Learning Randomized Reductions

Bridging Theoretical Computer Science, Machine Learning, and AI Agents

🏆 ICML 2026 Spotlight

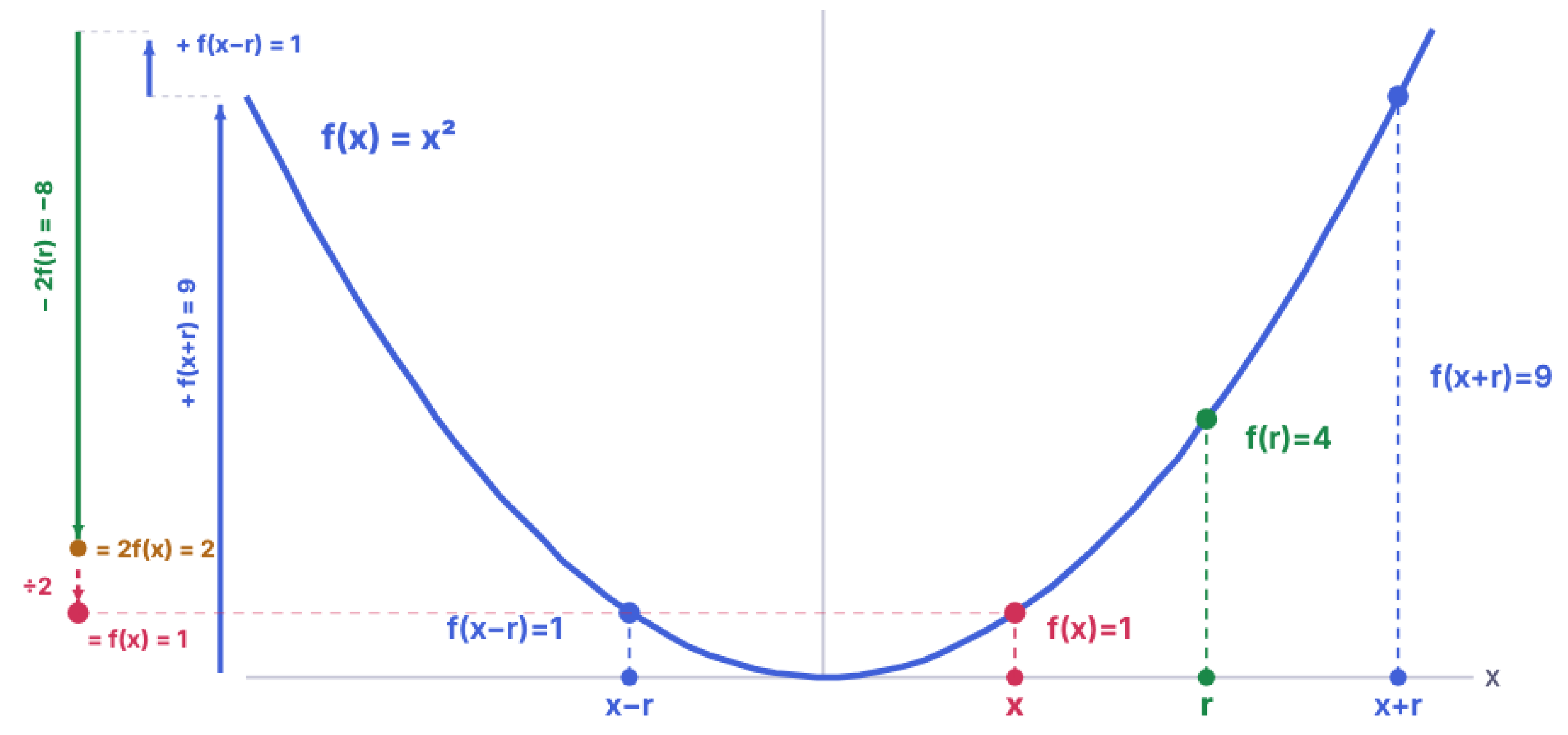
Ferhat Erata[1] · Orr Paradise[2] · Thanos Typaldos[1] · Timos Antonopoulos[1] · ThanhVu Nguyen[3] · **Shafi Goldwasser**[2] · **Ruzica Piskac**[1]

[1]Yale University · [2]UC Berkeley · [3]George Mason University

`RSR` `Self-Correction` `Neuro-Symbolic` `LLM Agents` `ICML 2026`

What is a Randomized Self-Reduction?

Recover $f(x)$ from f evaluated at **random correlated points** — each query point is uniformly random, revealing nothing about x.



EXAMPLE — QUADRATIC $F(X) = X^2$

Three query points give a **linear** recovery:

Queries	Recover
$f(x+r),\ f(x-r),\ f(r)$	$f(x) = \frac{f(x+r)+f(x-r)-2f(r)}{2}$

Lipton (1989): works for *all* low-degree polynomials.

Complexity Theory

Worst-case → average-case reductions

Goldwasser & Micali 1984

Instance-Hiding Protocols

Compute $f(x)$ without revealing x

Abadi et al. 1987

Self-Correcting Programs

90%-correct oracle → **99.99%**

Blum, Luby, Rubinfeld 1993

Two worlds, no bridge

TCS (SINCE 1984)

- **Goldwasser & Micali (1984)** — RSRs first introduced
- **Lipton (1989)** — all low-degree polynomials over finite fields
- **Blum, Luby, Rubinfeld (1993)** — self-correcting programs
- **40 years of manual expert derivation**
- **Handful of fixed queries:** $x+r,\ x-r,\ x\cdot r,\ r$

ML & AI (NOW)

- Powerful **regression** (linear, symbolic, MILP)
- **Symbolic solvers** (SymPy, Mathematica)
- **LLM agents** with tool use & reasoning
- **Neuro-symbolic** reasoning loops
- AI for **mathematical discovery**

Can we bridge them?

THE CHALLENGE

What about non-polynomial functions?

sigmoid $\sigma(x)$	**ReLU** $\max(0, x)$	**GELU** $x\Phi(x)$	**softmax** $e^x / \sum$	**tanh** $\frac{e^x - e^{-x}}{e^x + e^{-x}}$
log $\ln(x)$	**gamma** $\Gamma(x)$	**erf** $\mathrm{erf}(x)$	**Gudermannian** $\mathrm{gd}(x)$	**sinc** $\frac{\sin x}{x}$

No known RSRs for many of these.

For 40 years, nobody found them — because nobody **automated the search**.

SCALING UP

What about higher dimensions?

1D

```
def f(x):
  return 1/(1+exp(-x))
```

sigmoid
curve in 2D

2D

```
def f(x, y):
  return x**2 - y**2
```

saddle surface
surface in 3D

4D

```
def f(a, b, c, d):
  return a**2+b**2+c**2+d**2
```

quaternion norm
can't visualize!

8D

```
def f(a,b,c,d,e,f,g,h):
  return a**2+...+h**2
```

octonion norm
non-associative!

$\det(A)$
matrices

e_2, e_3, p_2
sym. polynomials

$\mathrm{Cl}(3,0)$
Clifford algebras

$\mathrm{tr}([A,B]^2)$
Lie algebras

Between only needs **black-box oracle access** — pass in a Python function, get RSRs back.

Bitween — Three Variants

V-Bitween

Vanilla Bitween

- **No LLM** — pure regression
- Fixed queries:

$\{x+r,\ x-r,\ x\cdot r,\ r\}$

- Backends: LR, PySR, GPLearn, MILP
- **Verifier** (symbolic_verify_tool)

43/80 functions (54%)
LR backend wins

N-Research

Neural Research (baseline)

- **LLM** — no regression
- sequential_thinking
- Relies on parametric knowledge
- **Verifier** (symbolic_verify_tool)

250 RSRs but **172 unverified**
Many false positives

A-Bitween

Agentic Bitween

- **LLM + V-Bitween tools**
- infer_property_tool (regression)
- Novel query function discovery
- **Verifier** (symbolic_verify_tool)

793 RSRs, only **26 unverified**
64/80 functions (80%)

V-Bitween discovers with fixed queries. N-Research lacks Bitween, only accesses verifier. **A-Bitween uses LLM's math prior + V-Bitween and Verifier as tools.**

ICML International Conference On Machine Learning

Vanilla Bitween — Architecture

INPUT

```python
def f(x):
    return 1/(1+math.exp(-x))

bitween.evaluate(
    infer_funcs  = [f],
    domain       = Domain.Real,
    distribution = Uniform(-5, 5),
    queries = [x+r, x-r, x*r, r],
    n=30, max_degree=2,
)
```

→ ① Queries → ② Sample → ③ Monomials → ④ Regression → ⑤ Sparsify → ⑥ Verify →

OUTPUT

$$\sigma(x) = \frac{\sigma(x+r)(\sigma(r)-1)}{2\sigma(x+r)\sigma(r)-\sigma(x+r)-\sigma(r)}$$

0.4s, 15 samples

④ MODULAR REGRESSION BACKEND

Regression	RSRs	Coverage	Time
PySR	61	38%	335s
GPLearn	48	32%	140s
MILP	74	51%	11s
LR	**87**	**54%**	**5s**

Linear regression wins. 43/80 (54%).

But **37 functions remain**. They need *novel query functions* beyond $\{x+r,\ x-r,\ x\cdot r,\ r\}$.

What if the LLM could propose new queries?

Agentic Bitween — LLM + Tools

LLM proposes hypotheses (novel query functions). **Tools ground them** (regression + formal verification).

3x more RSRs than N-Research (same LLM, no tools) with significantly **fewer false positives**.

ICML International Conference On Machine Learning

Novel Query Functions — beyond $\{x{+}r,\ x{-}r,\ x{\cdot}r,\ r\}$

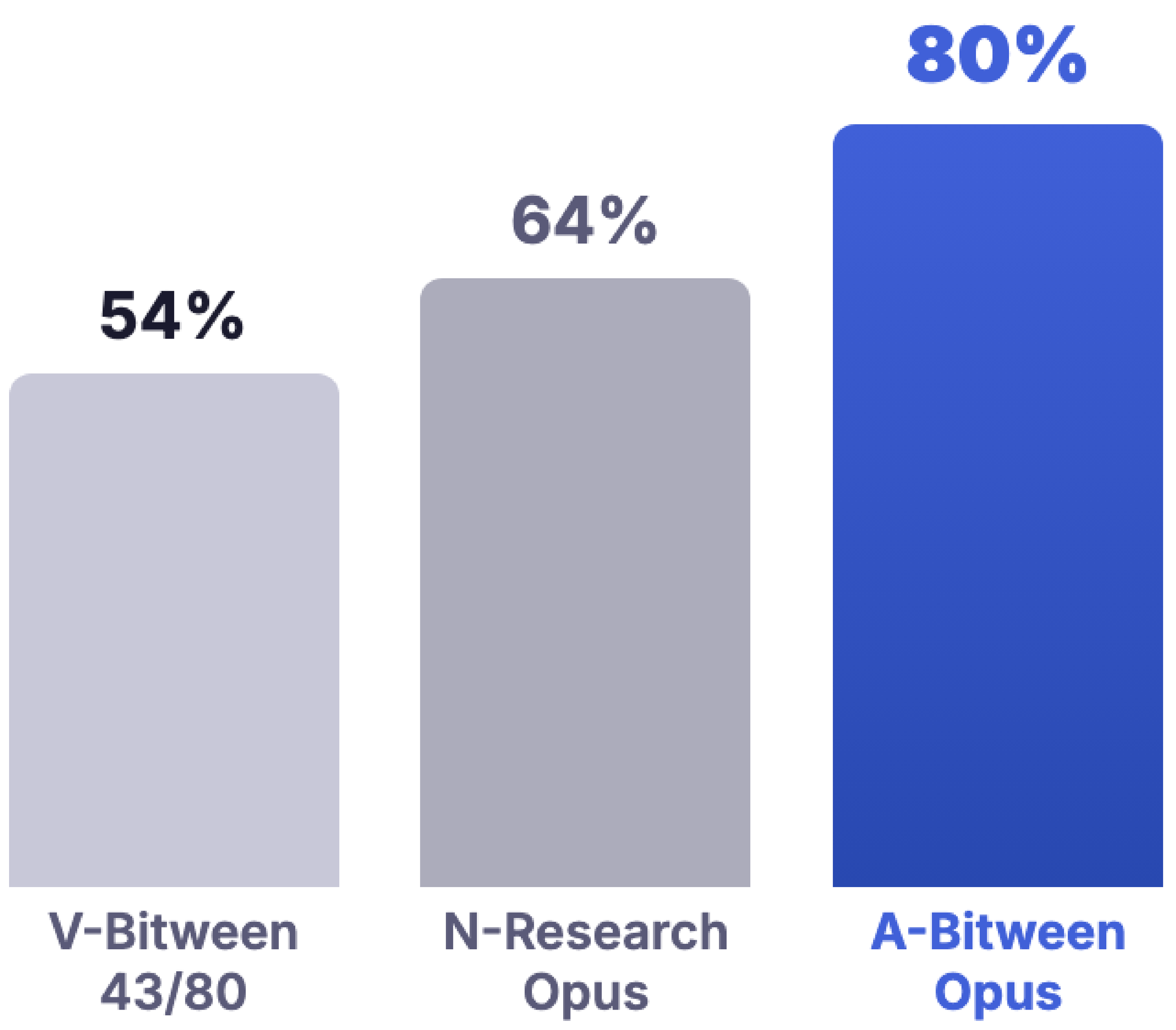
Function	Fixed queries V-Bitween uses	A-Bitween adds (novel queries + properties)
Sigmoid	$f(x{+}r), f(r)$	$f(x{+}\log k) - \frac{k\cdot f(x)}{1+(k-1)f(x)} = 0$ | $f(nx) - \frac{f^n(x)}{(1-f(x))^n+f^n(x)} = 0$
Logarithm	$f(x{\cdot}r), f(r)$	$f(x^n) - n\cdot f(x) = 0$ | $f(\sqrt{x\cdot y}) - \frac{f(x)+f(y)}{2} = 0$
Modulo	$f(x{+}y)$	$f(x{+}y) - f(f(x)+f(y)) = 0$ | $f(x{\cdot}y) - f(f(x)\cdot f(y)) = 0$
Gudermannian	No RSR found	$f(x{+}y) - \arctan(\sinh x\cosh y + \cosh x\sinh y) = 0$
Softmax	$f(x{+}r, y)$	$(ae^x + e^y)\cdot f(x{+}\log a,\ y) - e^y = 0$
Inverse	$f(x{\cdot}y)$	$f(rx) - f^2(x)\cdot f(r/x) = 0$

The LLM proposes **creative query functions** ($x{+}\log k$, $\sqrt{xy}$, $f(x){+}f(y)$) that the fixed query set could never reach.

Regression discovers; formal verification. **All properties above are symbolically verified.**

RESULTS

A-Bitween: 64/80 functions (80%)



WHAT THE TOOLS ADD

	N-Research	A-Bitween
RSRs found	250	**793 (3.2x)**
Unverified	**172**	**26**
Coverage	64%	**80%**

Same LLM (Claude Opus 4.1). The difference is **structured tools**.

WHERE THE 21 NEW FUNCTIONS COME FROM

V-Bitween finds 43/80 with fixed queries.

A-Bitween finds 21 more via **novel query functions** proposed by the LLM.

The gap = LLM's mathematical creativity.

The grounding = regression + verification.

Learning Randomized Reductions

80
functions in RSR-Bench

64
RSRs discovered

175
algebraic RSRs

ICML
2026

`RSR` `Self-Correction` `Neuro-Symbolic` `LLM Agents` `Regression` `Formal Verification`

40 years of theory → automated discovery → algebraic frontiers

Ferhat Erata · Yale University