$ and all $s \geq n$, there is an explicit hitting set of size

$$s^{n-\epsilon} \quad \textit{for} \quad \text{Formula}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \textit{for} \quad \text{Formula}(s, s, s).$$

Open Questions

- Can we get to hitting sets of size $\text{poly}(s)$?

Summary and Open Questions

Theorem* [Kumar, Saptharishi, T (arXiv)]
Suppose for a constant $n \geq 2$, some $\epsilon > 0$ and all $s \geq n$, there is an explicit hitting set of size

$$s^{n-\epsilon} \quad \textit{for} \quad \text{Formula}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \textit{for} \quad \text{Formula}(s, s, s).$$

Open Questions

- Can we get to hitting sets of size $\text{poly}(s)$?
- Can we bootstrap lower bounds? Similar to [CILM18].

Summary and Open Questions

Theorem* [Kumar, Saptharishi, T (arXiv)]
Suppose for a constant $n \geq 2$, some $\epsilon > 0$ and all $s \geq n$, there is an explicit hitting set of size

$$s^{n-\epsilon} \quad \textit{for} \quad \text{Formula}(n, s, s).$$

Then for all large s, there is an explicit hitting set of size

$$s^{\text{tiny}(s)} \quad \textit{for} \quad \text{Formula}(s, s, s).$$

Open Questions

- Can we get to hitting sets of size $\text{poly}(s)$?
- Can we bootstrap lower bounds? Similar to [CILM18].

Thank You!