The Solvability Complexity Index Hierarchy & Generalised Hardness of Approximation

On the role of foundations of computation in computer-assisted proofs

Alexander Bastounis (abastoun@ed.ac.uk)
University of Edinburgh

Matthew Colbrook (m.colbrook@damtp.cam.ac.uk)
University of Cambridge + École Normale Supérieure

http://www.damtp.cam.ac.uk/user/mjc249/home.html: slides, papers, and code

Example 1: Dirac-Schwinger conjecture

N: # of electrons, Z: charge of nucleus

$E(Z) =$ ground state energy of

$$H = \sum_{k=1}^{N} \left(-\Delta_{x_k} - Z|x_k|^{-1}\right) + \sum_{j \le k} \left|x_j - x_k\right|^{-1}.$$

Theorem: $E(Z) = -c_0 Z^{7/3} + \frac{1}{8} Z^2 - c_1 Z^{5/3} + O(Z^{5/3-1/2835})$, as $Z \to \infty$

Proof involves spectral analysis, analytic number theory, …,
computer-assisted bound involving solutions of an ODE.

- Fefferman, Seco, "*Aperiodicity of the Hamiltonian flow in the Thomas-Fermi potential,*" **Rev. Mat. Iberoamericana**, 1993.
- Fefferman, Seco, "*Interval arithmetic in quantum mechanics,*" **Applications of interval computations**, 1996.

Example 2: Kepler conjecture (Hilbert’s 18th problem)

Proof shows potential counterexamples would satisfy infeasible inequalities

relaxed to $\approx$ 10,000s linear programs

More on this later!

- Hales, *“A proof of the Kepler conjecture,”* **Ann. of Math.**, 2005.
- Hales et al., *“A formal proof of the Kepler conjecture,”* **Forum Math. Pi**, 2017.

Account of Flyspeck project (formal proof)

GOAL

Classify how and which computational problems can be used in computer-assisted proofs.

Part I: Infinite-dimensional problems (spectra, PDEs, etc.)

Part II: Finite-dimensional problems (LPs, optimisation, etc.)

Tool: The *Solvability Complexity Index Hierarchy*

- *Classes that allow verifiable error control.*
- *Phase transitions (e.g., $\in P \rightleftharpoons$ not comp.) dep. on the desired accuracy.*

The infinite-dimensional spectral problem

$$A \text{"="} \begin{pmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \qquad A\left(\sum_{k=1}^{\infty} x_k e_k\right) = \sum_{j=1}^{\infty}\left(\sum_{k=1}^{\infty} a_{jk} x_k\right) e_j$$

Canonical basis vectors of $l^2(\mathbb{N})$

Finite-dimensional	$\Longrightarrow$	**Infinite-dimensional**
Eigenvalues of $B \in \mathbb{C}^{n\times n}$	$\Longrightarrow$	Spectrum, $\mathrm{Spec}(A)$
$\{\lambda_j \in \mathbb{C} : \det(B - \lambda_j I) = 0\}$	$\Longrightarrow$	$\{\lambda \in \mathbb{C} : A - \lambda I \text{ is not invertible}\}$

"Most operators that arise in practice are not presented in a representation in which they are diagonalized, and it is often very hard to locate even a single point in the spectrum. Thus, one often has to settle for numerical approximations [...] Unfortunately, there is a dearth of literature on this basic problem and, so far as we have been able to tell, ***there are no proven [general] techniques****."*

W. Arveson, Berkeley (1994)

Why spectra?

Applications: Quantum mechanics, structural mechanics, optics, acoustics, statistical physics, number theory, matter physics, PDEs, data analysis, neural networks and AI, nuclear scattering, optics, computational chemistry, …

Rich history of **computational spectral theory:**

D. Arnold (Minnesota), W. Arveson (Berkeley), A. Böttcher (Chemnitz), W. Dahmen (South Carolina), E. B. Davies (KCL), P. Deift (NYU), L. Demanet (MIT), C. Fefferman (Princeton), G. Golub (Stanford), A. Iserles (Cambridge), I. Ipsen (NCS), S. Jitomirskaya (UCI), A. Laptev (Imperial), O. Nevanlinna (Aalto), W. Schlag (Yale), E. Schrödinger (DIAS), J. Schwinger (Harvard), N. Trefethen (Oxford), V. Varadarajan (UCLA), S. Varadhan (NYU), J. von Neumann (IAS), M. Zworski (Berkeley),…

Many computer-assisted proofs involve spectra: dynamical systems, hydrodynamics, atomic resonances, etc.

Motivating problem

In a series of papers in the 1950's and 1960's, J. Schwinger examined the foundations of quantum mechanics. A key problem he considered:

Given a self-adjoint Schrödinger operator $-\Delta + V$ on $\mathbb{R}$, can we approximate its spectrum?

Partial answer: T. Digernes, V. S. Varadarajan and S. R. S. Varadhan (1994) gave a convergent algorithm for a class of V generating compact resolvent.

For which classes of differential operators on unbounded domains do there exist algorithms that converge to the spectrum? Can we guarantee that the output is in the spectrum up to an arbitrarily small tolerance?

- Digernes, Varadarajan, Varadhan, "*Finite approximations to quantum systems,*" **Rev. Math. Phys.**, 1994.

Warm-up: bounded diagonal operators

$$A = \begin{pmatrix} a_1 & & \\ & a_2 & \\ & & \ddots \end{pmatrix}$$

Assumption: Algorithm can query entries of A.

Algorithm: $\Gamma_n(A) = \{a_1, a_2, \ldots, a_n\} \to \mathrm{Spec}(A) = \overline{\{a_1, a_2, \ldots\}}$ in Haus. Metric.

One-sided error control: $\Gamma_n(A) \subset \mathrm{Spec}(A)$

Optimal: Can't obtain $\hat{\Gamma}_n(A) \to \mathrm{Spec}(A)$ with $\mathrm{Spec}(A) \subset \hat{\Gamma}_n(A)$.

Example: compact operators (still easy?)

classic method

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Algorithm: $\Gamma_n(A) = \mathrm{Spec}(P_n A P_n)$ converges to $\mathrm{Spec}(A)$ in Haus. Metric.

Question: Can we verify the output?

i.e., Does there exist $\hat{\Gamma}_n(A) \to \mathrm{Spec}(A)$ with $\hat{\Gamma}_n(A) \subset \mathrm{Spec}(A) + B_{2^{-n}}$?

Example: compact operators (still easy?)

classic method

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Algorithm: $\Gamma_n(A) = \mathrm{Spec}(P_n A P_n)$ converges to $\mathrm{Spec}(A)$ in Haus. Metric.

Question: Can we verify the output?

i.e., Does there exist $\hat{\Gamma}_n(A) \to \mathrm{Spec}(A)$ with $\hat{\Gamma}_n(A) \subset \mathrm{Spec}(A) + B_{2^{-n}}$?

Answer: No!

No alg. can do this on whole class, even for self-adjoint compact operators.

What about Jacobi operators?

$$A = \begin{pmatrix} a_1 & b_1 & & \\ b_1 & a_2 & b_2 & \\ & b_2 & a_3 & \ddots \\ & & \ddots & \ddots \end{pmatrix}, \qquad b_k > 0, \qquad a_k \in \mathbb{R}$$

Non-trivial, e.g., spurious eigenvalues.

What about Jacobi operators?

$$A = \begin{pmatrix} a_1 & b_1 & & \\ b_1 & a_2 & b_2 & \\ & b_2 & a_3 & \ddots \\ & & \ddots & \ddots \end{pmatrix}, \qquad b_k > 0, \qquad a_k \in \mathbb{R}$$

Non-trivial, e.g., spurious eigenvalues.

Enlarge class to **sparse normal operators** - surely now much harder?!

What about Jacobi operators?

$$A = \begin{pmatrix} a_1 & b_1 & & \\ b_1 & a_2 & b_2 & \\ & b_2 & a_3 & \ddots \\ & & \ddots & \ddots \end{pmatrix}, \qquad b_k > 0, \qquad a_k \in \mathbb{R}$$

Non-trivial, e.g., spurious eigenvalues.

Enlarge class to **sparse normal operators** - surely now much harder?!

Answer: $\exists\{\Gamma_n\}$ s.t. $\lim_{n\to\infty} \Gamma_n(A) = \mathrm{Spec}(A)$ and $\Gamma_n(A) \subset \mathrm{Spec}(A) + B_{2^{-n}}$,

for any sparse normal operator A

- C., Roman, Hansen, *"How to compute spectra with error control,"* **Phys. Rev. Lett.**, 2019.
- Ben-Artzi, C., Hansen, Nevanlinna, Seidel, *"On the solvability complexity index hierarchy and towers of algorithms,"* preprint.

A curious case of limits

General bounded: $$A = \begin{pmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Algorithm: $\exists\{\Gamma_{n_3,n_2,n_1}\}$ s.t. $\lim\limits_{n_3\to\infty}\lim\limits_{n_2\to\infty}\lim\limits_{n_1\to\infty}\Gamma_{n_3,n_2,n_1}(A) = \mathrm{Spec}(A)$

Question: Can we do better?

- Hansen, *"On the solvability complexity index, the n-pseudospectrum and approximations of spectra of operators,"* **J. Amer. Math. Soc.**, 2011.
- Ben-Artzi, C., Hansen, Nevanlinna, Seidel, *"On the solvability complexity index hierarchy and towers of algorithms,"* preprint.

A curious case of limits

General bounded: $$A = \begin{pmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Algorithm: $\exists\{\Gamma_{n_3,n_2,n_1}\}$ s.t. $\lim_{n_3\to\infty}\lim_{n_2\to\infty}\lim_{n_1\to\infty}\Gamma_{n_3,n_2,n_1}(A) = \mathrm{Spec}(A)$

Question: Can we do better?

Answer: No! Canonically embed problems such as:

Given $B \in \{0,1\}^{\mathbb{N}\times\mathbb{N}}$, does B have a column with infinitely many 1's?

$\Longrightarrow$ lower bound on number of "successive limits" needed (ind. of comp. model).

- Hansen, *"On the solvability complexity index, the n-pseudospectrum and approximations of spectra of operators,"* **J. Amer. Math. Soc.**, 2011.
- Ben-Artzi, C., Hansen, Nevanlinna, Seidel, *"On the solvability complexity index hierarchy and towers of algorithms,"* preprint.
- C., *"On the computation of geometric features of spectra of linear operators on Hilbert spaces,"* **Found. Comput. Math.**, to appear.

Solvability Complexity Index Hierarchy

Class $\Omega \ni A$, want to compute $\Xi\colon \Omega \to (\mathcal{M}, d)$ metric space

- Δ_0: Problems solved in finite time (v. rare for cts problems).
- Δ_1: Problems solved in "one limit" with full error control:

$$d(\Gamma_n(A), \Xi(A)) \leq 2^{-n}$$

- Δ_2: Problems solved in "one limit":

$$\lim_{n\to\infty} \Gamma_n(A) = \Xi(A)$$

- Δ_3: Problems solved in "two successive limits":

$$\lim_{n\to\infty} \lim_{m\to\infty} \Gamma_{n,m}(A) = \Xi(A)$$

$\vdots$

- Ben-Artzi, C., Hansen, Nevanlinna, Seidel, "*On the solvability complexity index hierarchy and towers of algorithms,*" preprint.
- Hansen, "*On the solvability complexity index, the n-pseudospectrum and approximations of spectra of operators,*" **J. Amer. Math. Soc.**, 2011.
- McMullen, "*Families of rational maps and iterative root-finding algorithms,*" **Ann. of Math.**, 1987.
- Doyle, McMullen, "*Solving the quintic by iteration,*" **Acta Math.**, 1989.
- Smale, "*The fundamental theorem of algebra and complexity theory,*" **Bull. Amer. Math. Soc.**, 1981.

Solvability Complexity Index Hierarchy

Class $\Omega \ni A$, want to compute $\Xi: \Omega \to (\mathcal{M}, d)$ metr… …ace

- Δ_0: Problems solved in finite time (v. rare for cts problem…
- Δ_1: Problems solved in "one limit" with full erro…

$$d(\Gamma_n(A), \Xi(A))…$$

- Δ_2: Problems solved in "one li…

$$…(A) = \Xi(A)$$

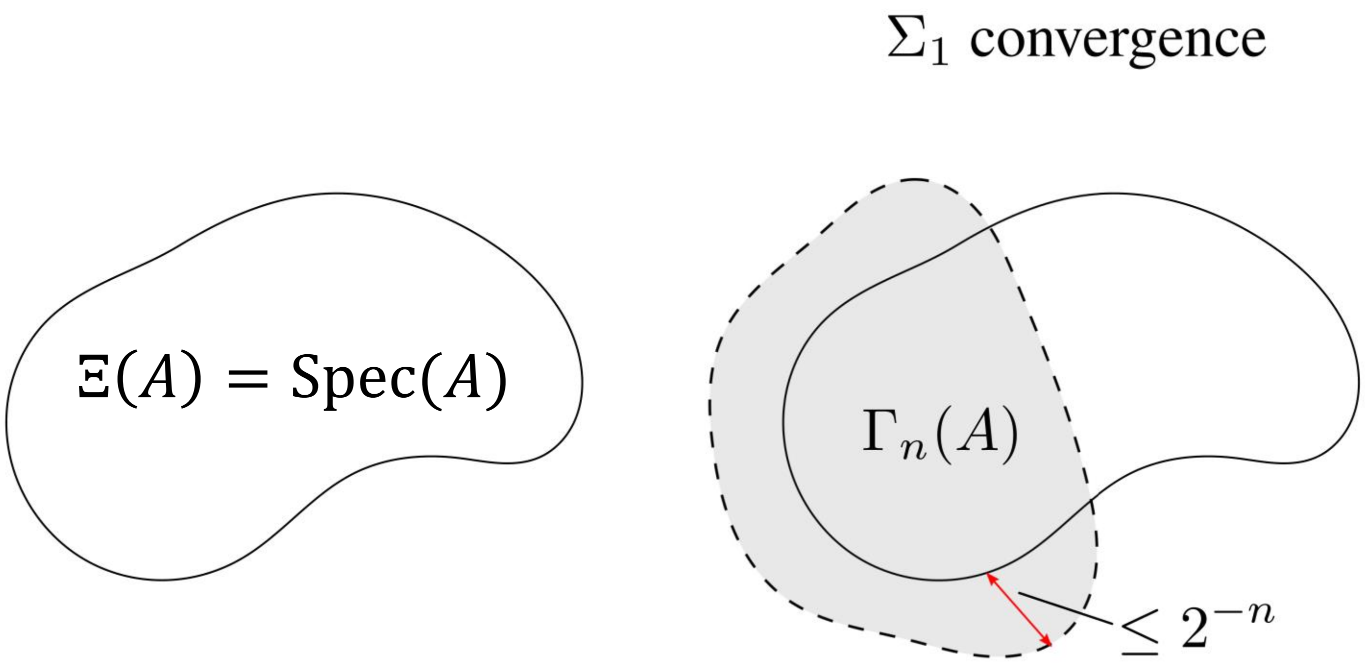
- Δ_3: Probl… …in "two successive limits":

$$\lim_{n\to\infty} \lim_{m\to\infty} \Gamma_{n,m}(A) = \Xi(A)$$

Not the full picture for computer-assisted proofs…

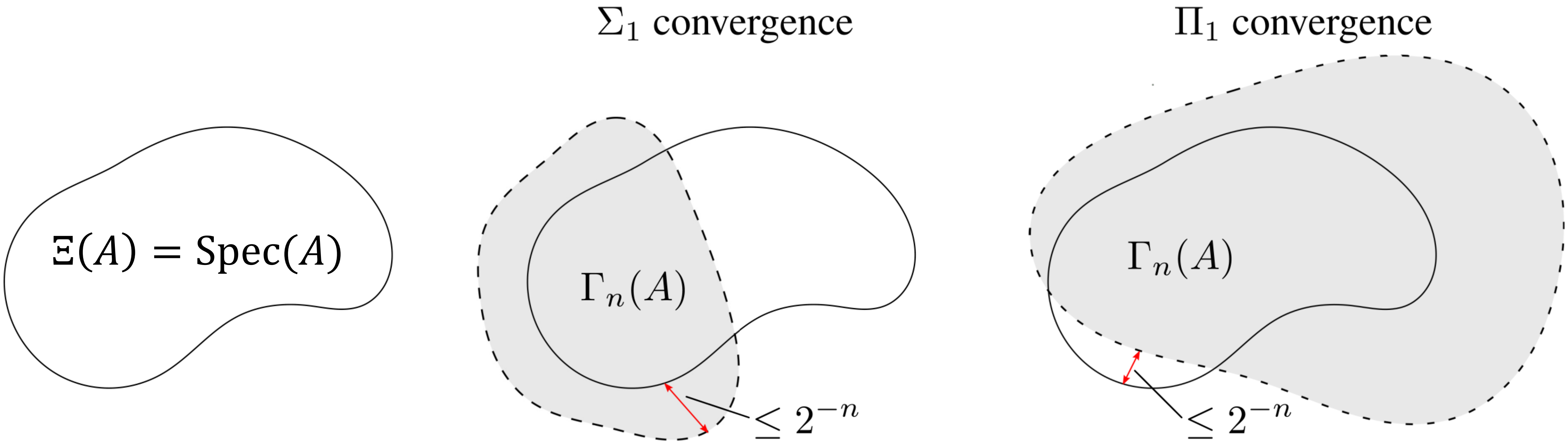
- Ben-Artzi, C., Hansen, Nevanlinna, Seidel, "*On the solvability complexity index hierarchy and towers of algorithms,*" preprint.
- Hansen, "*On the solvability complexity index, the n-pseudospectrum and approximations of spectra of operators,*" **J. Amer. Math. Soc.**, 2011.
- McMullen, "*Families of rational maps and iterative root-finding algorithms,*" **Ann. of Math.**, 1987.
- Doyle, McMullen, "*Solving the quintic by iteration,*" **Acta Math.**, 1989.
- Smale, "*The fundamental theorem of algebra and complexity theory,*" **Bull. Amer. Math. Soc.**, 1981.

Error control for spectral problems



- Σ_1: $\exists$ alg. $\{\Gamma_n\}$ s.t. $\lim_{n\to\infty} \Gamma_n(A) = \Xi(A)$, $\max_{z\in\Gamma_n(A)} \mathrm{dist}\big(z, \Xi(A)\big) \leq 2^{-n}$

Error control for spectral problems



- Σ_1: $\exists$ alg. $\{\Gamma_n\}$ s.t. $\lim_{n\to\infty} \Gamma_n(A) = \Xi(A)$, $\max_{z\in\Gamma_n(A)} \mathrm{dist}\big(z, \Xi(A)\big) \leq 2^{-n}$
- Π_1: $\exists$ alg. $\{\Gamma_n\}$ s.t. $\lim_{n\to\infty} \Gamma_n(A) = \Xi(A)$, $\max_{z\in\Xi(A)} \mathrm{dist}(z, \Gamma_n(A)) \leq 2^{-n}$

Such problems can be used in a proof!

Sample: some results for bounded op. on $l^2(\mathbb{N})$

Increasing difficulty

Error control, comp.-assisted proofs

$$
\begin{array}{ccccccccc}
\Pi_0 & & & & & \Pi_1 & & & \\
\| & & & \subsetneq & & & \subsetneq & & \\
\Delta_0 & \subsetneq & \Delta_1 & \subsetneq & \Sigma_1 & \cup & \Pi_1 & \subsetneq & \\
\| & & & \subsetneq & & & \subsetneq & & \\
\Sigma_0 & & & & & \Sigma_1 & & &
\end{array}
$$

Sample: some results for bounded op. on $l^2(\mathbb{N})$

Increasing difficulty

Error control, comp.-assisted proofs

One limit, no error control.

$$\begin{array}{ccccccc}
\Pi_0 & & & & \Pi_1 & & \\
\| & & & \subsetneq & & \subsetneq & \\
\Delta_0 & \subsetneq & \Delta_1 & \subsetneq & \Sigma_1 \cup \Pi_1 & \subsetneq & \Delta_2 \\
\| & & & \subsetneq & & \subsetneq & \\
\Sigma_0 & & & & \Sigma_1 & &
\end{array}$$

One limit: SCI≤1

Sample: some results for bounded op. on $l^2(\mathbb{N})$

Increasing difficulty

Error control, comp.-assisted proofs

$$\begin{array}{ccccccccccccccc}
\Pi_0 & & & & & \Pi_1 & & & & & \Pi_2 & & & \\
\| & & & \rotatebox{45}{$\subsetneq$} & & & \rotatebox{-45}{$\subsetneq$} & & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \\
\Delta_0 & \subsetneq & \Delta_1 & \subsetneq & \Sigma_1 & \cup & \Pi_1 & \subsetneq & \Delta_2 & \subsetneq & \Sigma_2 \cup \Pi_2 & \subsetneq & \Delta_3 \\
\| & & & \rotatebox{-45}{$\subsetneq$} & & & \rotatebox{45}{$\subsetneq$} & & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \\
\Sigma_0 & & & & & \Sigma_1 & & & & & \Sigma_2 & & &
\end{array}$$

Two limits: SCI ≤ 2

Sample: some results for bounded op. on $l^2(\mathbb{N})$

Increasing difficulty

Error control, comp.-assisted proofs

$$
\begin{array}{ccccccccccccccc}
\Pi_0 & & & & \Pi_1 & & & & \Pi_2 & & & & \Pi_3 & & \\
\| & & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \cdots & \\
\Delta_0 & \subsetneq & \Delta_1 & \subsetneq & \Sigma_1 \cup \Pi_1 & \subsetneq & \Delta_2 & \subsetneq & \Sigma_2 \cup \Pi_2 & \subsetneq & \Delta_3 & \subsetneq & \Sigma_3 \cup \Pi_3 & \cdots & \\
\| & & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \cdots & \\
\Sigma_0 & & & & \Sigma_1 & & & & \Sigma_2 & & & & \Sigma_3 & &
\end{array}
$$

Three limits: SCI ≤ 3 …

Sample: some results for bounded op. on $l^2(\mathbb{N})$

Increasing difficulty

Error control, comp.-assisted proofs

"Sparse" operators

Compact operators

$$
\begin{array}{ccccccccccccccc}
\Pi_0 & & & & \Pi_1 & & & & \Pi_2 & & & & \Pi_3 & & \\
\| & & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \ddots & \\
\Delta_0 & \subsetneq & \Delta_1 & \subsetneq & \Sigma_1 \cup \Pi_1 & \subsetneq & \Delta_2 & \subsetneq & \Sigma_2 \cup \Pi_2 & \subsetneq & \Delta_3 & \subsetneq & \Sigma_3 \cup \Pi_3 & \cdots & \\
\| & & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \iddots & \\
\Sigma_0 & & & & \Sigma_1 & & & & \Sigma_2 & & & & \Sigma_3 & &
\end{array}
$$

General operators

Normal operators

Sample: some results for bounded op. on $l^2(\mathbb{N})$

Increasing difficulty

Error control, comp.-assisted proofs

"Sparse" operators

Compact operators

$$\Pi_0 = \Delta_0 = \Sigma_0$$

$$\begin{array}{ccccccccccccccc}
 & & & & \Pi_1 & & & & \Pi_2 & & & & \Pi_3 & & \\
 & & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & & \\
\Delta_0 & \subsetneq & \Delta_1 & \subsetneq & \Sigma_1 \cup \Pi_1 & \subsetneq & \Delta_2 & \subsetneq & \Sigma_2 \cup \Pi_2 & \subsetneq & \Delta_3 & \subsetneq & \Sigma_3 \cup \Pi_3 & \cdots & \\
 & & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & & \\
 & & & & \Sigma_1 & & & & \Sigma_2 & & & & \Sigma_3 & &
\end{array}$$

"Sparse" normal operators

Normal operators

General operators

Sample: some results for bounded op. on $l^2(\mathbb{N})$

Increasing difficulty

Error control, comp.-assisted proofs

Compact operators

"Sparse" operators

$$\begin{tikzcd}
\Pi_0 \arrow[d, phantom, "\|"] & & \Pi_1 \arrow[dr, phantom, "\subsetneq"] & & \Pi_2 \arrow[dr, phantom, "\subsetneq"] & & \Pi_3 & \ddots \\
\Delta_0 \arrow[r, phantom, "\subsetneq"] \arrow[d, phantom, "\|"] & \Delta_1 \arrow[ur, phantom, "\subsetneq"] \arrow[dr, phantom, "\subsetneq"] \arrow[r, phantom, "\subsetneq"] & \Sigma_1 \cup \Pi_1 \arrow[r, phantom, "\subsetneq"] & \Delta_2 \arrow[ur, phantom, "\subsetneq"] \arrow[dr, phantom, "\subsetneq"] \arrow[r, phantom, "\subsetneq"] & \Sigma_2 \cup \Pi_2 \arrow[r, phantom, "\subsetneq"] & \Delta_3 \arrow[ur, phantom, "\subsetneq"] \arrow[dr, phantom, "\subsetneq"] \arrow[r, phantom, "\subsetneq"] & \Sigma_3 \cup \Pi_3 & \cdots \\
\Sigma_0 & & \Sigma_1 \arrow[ur, phantom, "\subsetneq"] & & \Sigma_2 \arrow[ur, phantom, "\subsetneq"] & & \Sigma_3 & \iddots
\end{tikzcd}$$

"Sparse" normal operators

Normal operators

General operators

Zoo of problems: spectral type (pure point, absolutely continuous, singularly continuous), Lebesgue measure and fractal dimensions of spectra, discrete spectra, essential spectra, eigenspaces + multiplicity, spectral radii, essential numerical ranges, geometric features of spectrum (e.g., capacity), spectral gap problem, resonances ...

- C., "*The foundations of infinite-dimensional spectral computations,*" PhD diss., University of Cambridge, 2020.

Example (local uniform convergence)

Theorem: Let Ω be class of self-adjoint diff. operators on $L^2(\mathbb{R}^d)$ of the form

$$T = \sum_{k \in \mathbb{Z}_{\geq 0}^d, |k| \leq N} c_k(x)\, \partial^k \qquad \text{s.t.}$$

- Smooth compactly supported functions form a core of T.
- $\{c_k\}$ are polynomially bounded and of locally bounded total variation.

Assume algorithm can:

- Point sample $\{c_k(q)\}$ for $q \in \mathbb{Q}^d$ to arbitrary prec.
- Evaluate a polynomial that bounds $\{c_k\}$ on $\mathbb{R}^d$.

Then…

- C., Hansen, *"The foundations of spectral computations via the solvability complexity index hierarchy,"* **J. Eur. Math. Soc.**, 2022

Example (local uniform convergence)

Theorem: Let Ω be class of self-adjoint diff. operators on $L^2(\mathbb{R}^d)$ of the form

$$T = \sum_{k\in\mathbb{Z}^d_{\geq 0}, |k|\leq N} c_k(x)\,\partial^k \qquad \text{s.t.}$$

- Smooth compactly supported functions form a core of T.
- $\{c_k\}$ are polynomially bounded and of locally bounded total variation.

Assume algorithm can:

- Point sample $\{c_k(q)\}$ for $q\in\mathbb{Q}^d$ to arbitrary prec.
- Evaluate a polynomial that bounds $\{c_k\}$ on $\mathbb{R}^d$.

Then

Verifiable

Not verifiable

(a) Know bound $\mathrm{TV}_{[-n,n]^d}(c_k)\leq b_n \Longrightarrow \{\mathrm{Sp},\Omega\}\in\Sigma_1$.

(b) Only know asymp. bound $\mathrm{TV}_{[-n,n]^d}(c_k)=O(b_n)\Longrightarrow \{\mathrm{Sp},\Omega\}\in\Delta_2\backslash(\Sigma_1\cup\Pi_1)$.

- C., Hansen, *"The foundations of spectral computations via the solvability complexity index hierarchy,"* **J. Eur. Math. Soc.**, 2022

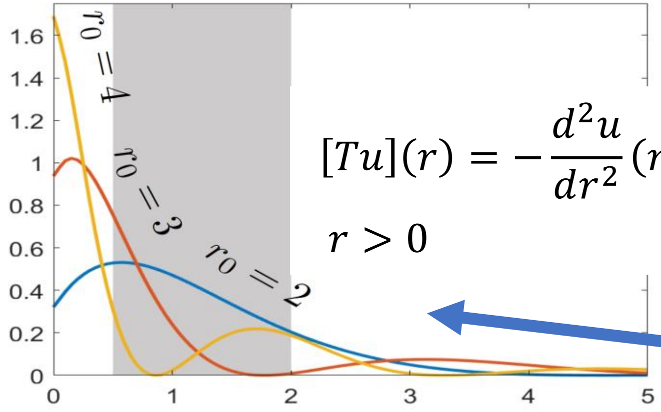
Further examples

Examples with discrete spectra

$$T = -\nabla^2 + x^2 + V(x) \text{ on } \mathbb{R}^1$$

V	$\cos(x)$	$\tanh(x)$	$\exp(-x^2)$	$1/(1+x^2)$
E_0	1.7561051579	0.8703478514	1.6882809272	1.7468178026
E_1	3.3447026910	2.9666370800	3.3395578680	3.4757613534
E_2	5.0606547136	4.9825969775	5.2703748823	5.4115076464
E_3	6.8649969390	6.9898951678	7.2225903394	7.3503220313
E_4	8.7353069954	8.9931317537	9.1953373991	9.3168983920

Spectral measures



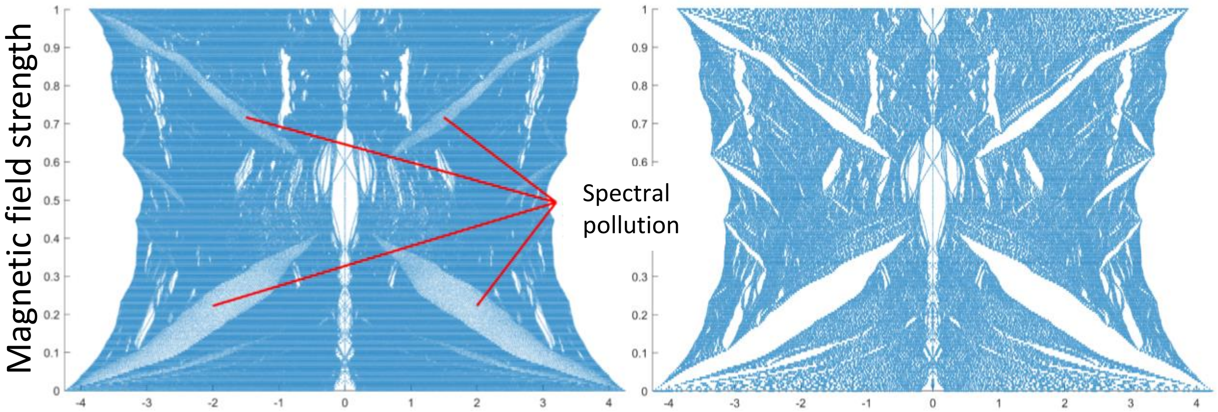
$$[Tu](r) = -\frac{d^2u}{dr^2}(r) + \left(\frac{l(l+1)}{r^2} + \frac{e^{-r}-1}{r}\right)u(r),$$

$r > 0$

Wavefunction $\propto e^{-(r-r_0)^2}$

Example with mixed spectra: aperiodic material + mag. field

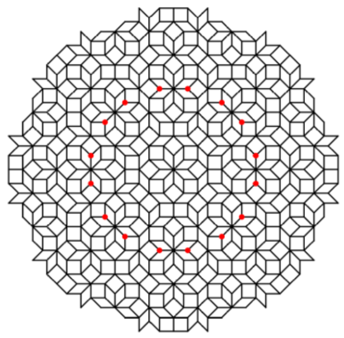
Naïve method **Verified method**

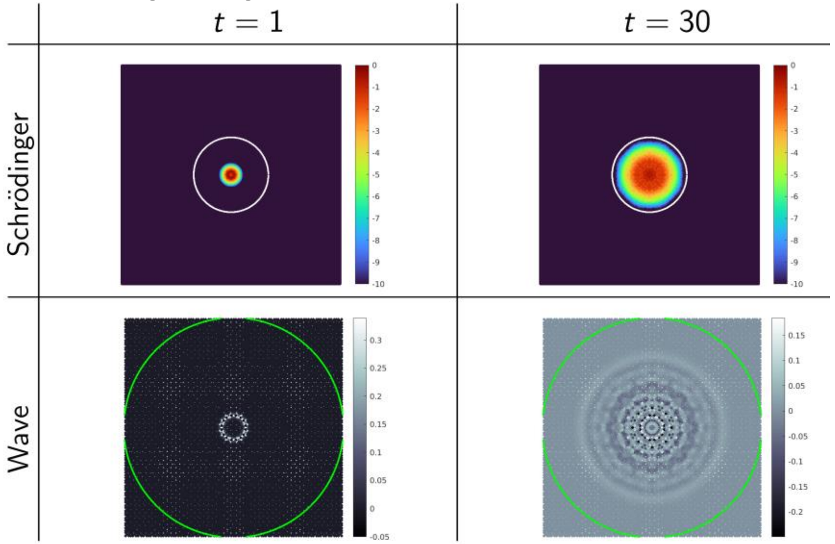


Approx. of spectrum Approx. of spectrum

Transport phenomena

Inf. aperiodic tile





- C., *"Computing spectral measures and spectral types,"* **Comm. Math. Phys.**, 2021

Why study these foundations?

- Classifications with SCI>1 often tell us assumptions we need to lower SCI.
- Σ_1 and Π_1 classifications $\Longrightarrow$ look-up table for computer-assisted proofs.
- Negative results prevent us from trying to prove too much.
- Much of computational literature does not prove sharp results.

Remarks:

- Can use with any model of computation.
- Existing hierarchies included as particular cases.

What if we know a priori that we only need an ε-accurate solution for a computer-assisted proof?

Linear Programs (LPs)

The proof of Kepler's conjecture involves solving 10,000s of LPs.

Problem: Find algorithm that for input $A \in \mathbb{R}^{m \times N}, y \in \mathbb{R}^m, c \in \mathbb{R}^N$, computes

$$z \in \underset{x}{\operatorname{argmin}} \langle x, c \rangle \qquad \text{s.t.} \quad Ax = y, \qquad x \geq 0.$$

Linear Programs (LPs)

The proof of Kepler's conjecture involves solving 10,000s of LPs.

Problem: Find algorithm that for input $A \in \mathbb{R}^{m \times N}, y \in \mathbb{R}^m, c \in \mathbb{R}^N$, computes

$$z \in \underset{x}{\operatorname{argmin}} \langle x, c \rangle \qquad \text{s.t.} \quad Ax = y, \qquad x \geq 0.$$

Often, we want **minimisers**, and not the objective function.

Linear Programs (LPs)

The proof of Kepler's conjecture involves solving 10,000s of LPs.

Problem: Find algorithm that for input $A \in \mathbb{R}^{m\times N}, y \in \mathbb{R}^m, c \in \mathbb{R}^N$, computes

$$z \in \operatorname*{argmin}_x \langle x, c\rangle \qquad \text{s.t.} \quad Ax = y, \qquad x \geq 0.$$

Often, we want **minimisers**, and not the objective function.

LP in P


"All the News That's Fit to Print"

The New York Times

LATE CITY EDITION

Similar Transportation Bond Issues Are Passed in New York and Jersey

CARTER SAID TO ASK FOR LOYALTY CHECK OF AIDES ON POLITICS

IRAN'S CIVIL GOVERNMENT OUT; HOSTAGES FACE DEATH THREAT; OIL EXPORTS BELIEVED HALTED

STUDENTS WARN U.S.

Ayatollah Instructs Secret Revolutionary Council to Form a Cabinet

F.A.A. Seeks Fine of $1.5 Million For 'Unairworthy' Braniff Flights

Main Iran Oil Port Reported Closed And Prices of Spot Petroleum Soar

Energy Dept. Accuses 9 Refiners Of $1.18 Billion Oil Overcharges

A Soviet Discovery Rocks World of Mathematics

NY Times 1979. Proved by L. Khachiyan
– based on work by N. Shor, D. Yudin, A. Nemirovski.

LP in P

NY Times 1979. Proved by L. Khachiyan
– based on work by N. Shor, D. Yudin, A. Nemirovski.

Example (Karmarkar's) algorithm:

- n = number of variables
- L = number of bits

$O(n^{3.5}L^2 \log(L) \log(\log(L)))$ operations

Weakly polynomial time

LP in P


The New York Times

Similar Transportation Bond Issues Are Passed in New York and Jersey

CARTER SAID TO ASK FOR LOYALTY CHECK OF AIDES ON POLITICS

IRAN'S CIVIL GOVERNMENT OUT; HOSTAGES FACE DEATH THREAT; OIL EXPORTS BELIEVED HALTED

STUDENTS WARN U.S.

F.A.A. Seeks Fine of $1.5 Million For 'Unairworthy' Braniff Flights

Main Iran Oil Port Reported Closed And Prices of Spot Petroleum Soar

Energy Dept. Accuses 9 Refiners Of $1.18 Billion Oil Overcharges

A Soviet Discovery Rocks World of Mathematics

NY Times 1979. Proved by L. Khachiyan
– based on work by N. Shor, D. Yudin, A. Nemirovski.

Example (Karmarkar's) algorithm:

- n = number of variables
- L = number of bits

$O(n^{3.5}L^2 \log(L) \log(\log(L)))$ operations

Weakly polynomial time

Numerical example: Random matrices

Basis pursuit (e.g., compressed sensing, type of LP):

$$z \in \underset{x}{\operatorname{argmin}} \ \|x\|_1 \quad \text{s.t.} \quad Ax = y.$$

$A \in \mathbb{R}^{1\times N}$ i.i.d. according to prob. dist., $y = A_{1i}$, i unif. in $\{1,2,\dots,N\}$

Solve using $\mathrm{spgl1}$ (state-of-the-art basis pursuit solver).

Numerical example: Random matrices

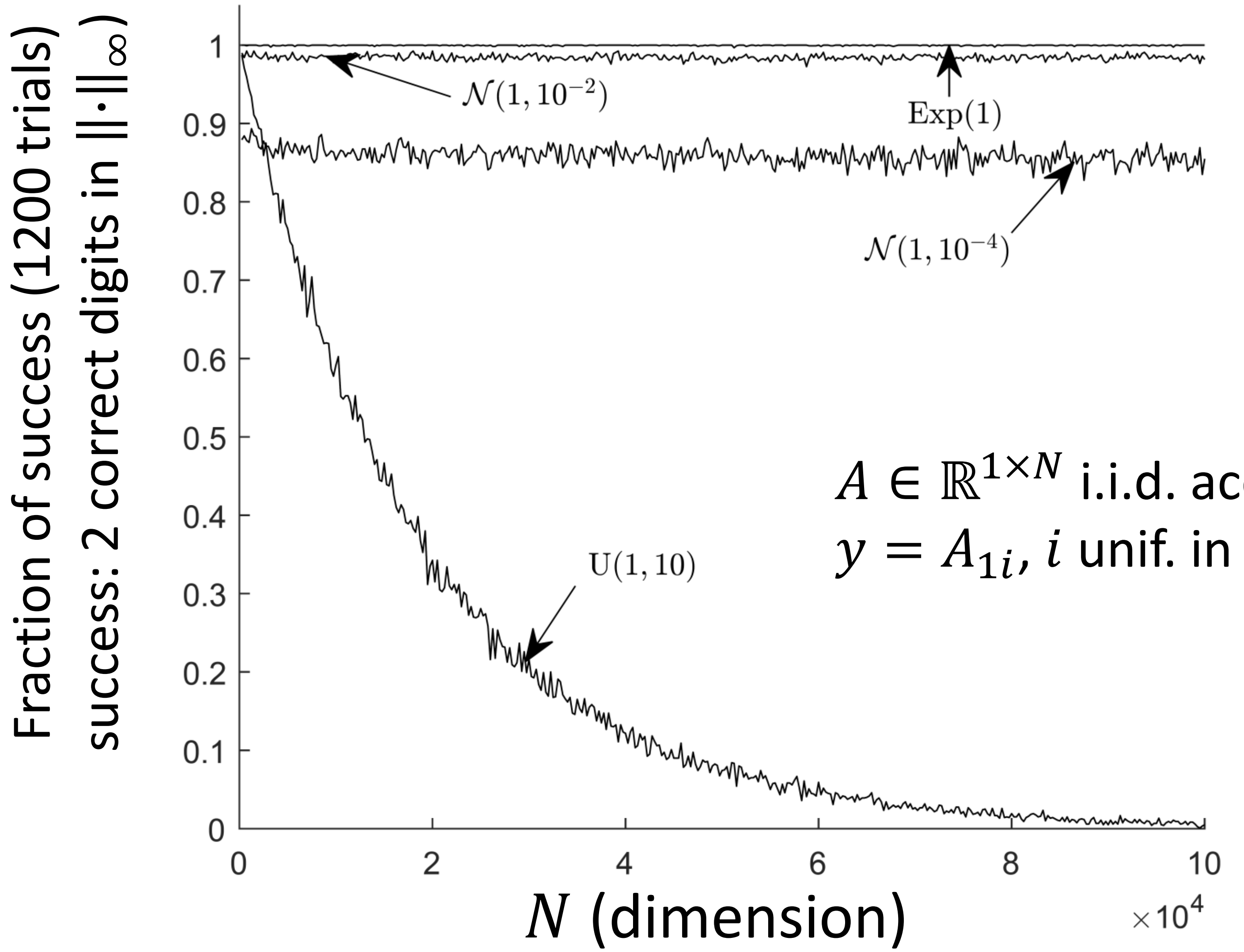
Basis pursuit (e.g., compressed sensing, type of LP):

$$z \in \underset{x}{\operatorname{argmin}} \ \|x\|_1 \quad \text{s.t.} \quad Ax = y.$$

$A \in \mathbb{R}^{1\times N}$ i.i.d. according to prob. dist., $y = A_{1i}$, i unif. in $\{1,2,\dots,N\}$

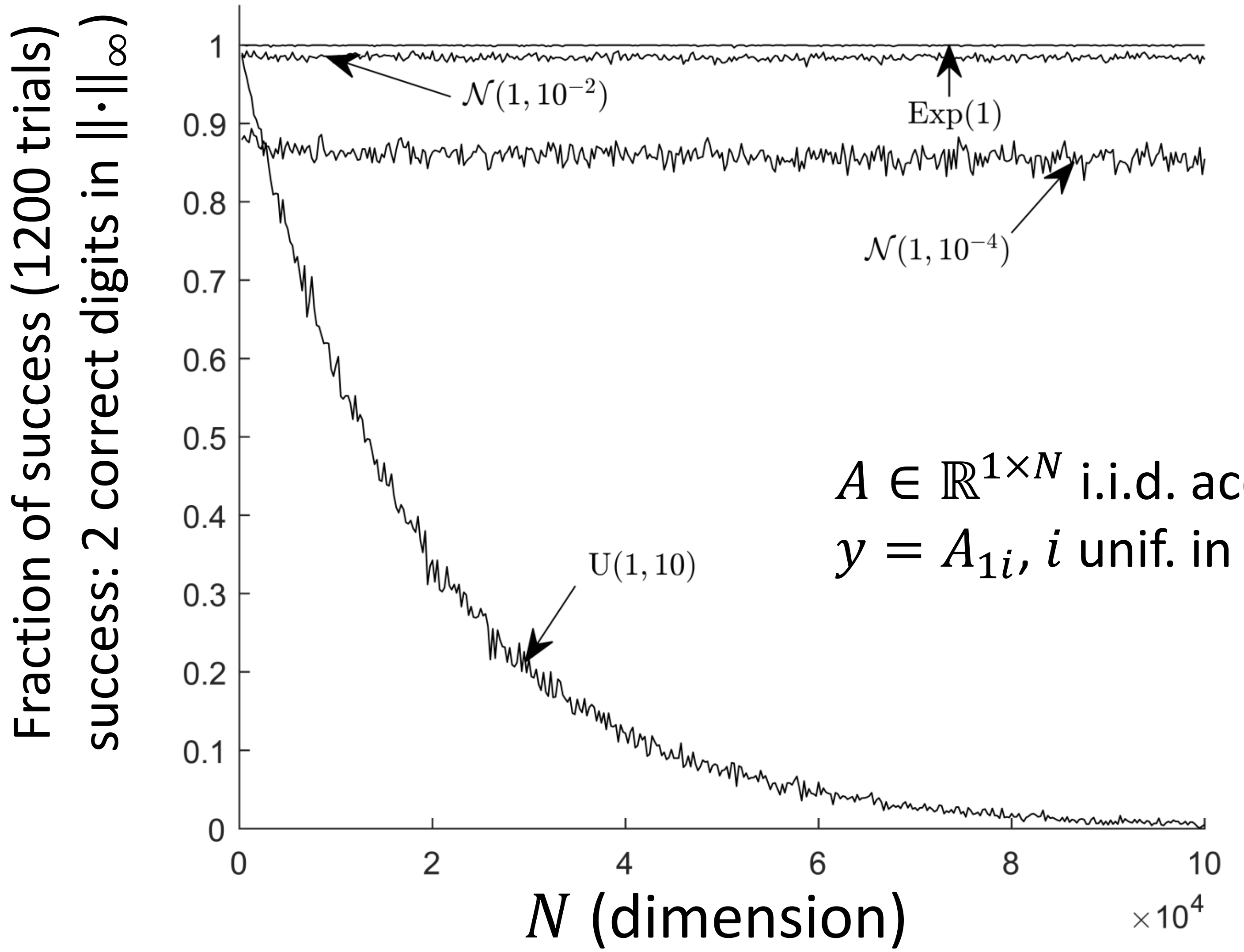
Solve using $\mathrm{spgl1}$ (state-of-the-art basis pursuit solver).

Numerical example: Random matrices



$A \in \mathbb{R}^{1\times N}$ i.i.d. according to prob. dist.,
$y = A_{1i}$, i unif. in $\{1,2,\ldots,N\}$

Numerical example: Random matrices



$A \in \mathbb{R}^{1\times N}$ i.i.d. according to prob. dist.,
$y = A_{1i}$, i unif. in $\{1,2,\dots,N\}$

Numerical example: Innocent LP

$$\underset{x \in \mathbb{R}^2}{\operatorname{argmin}} \; x_1 + x_2 \qquad \text{s.t.} \quad x_1 + (1-\delta)x_2 = 1, \qquad x_j \geq 0$$

MATLAB's linprog has EXITFLAG:

+3: Solution feasible w.r.t. rel. constraint tol., but not abs. tol.

+1: Converged to a solution.

0: Number of iterations or time exceeded maximum.

-2: No feasible point found.

-3: Problem unbounded.

-4: NaN encountered.

-5: Both primal and dual problems are infeasible.

-7: Search direction became too small.

-9: Solver lost feasibility.

Numerical example: Innocent LP

$$\operatorname*{argmin}_{x\in\mathbb{R}^2} \; x_1 + x_2 \qquad \text{s.t.} \quad x_1 + (1-\delta)x_2 = 1, \qquad x_j \geq 0$$

MATLAB's linprog has EXITFLAG:

+3: Solution feasible w.r.t. rel. constraint tol., but not abs. tol.

+1: Converged to a solution.

0: Number of iterations or time exceeded maximum.

-2: No feasible point found.

-3: Problem unbounded.

-4: NaN encountered.

-5: Both primal and dual problems are infeasible.

-7: Search direction became too small.

-9: Solver lost feasibility.

Numerical example: Innocent LP

Algorithms →

	'dual-simplex'		'interior-point'		'interior-point-legacy'	
δ	$\|\hat{x}-x\|_2$	`EXITFLAG`	$\|\hat{x}-x\|_2$	`EXITFLAG`	$\|\hat{x}-x\|_2$	`EXITFLAG`
2^{-1}	0	1	0	1	$6.0 \cdot 10^{-12}$	1
2^{-15}	0	1	0	1	$3.0 \cdot 10^{-5}$	1
2^{-20}	0	1	0	1	$7.0 \cdot 10^{-7}$	1
2^{-24}	0	1	0	1	$7.1 \cdot 10^{-8}$	1
2^{-26}	1.4	1	1.4	1	$1.2 \cdot 10^{-1}$	1
2^{-28}	1.4	1	1.4	1	$4.6 \cdot 10^{-1}$	1
2^{-30}	1.4	1	1.4	1	$7.1 \cdot 10^{-1}$	1

$x =$ computed "solution", $\hat{x} =$ true solution, $\varepsilon_{\text{mach}} = 2^{-52}$

Numerical example: Innocent LP

Algorithms →

δ	'dual-simplex' $\|\hat{x}-x\|_2$	EXITFLAG	'interior-point' $\|\hat{x}-x\|_2$	EXITFLAG	'interior-point-legacy' $\|\hat{x}-x\|_2$	EXITFLAG
2^{-1}	0	1	0	1	$6.0 \cdot 10^{-12}$	1
2^{-15}	0	1	0	1	$3.0 \cdot 10^{-5}$	1
2^{-20}	0	1	0	1	$7.0 \cdot 10^{-7}$	1
2^{-24}	0	1	0	1	$7.1 \cdot 10^{-8}$	1
2^{-26}	1.4	1	1.4	1	$1.2 \cdot 10^{-1}$	1
2^{-28}	1.4	1	1.4	1	$4.6 \cdot 10^{-1}$	1
2^{-30}	1.4	1	1.4	1	$7.1 \cdot 10^{-1}$	1

$x =$ computed "solution", $\hat{x} =$ true solution, $\varepsilon_{\mathrm{mach}} = 2^{-52}$

What went wrong?

What went wrong?

1st issue: Smale's 9th problem*

Recall complexity of Karmarkar's algorithm: $O(n^{3.5}L^2 \log(L) \log(\log(L)))$

Question: What happens with $L = \infty$?

E.g., irrational inputs, computer-assisted proofs etc.

*Steve Smale's list of problems for the 21st century (requested by Vladimir Arnold), inspired by Hilbert's famous list.

1st issue: Smale's 9th problem*

Recall complexity of Karmarkar's algorithm: $O(n^{3.5}L^2 \log(L) \log(\log(L)))$

Question: What happens with $L = \infty$?

E.g., irrational inputs, computer-assisted proofs etc.

Smale's 9th problem: *"Is there a polynomial time algorithm over the real numbers which decides the feasibility of the linear system of inequalities $Ax \geq y$, and if so, outputs such an x?"*

*Steve Smale's list of problems for the 21st century (requested by Vladimir Arnold), inspired by Hilbert's famous list.

1st issue: Smale's 9th problem*

Recall complexity of Karmarkar's algorithm: $O(n^{3.5}L^2\log(L)\log(\log(L)))$

Question: What happens with $L=\infty$?

E.g., irrational inputs, computer-assisted proofs etc.

Smale's 9th problem: *"Is there a polynomial time algorithm over the real numbers which decides the feasibility of the linear system of inequalities $Ax \geq y$, and if so, outputs such an x?"*

*Steve Smale's list of problems for the 21st century (requested by Vladimir Arnold), inspired by Hilbert's famous list.

2nd issue: inexactness

“Real number computations and algorithms which work only in exact arithmetic can offer only limited understanding. Models which process approximate inputs and which permit round-off computations are called for.”

- S. Smale (from the list of mathematical problems for the 21st century)

There will always be numbers which you can't work with exactly!

2nd issue: inexactness

"Real number computations and algorithms which work only in exact arithmetic can offer only limited understanding. Models which process approximate inputs and which permit round-off computations are called for."

- S. Smale (from the list of mathematical problems for the 21st century)

There will always be numbers which you can't work with exactly!

Extended model: Given domain Ω, for any $\iota \in \Omega$ and $k \in \mathbb{N}$, the algorithm can't access ι. Instead, it accesses oracle $\mathcal{O}(\iota, k)$ (cost poly in k) s.t.

$$\|\iota - \mathcal{O}(\iota, k)\|_\infty \leq 2^{-k}$$

2^{nd} issue: inexactness

"Real number computations and algorithms which work only in exact arithmetic can offer only limited understanding. Models which process approximate inputs and which permit round-off computations are called for."

- S. Smale (from the list of mathematical problems for the 21st century)

There will always be numbers which you can't work with exactly!

Extended model: Given domain Ω, for any $\iota \in \Omega$ and $k \in \mathbb{N}$, the algorithm can't access ι. Instead, it accesses oracle $\mathcal{O}(\iota, k)$ (cost poly in k) s.t.

$$\|\iota - \mathcal{O}(\iota, k)\|_\infty \leq 2^{-k}$$

Extended Smale's 9th problem

1st issue: Smale's 9th problem

Polytime alg. for feasibility of linear system of inequalities over $\mathbb{R}$

Extended Smale's 9th problem

1st issue: Smale's 9th problem
Polytime alg. for feasibility of linear system of inequalities over $\mathbb{R}$

+

2nd issue:
Inexactness of input and computations

Extended Smale's 9th problem

1st issue: Smale's 9th problem
Polytime alg. for feasibility of linear system of inequalities over $\mathbb{R}$

+

2nd issue:
Inexactness of input and computations

Extended Smale's 9th problem: *Given a class of LPs that contain irrational numbers (given by an oracle in polynomial time), what is the computational cost of computing a K-digit approximate minimiser? Is the problem solvable in polynomial time in the number of variables?*

Extended Smale's 9th problem

1st issue: Smale's 9th problem
Polytime alg. for feasibility of linear system of inequalities over $\mathbb{R}$

+

2nd issue:
Inexactness of input and computations

Extended Smale's 9th problem: *Given a class of LPs that contain irrational numbers (given by an oracle in polynomial time), what is the computational cost of computing a K-digit approximate minimiser? Is the problem solvable in polynomial time in the number of variables?*

Other common optimization problems

Talked about LPs and basis pursuit. Also have:

- Semidefinite prog.:

$$\underset{X \in \mathbb{S}^n}{\operatorname{argmin}} \langle X, C \rangle \text{ s.t. } \langle X, A_k \rangle = b_k, k = 1, \ldots, m, X \succcurlyeq 0.$$

- Unconstrained Lasso:

$$\underset{x}{\operatorname{argmin}} \ \|Ax - y\|_2^2 + \lambda \mathcal{J}(x), \qquad \mathcal{J}(x) = \|x\|_1 \text{or } \mathrm{TV}(x).$$

- Constrained Lasso:

$$\underset{x}{\operatorname{argmin}} \ \|Ax - y\|_2^2 \text{ s.t. } \mathcal{J}(x) \le \delta, \qquad \mathcal{J}(x) = \|x\|_1 \text{or } \mathrm{TV}(x).$$

$$A \in \mathbb{C}^{m \times N}, \qquad y \in \mathbb{C}^m$$

Other common optimization problems

Talked about LPs and basis pursuit. Also have:

- Semidefinite prog.:

$$\operatorname*{argmin}_{X \in \mathbb{S}^n} \langle X, C \rangle \text{ s.t. } \langle X, A_k \rangle = b_k, k = 1, \dots, m, X \succcurlyeq 0.$$

- Unconstrained Lasso:

$$\operatorname*{argmin}_{x} \ \|Ax - y\|_2^2 + \lambda \mathcal{J}(x), \qquad \mathcal{J}(x) = \|x\|_1 \text{or } \mathrm{TV}(x).$$

- Constrained Lasso:

$$\operatorname*{argmin}_{x} \ \|Ax - y\|_2^2 \ \text{ s.t. } \mathcal{J}(x) \le \delta, \qquad \mathcal{J}(x) = \|x\|_1 \text{or } \mathrm{TV}(x).$$

$$A \in \mathbb{C}^{m \times N}, \qquad y \in \mathbb{C}^m$$

Theorem: For any of prev. problems, integer $K \geq 3$, $\exists$ class $\Omega(K)$ of inputs s.t. simultaneously

NB: Extends to problem of feasibility.

- B., Hansen, Vlačić, "*The extended Smale's 9th problem,*" preprint.

Theorem: For any of prev. problems, integer $K \geq 3$, $\exists$ class $\Omega(K)$ of inputs s.t. simultaneously

1. No (random) alg. can produce $\boldsymbol{K}$ **correct digits** over all of Ω (with prob. $\geq p$, for any $p > 1/2$).

NB: Extends to problem of feasibility.

- B., Hansen, Vlačić, "*The extended Smale's 9th problem,*" preprint.

Theorem: For any of prev. problems, integer $K \geq 3$, $\exists$ class $\Omega(K)$ of inputs s.t. simultaneously

1. No (random) alg. can produce $\boldsymbol{K}$ **correct digits** over all of Ω (with prob. $\geq p$, for any $p > 1/2$).
2. If we allow random alg. with non-zero prob. of not halting, then, for any $p > 2/3$, no alg. produces $\boldsymbol{K}$ **correct digits** over all of Ω with prob. $\geq p$. However, $\exists$ such an alg. for $p = 2/3$.

NB: Extends to problem of feasibility.

- B., Hansen, Vlačić, "*The extended Smale's 9th problem,*" preprint.

Theorem: For any of prev. problems, integer $K \geq 3$, $\exists$ class $\Omega(K)$ of inputs s.t. simultaneously

1. No (random) alg. can produce ***K* correct digits** over all of Ω (with prob. $\geq p$, for any $p > 1/2$).
2. If we allow random alg. with non-zero prob. of not halting, then, for any $p > 2/3$, no alg. produces ***K* correct digits** over all of Ω with prob. $\geq p$. However, $\exists$ such an alg. for $p = 2/3$.
3. (a) $\exists$ alg. that produces $\boldsymbol{K-1}$ **correct digits** over all of Ω. However, for any such alg., $T > 0$, fixed input dim. (m, N), $\exists \iota \in \Omega$ of dim. (m, N) s.t. the runtime on input ι exceeds T.
   (b) For any random alg. Γ, $T > 0$, fixed input dim. (m, N) and $p < 1/2$, $\exists \iota \in \Omega$ of dim. (m, N) s.t.
   $$\mathbb{P}(\Gamma(\iota) \text{ does not have } \boldsymbol{K-1} \textbf{ correct digits} \text{ or run time } > T) > p.$$

NB: Extends to problem of feasibility.

- B., Hansen, Vlačić, *"The extended Smale's 9th problem,"* preprint.

Theorem: For any of prev. problems, integer $K \geq 3$, $\exists$ class $\Omega(K)$ of inputs s.t. simultaneously

1. No (random) alg. can produce $\boldsymbol{K}$ **correct digits** over all of Ω (with prob. $\geq p$, for any $p > 1/2$).
2. If we allow random alg. with non-zero prob. of not halting, then, for any $p > 2/3$, no alg. produces $\boldsymbol{K}$ **correct digits** over all of Ω with prob. $\geq p$. However, $\exists$ such an alg. for $p = 2/3$.
3. (a) $\exists$ alg. that produces $\boldsymbol{K} - \mathbf{1}$ **correct digits** over all of Ω. However, for any such alg., $T > 0$, fixed input dim. (m, N) , $\exists \iota \in \Omega$ of dim. (m, N) s.t. the runtime on input ι exceeds T.
   (b) For any random alg. Γ, $T > 0$, fixed input dim. (m, N) and $p < 1/2$, $\exists \iota \in \Omega$ of dim. (m, N) s.t.
   $$\mathbb{P}(\Gamma(\iota) \text{ does not have } \boldsymbol{K} - \mathbf{1} \textbf{ correct digits} \text{ or run time } > T) > p.$$
4. $\exists$ alg. that produces $\boldsymbol{K} - \mathbf{2}$ **correct digits** over all of Ω s.t. on an input with dim. (m, N) (arbitrary)
   (a) The runtime (and, in the Turing case, the space complexity) is O(poly($m + N$)).
   (b) The number of digits required from the oracle is O(poly(log($m + N$))).

NB: Extends to problem of feasibility.

- B., Hansen, Vlačić, *"The extended Smale's 9th problem,"* preprint.

Theorem: For any of prev. problems, integer $K \geq 3$, $\exists$ class $\Omega(K)$ of inputs s.t. simultaneously

1. No (random) alg. can produce $\boldsymbol{K}$ **correct digits** over all of Ω (with prob. $\geq p$, for any $p > 1/2$).
2. If we allow random alg. with non-zero prob. of not halting, then, for any $p > 2/3$, no alg. produces $\boldsymbol{K}$ **correct digits** over all of Ω with prob. $\geq p$. However, $\exists$ such an alg. for $p = 2/3$.
3. (a) $\exists$ alg. that produces $\boldsymbol{K-1}$ **correct digits** over all of Ω. However, for any such alg., $T > 0$, fixed input dim. (m, N) , $\exists \iota \in \Omega$ of dim. (m, N) s.t. the runtime on input ι exceeds T.
   (b) For any random alg. Γ, $T > 0$, fixed input dim. (m, N) and $p < 1/2$, $\exists \iota \in \Omega$ of dim. (m, N) s.t.
   $$\mathbb{P}(\Gamma(\iota) \text{ does not have } \boldsymbol{K-1} \textbf{ correct digits} \text{ or run time } > T) > p.$$
4. $\exists$ alg. that produces $\boldsymbol{K-2}$ **correct digits** over all of Ω s.t. on an input with dim. (m, N) (arbitrary)
   (a) The runtime (and, in the Turing case, the space complexity) is O(poly($m + N$)).
   (b) The number of digits required from the oracle is O(poly(log($m + N$))).
5. It is impossible to decide if a given alg. fails to produce K correct digits on a given input (with prob. >1/2 in randomised case), even when given an oracle that solves the original problem accurately.
   Hence, producing an $\mathrm{EXITFLAG}$ **is strictly harder than solving the original problem.**

NB: Extends to problem of feasibility.

- B., Hansen, Vlačić, *"The extended Smale's 9th problem,"* preprint.

Theorem: For any of prev. problems, integer $K \geq 3$, $\exists$ class $\Omega(K)$ of inputs s.t. simultaneously

1. No (random) alg. can produce ***K* correct digits** over all of Ω (with prob. $\geq p$, for any $p > 1/2$).
2. If we [allow] random alg. with non-zero prob. of not halting, then, for any $p > 2/3$, no alg. produces ***K* correc[t] digits** over all of Ω with prob. $\geq p$. However, $\exists$ such an alg. for $p = 2/3$.
3. (a) $\exists$ alg. t… **digits** over all of Ω. However, for any such alg., $T > 0$, fixed input dim. … N) s.t. the runtime on input ι exceeds T.
   (b) For any random alg. Γ, $T > 0$, fixed input dim. (m, N) and $p < 1/2$, $\exists \iota$ …
   $\mathbb{P}(\Gamma(\iota)$ does not have $\boldsymbol{K-1}$ **correct digits** or run time > …
4. $\exists$ alg. that produces $\boldsymbol{K-2}$ **correct digits** over all of Ω s.t. on an input with dim. $(m,$ …$)$ (arbitrary)
   (a) The runtime (and, in the Turing case, the space complexity) is O(poly($m + N$)).
   (b) The number of digits required from the oracle is O(poly(log($m + N$))).
5. It is impossible to decide if a given alg. fails to produce K correct digits on a given input (with prob. >1/2 in randomised case), even when given an oracle that solves the original problem accurately.
   Hence, producing an $\mathrm{EXITFLAG}$ is strictly harder than solving the original problem.

1st numerical example
$\Longrightarrow$ happens in practice!

Related to 2nd numerical example

NB: Extends to problem of feasibility.

- B., Hansen, Vlačić, *"The extended Smale's 9th problem,"* preprint.

Theorem: For any of prev. problems, integer $K \geq 3$, $\exists$ class $\Omega(K)$ of inputs s.t. simultaneously

1. No (random) alg. can produce ***K* correct digits** over all of Ω (with prob. $\geq p$, for any $p > 1/2$).
2. If we [...]ow random alg. with non-zero prob. of not halting, then, for any $p > 2/3$, no alg. produces ***K* correc[...] digits** over all of Ω with prob. $\geq p$. However, $\exists$ such an alg. for $p = 2/3$.
3. (a) $\exists$ alg. t[...] **digits** over all of Ω. However, for any such alg., $T > 0$, fixed input dim. [...]$N)$ s.t. the runtime on input ι exceeds T.
   (b) For any random alg. Γ, $T > 0$, fixed input dim. (m, N) and $p < 1/2$, $\exists\iota$ [...]

   $$\mathbb{P}(\Gamma(\iota) \text{ does not have } \boldsymbol{K-1} \textbf{ correct digits} \text{ or run time} > [...]$$
4. $\exists$ alg. that produces $\boldsymbol{K-2}$ **correct digits** over all of Ω s.t. on an input with dim. $(m,$ [...]$)$ (arbitrary)
   (a) The runtime (and, in the Turing case, the space complexity) is O(poly($m + N$)).
   (b) The number of digits required from the oracle is O(poly(log($m + N$))).
5. It is impossible to decide if a given alg. fails to produce K correct digits on a given input (with prob. >1/2 in randomised case), even when given an oracle that solves the original problem accurately.
   Hence, producing an $\mathrm{EXITFLAG}$ **is strictly harder than solving the original problem.**

1st numerical example
$\Longrightarrow$ happens in practice!

Related to 2nd
numerical example

NB: Extends to problem of feasibility.

- B., Hansen, Vlačić, *"The extended Smale's 9th problem,"* preprint.

Breakdown epsilons (computational limitations)

Strong breakdown-epsilon ($\mathrm{e.g.}, 10^{-K}$ in prev. thm.):

$$\varepsilon_B^s = \sup\left\{\varepsilon \geq 0 : \forall \text{ alg. } \Gamma \ \exists \iota \in \Omega \quad \text{s.t.} \quad \operatorname{dist}\big(\Gamma(\iota), \Xi(\iota)\big) > \varepsilon\right\}$$

Weak breakdown-epsilon ($\mathrm{e.g.}, 10^{-(K-1)}$ in prev. thm.):

$$\varepsilon_B^w = \sup\left\{\begin{array}{r} \varepsilon \geq 0 : \forall \text{ alg. } \Gamma, \forall T > 0, \exists \iota \in \Omega \quad \text{s.t.} \quad \operatorname{dist}\big(\Gamma(\iota), \Xi(\iota)\big) > \varepsilon \\ \text{or runtime}(\Gamma, \iota) > T \end{array}\right\}$$

Breakdown epsilons (computational limitations)

Strong breakdown-epsilon ($\mathrm{e.g.}, 10^{-K}$ in prev. thm.):

$$\varepsilon_B^s = \sup\left\{\varepsilon \geq 0 : \forall \text{ alg. } \Gamma \; \exists \iota \in \Omega \quad \text{s.t.} \quad \mathrm{dist}\big(\Gamma(\iota), \Xi(\iota)\big) > \varepsilon\right\}$$

Weak breakdown-epsilon ($\mathrm{e.g.}, 10^{-(K-1)}$ in prev. thm.):

$$\varepsilon_B^w = \sup\left\{\begin{aligned} \varepsilon \geq 0 : \forall \text{ alg. } \Gamma, \forall T > 0, \exists \iota \in \Omega \quad \text{s.t.} \quad \mathrm{dist}\big(\Gamma(\iota), \Xi(\iota)\big) > \varepsilon \\ \text{or runtime}(\Gamma, \iota) > T \end{aligned}\right\}$$

How much accuracy do we need in a computer assisted proof?

Is the breakdown epsilon for the problem below that threshold?

A related story: Hardness of approximation

Given $\iota \in \Omega \subseteq \mathbb{R}^n$, have feasible set $F(\iota)$ & objective fun. f_ι.

Compute: $\mathrm{OPT}(\iota) = \min\limits_{x \in F(\iota)} f_\iota(x)$

A related story: Hardness of approximation

Given $\iota \in \Omega \subseteq \mathbb{R}^n$, have feasible set $F(\iota)$ & objective fun. f_ι.

Compute: $\mathrm{OPT}(\iota) = \min_{x \in F(\iota)} f_\iota(x)$

Does there exist alg. Γ s.t. $\forall \iota \in \Omega$, $\Gamma(\iota) \in F(\iota)$ and

$$f_\iota\big(\Gamma(\iota)\big) \leq (1+\varepsilon)\mathrm{OPT}(\iota), \qquad \mathrm{runtime}(\Gamma, \iota) = O(\mathrm{poly}(n))$$

$\Gamma(\iota)$: ε-approx. solution

A related story: Hardness of approximation

Given $\iota \in \Omega \subseteq \mathbb{R}^n$, have feasible set $F(\iota)$ & objective fun. f_ι.

$$\textbf{Compute: } \mathrm{OPT}(\iota) = \min_{x \in F(\iota)} f_\iota(x)$$

Does there exist alg. Γ s.t. $\forall \iota \in \Omega$, $\Gamma(\iota) \in F(\iota)$ and

$$\left.\begin{array}{l} \\ f_\iota\big(\Gamma(\iota)\big) \le (1+\varepsilon)\mathrm{OPT}(\iota), \qquad \mathrm{runtime}(\Gamma, \iota) = O(\mathrm{poly}(n)) \end{array}\right\}$$

$\Gamma(\iota)$: ε-approx. solution

Typically, for combinatorial prob., $\exists \varepsilon_A > 0$ s.t. computing ε-approx. solution is

(a) $\in$ P if $\varepsilon_A < \varepsilon$ and (b) NP-hard if $\varepsilon < \varepsilon_A$.

A related story: Hardness of approximation

Given $\iota \in \Omega \subseteq \mathbb{R}^n$, have feasible set $F(\iota)$ & objective fun. f_ι.

$\Gamma(\iota)$: ε-approx. solution

Compute: $\mathrm{OPT}(\iota) = \min_{x \in F(\iota)} f_\iota(x)$

Does there exist alg. Γ s.t. $\forall \iota \in \Omega$, $\Gamma(\iota) \in F(\iota)$ and

$$f_\iota\big(\Gamma(\iota)\big) \leq (1+\varepsilon)\mathrm{OPT}(\iota), \qquad \mathrm{runtime}(\Gamma, \iota) = O(\mathrm{poly}(n))$$

Typically, for combinatorial prob., $\exists \varepsilon_A > 0$ s.t. computing ε-approx. solution is

(a) $\in$ P if $\varepsilon_A < \varepsilon$ and (b) NP-hard if $\varepsilon < \varepsilon_A$.

Phase transitions

Hardness of Approximation

PCP Theorem* often leads to threshold $\varepsilon_A > 0$.
Assuming P ≠ NP, often have phase transition:

$\varepsilon_A < \varepsilon$
Comp. ε-approx. ∈ P

↕

$\varepsilon < \varepsilon_A$
Comp. ε-approx. is NP-hard (thus ∉ P)

*2001 Gödel Prize awarded to S. Arora, U. Feige, S. Goldwasser, C. Lund, L. Lovász, R. Motwani, S. Safra, M. Sudan, and M. Szegedy for work on the PCP theorem and its connection to hardness of approximation.

Phase transitions

Hardness of Approximation

PCP Theorem* often leads to threshold $\varepsilon_A > 0$.
Assuming P $\neq$ NP, often have phase transition:

$\varepsilon_A < \varepsilon$
Comp. ε-approx. $\in$ P

$\varepsilon < \varepsilon_A$
Comp. ε-approx. is NP-hard (thus $\notin$ P)

*2001 Gödel Prize awarded to S. Arora, U. Feige, S. Goldwasser, C. Lund, L. Lovász, R. Motwani, S. Safra, M. Sudan, and M. Szegedy for work on the PCP theorem and its connection to hardness of approximation.

Generalised Hardness of Approximation (GHA)

Extended Smale's 9th problem shows:

$\varepsilon_B^w < \varepsilon$
Comp. ε-approx. $\in$ P

$\varepsilon_B^s < \varepsilon < \varepsilon_B^w$
Comp. ε-approx. $\notin$ k-EXPTIME $\forall k$
Possible to comp. ε-approx.

$\varepsilon < \varepsilon_B^s$
Impossible to comp. ε-approx.

$\varepsilon_B^s = \varepsilon_B^w$

Holds even if P $=$ NP!

The SCI hierarchy for computer-assisted proofs

Arbitrary prec.
or decision problems

$$\begin{array}{ccccc}
 & & \vdots & & \\
\Sigma_2 & \neq & \Sigma_2 \cup \Pi_2 & \neq & \Pi_2 \\
\rotatebox{45}{$\subsetneq$} & & \rotatebox{90}{$\subsetneq$} & & \rotatebox{135}{$\subsetneq$} \\
 & & \Delta_2 & & \\
\rotatebox{135}{$\subsetneq$} & & \rotatebox{90}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} \\
\Sigma_1 & \neq & \Sigma_1 \cup \Pi_1 & \neq & \Pi_1 \\
\rotatebox{45}{$\subsetneq$} & & \rotatebox{90}{$\subsetneq$} & & \rotatebox{135}{$\subsetneq$} \\
 & & \Delta_1 & & \\
 & & \rotatebox{90}{$\subsetneq$} & & \\
 & & \Pi_0 = \Delta_0 = \Sigma_0 & &
\end{array}$$

$\varepsilon > 0$ precision

$\varepsilon_B^w < \varepsilon$

Comp. ε-approx. $\in$ P

$\varepsilon_B^s < \varepsilon < \varepsilon_B^w$

Comp. ε-approx. $\notin$ k-EXPTIME $\forall k$

Possible to comp. ε-approx.

$\varepsilon < \varepsilon_B^s$

Impossible to comp. ε-approx.

$\varepsilon_B^s = \varepsilon_B^w$

The SCI hierarchy for computer-assisted proofs

Arbitrary prec. or decision problems

$$\vdots$$

$$\Sigma_2 \neq \Sigma_2 \cup \Pi_2 \neq \Pi_2$$

$$\Delta_2$$

$$\Sigma_1 \neq \Sigma_1$$

$$\Pi_0 = \Delta_0 = \Sigma_0$$

$\varepsilon > 0$ precision

$\in$ P

$< \varepsilon < \varepsilon_B^w$

...p. ε-approx. $\notin$ k-EXPTIME $\forall k$

Possible to comp. ε-approx.

$\varepsilon < \varepsilon_B^s$

Impossible to comp. ε-approx.

$\varepsilon_B^s = \varepsilon_B^w$

Can we build the "SCI-library" for computer-assisted proofs?

Further examples and questions I

- The SCI hierarchy appears throughout computational mathematics.
- **Results so far:** spectral theory, PDEs and ODEs, iterative rational maps, generalized Collatz problem (and dec. problems), topology, inverse problems, optimization, AI, …
- *Can we classify which PDEs allow $\Sigma_1 \cup \Pi_1$ verification of blow-up?*
- Nonlinear dyn. systems can be studied through transfer operators. *Can we develop a foundations theory for their spectral properties? Can this be done by simply observing the dynamical system?*

- McMullen, *"Families of rational maps and iterative root-finding algorithms,"* **Ann. of Math.**, 1987.
- Doyle, McMullen, *"Solving the quintic by iteration,"* **Acta Math.**, 1989.
- Smale, *"The fundamental theorem of algebra and complexity theory,"* **Bull. Amer. Math. Soc.**, 1981.
- B., Hansen, Vlačić, *"The extended Smale's 9th problem,"* preprint.
- Hansen, Becker, *"Computing solutions of Schrödinger equations on unbounded domains,"* preprint.
- C., *"Computing semigroups with error control,"* **SIAM J. Numer. Anal.**, 2022.
- Koopman, v. Neumann, *"Dynamical systems of continuous spectra,"* **Proc. Natl. Acad. Sci. USA**., 1932.
- C., Townsend, *"Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,"* preprint.

Further examples and questions II

- GHA appears in learning problems (adaptive/probabilistic training data):

Smale's 18th problem: ***What are the limits of artificial intelligence?***

Image reconstruction and inverse problems where:

  - Stable and accurate neural networks **exist** **BUT…**
  - Whether we can **train** them depends on ε, amount of training data, stability requirements.

- *Can we develop a theory for characterisations of phase transitions in GHA?*
  NB: Specific cases known for compressed sensing etc.
- *What other types of problems have this phenomena?*
- Re computer-assisted proofs: *For which problems can we compute exit flags?*

- C., Antun, Hansen, *"The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem,"* **Proc. Natl. Acad. Sci. USA**.

Summary

Classifications $\Longrightarrow$ assumptions needed for use in a computer-assisted proof.

SCI hierarchy: a tool that allows us to

- Classify the difficulty of both continuous and discrete computational problems.
- Prove that algorithms are optimal and find them.
- Σ_1 and Π_1 classifications allow verified computations.

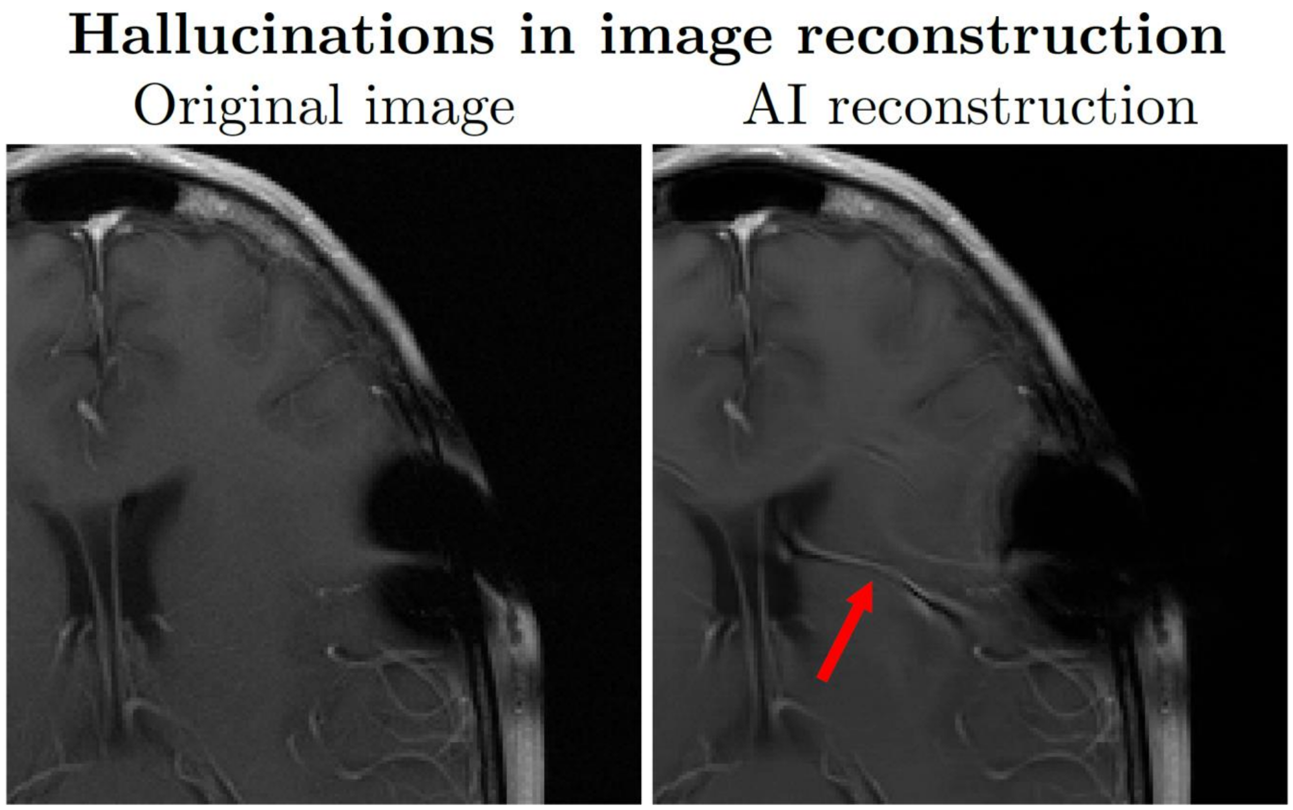
GHA: phase transitions (for inexact input) depending on the accuracy goal ε

- $\varepsilon_B^w < \varepsilon$: $\in P$.
- $\varepsilon_B^s < \varepsilon < \varepsilon_B^w$: computable, but $\notin$ k−EXPTIME $\forall k$.
- $\varepsilon < \varepsilon_B^s$: impossible to compute.

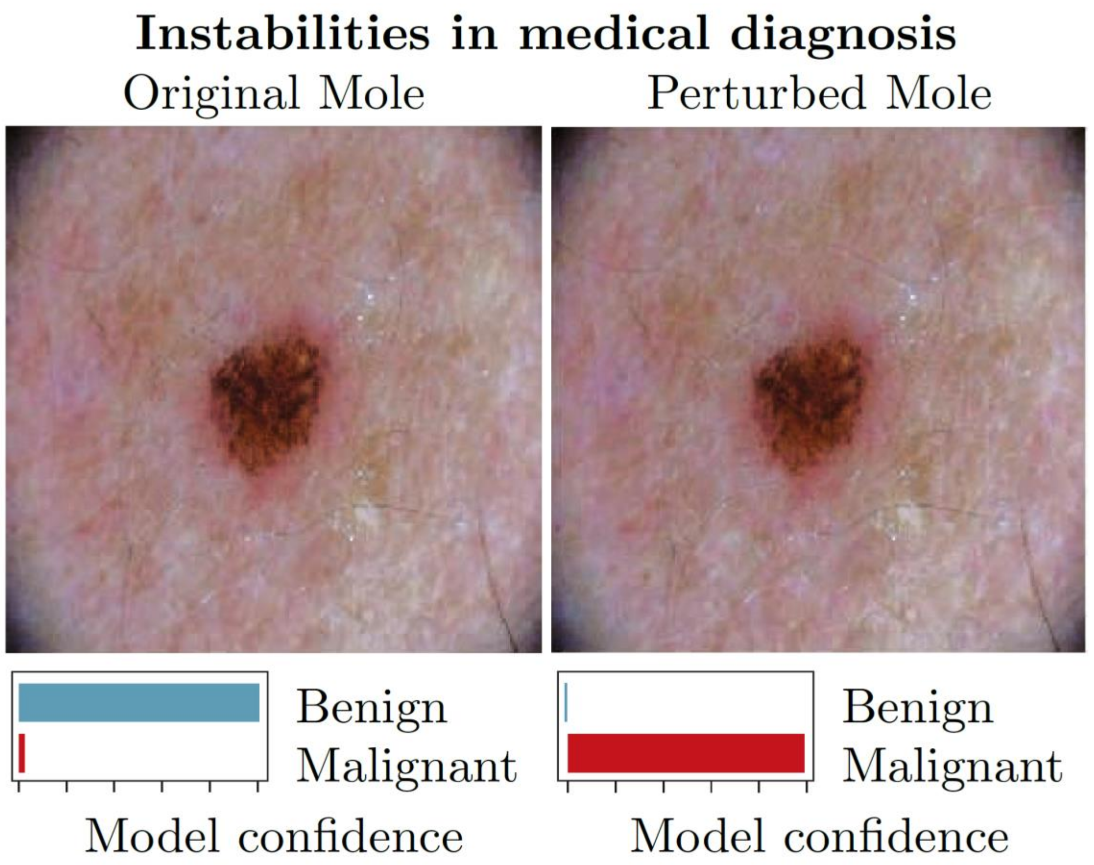
abastoun@ed.ac.uk m.colbrook@damtp.cam.ac.uk

Additional slides

Problem: hallucinations and instability



"**AI generated hallucination**", from Facebook and NYU's ***FastMRI challenge*** 2020



From Finlayson et al., "*Adversarial attacks on medical machine learning*," **Science**, 2019.

When can we make AI robust and trustworthy?

Example of the limits of deep learning

Paradox: "Nice" linear inverse problems where a *stable* and *accurate* neural network for image reconstruction <u>exists</u>, but it <u>can never be trained</u>!

E.g., suppose we want to solve (holds for much more general problems)

$$(P_1) \quad \operatorname{argmin}_{x \in \mathbb{C}^N} F_1^A(x) := \|x\|_{l_w^1}, \text{ such that } \|Ax - y\|_{l^2} \leq \epsilon,$$
$$(P_2) \quad \operatorname{argmin}_{x \in \mathbb{C}^N} F_2^A(x, y, \lambda) := \lambda \|x\|_{l_w^1} + \|Ax - y\|_{l^2}^2,$$
$$(P_3) \quad \operatorname{argmin}_{x \in \mathbb{C}^N} F_3^A(x, y, \lambda) := \lambda \|x\|_{l_w^1} + \|Ax - y\|_{l^2}.$$

$$A \in \mathbb{C}^{m \times N} \text{ (modality, } m < N), \qquad S = \{y_j\}_{j=1}^R \text{ (samples)}$$

Arises when given $y \approx Ax + e$.

Arbitrary precision of training data

In practice, A not known exactly or cannot be stored to infinite precision.

Assume access to: $\{y_{k,n}\}_{k=1}^{R}$ and A_n (rational approximations, e.g., floats) such that

$$\|y_{k,n} - y_k\| \leq 2^{-n}, \quad \|A_n - A\| \leq 2^{-n}, \quad \forall n \in \mathbb{N}.$$

Training set for $(A, \mathcal{S}) \in \Omega$:

$$\iota_{A,\mathcal{S}} := \{(y_{k,n}, A_n) \mid k = 1, \ldots, R \text{ and } n \in \mathbb{N}\}.$$

In a nutshell: allow access to arbitrary precision training data.

Question: Given a collection Ω of $(A, \mathcal{S})$, does there <u>exist</u> a neural network approximating Ξ (solution map of (P_j)), and <u>can it be trained</u> by an algorithm?

Condition numbers

Given $\Omega \subseteq \mathbb{C}^n$, define

$$\mathrm{Act}(\Omega) = \left\{j : \exists x, y \in \Omega, x_j \neq y_j\right\}, \qquad \Omega^{\mathrm{Act}} = \left\{x : \exists y \in \Omega, x_{\mathrm{Act}(\Omega)^c} = y_{\mathrm{Act}(\Omega)^c}\right\}$$

- Condition of a mapping $\Xi : \widehat{\Omega} \rightrightarrows \mathbb{C}^m$ with $\Omega \subseteq \widehat{\Omega}$:

$$\mathrm{Cond}(\Xi, \Omega) = \sup_{x \in \Omega} \lim_{\varepsilon \to 0_+} \sup_{\substack{x+z \in \Omega^{\mathrm{Act}} \cap \widehat{\Omega} \\ 0 < \|z\|_\infty < \varepsilon}} \frac{\mathrm{dist}(\Xi(x+z), \Xi(x))}{\|z\|_\infty}$$

- For problems with constraints (e.g., basis pursuit P_1 or LPs)

$$\nu(A, y) = \inf\left\{\varepsilon \geq 0 : \|\hat{y} - y\|_2, \left\|\hat{A} - A\right\| \leq \varepsilon, \left(\hat{A}, \hat{y}\right) \in \Omega^{\mathrm{Act}} \text{ and infeasible}\right\}$$

$$C_{\mathrm{FP}}(A, y) = \frac{\max\{\|y\|_2, \|A\|\}}{\nu(A, y)}$$

- Renegar condition number

$$\mu(A, y) = \inf\left\{\varepsilon \geq 0 : \|\hat{y} - y\|_2, \left\|\hat{A} - A\right\| \leq \varepsilon, \left(\hat{A}, \hat{y}\right) \in \Omega^{\mathrm{Act}}, \Xi \text{ multivalued}\right\}$$

$$C_{\mathrm{RCC}}(A, y) = \frac{\max\{\|y\|_2, \|A\|\}}{\mu(A, y)}$$

Theorem: For any of prev. problems, integer $K \geq 3$ and $L \in \mathbb{N}$, $\exists$ a well-conditioned class $\Omega(K)$ of inputs s.t. simultaneously

1. No deterministic alg. can, given a training set $\iota_{A,S} \in \Omega_{\mathcal{T}}$, produce a neural network (NN) ϕ with
$$(1) \qquad \min_{y \in S} \inf_{x^* \in \Xi(A,y)} \|\phi(y) - x^*\|_2 \leq 10^{-K} \qquad \forall (A,S) \in \Omega(K).$$
For any $p > 1/2$, no random alg. (any model of comp.) can produce a NN ϕ s.t. (1) holds with prob. $\geq p$.

2. (a) $\exists$ deterministic alg. that , given a training set $\iota_{A,S} \in \Omega_{\mathcal{T}}$, produces a neural network (NN) ϕ with
$$(2) \qquad \max_{y \in S} \inf_{x^* \in \Xi(A,y)} \|\phi(y) - x^*\|_2 \leq 10^{-(K-1)} \qquad \forall (A,S) \in \Omega(K).$$
(b) However, for any probabilistic Turing Machine that produces such a NN, any $M \in \mathbb{N}$ and $p \in \left[0, \frac{N-m}{N+1-m}\right)$, there exists a training set $\iota_{A,S} \in \Omega_{\mathcal{T}}$ s.t. $\forall y \in S$
$$\mathbb{P}\left(\inf_{x^* \in \Xi(A,y)} \|\phi(y) - x^*\|_2 > 10^{-(K-1)} \text{ or size of training data to construct } \phi \text{ exceeds } M\right) > p.$$

3. $\exists$ deterministic alg. that, given a training set $\iota_{A,S} \in \Omega_{\mathcal{T}}$, produces a NN ϕ accessing at most L training samples of $\iota_{A,S}$ s.t.
$$(3) \qquad \max_{y \in S} \inf_{x^* \in \Xi(A,y)} \|\phi(y) - x^*\|_2 \leq 10^{-(K-2)} \qquad \forall (A,S) \in \Omega(K).$$

- C., Antun, Hansen, *"The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem,"* **Proc. Natl. Acad. Sci. USA**.

Theorem: For any of prev. problems, integer $K \geq 3$ and $L \in \mathbb{N}$, $\exists$ a well-conditioned class $\Omega(K)$ of inputs s.t. simultaneously

1. No deterministic alg. can, given a training set $\iota_{A,S} \in \Omega_{\mathcal{T}}$, produce a neural network (NN) ϕ with

$$(1) \qquad \min_{y \in S} \inf_{x^* \in \Xi(A,y)} \|\phi(y) - x^*\|_2 \leq 10^{-K} \qquad \forall (A,S) \in \Omega(K).$$

For any $p > 1/2$, no random alg. (any model of comp.) can produce a NN ϕ s.t. (1) holds with prob. $\geq p$.

2. (a) $\exists$ deterministic alg. that , given a training set $\iota_{A,S} \in \Omega_{\mathcal{T}}$, produces a neural network (NN) ϕ with

$$(2) \qquad \max_{y \in S} \inf_{x^* \in \Xi(A,y)} \|\phi(y) - x^*\|_2 \leq 10^{-(K-1)} \qquad \forall (A,S) \in \Omega(K).$$

Holds for any architecture, any precision of training data.

$\Longrightarrow$ Classification theory telling us what can and cannot be done

$$\mathbb{P}\left(\inf_{x^* \in \Xi(A,y)} \|\phi(y) - x^*\|_2 > 10^{-(K-1)} \text{ or size of training data to construct } \phi \text{ exceeds } M \right) > p.$$

3. $\exists$ deterministic alg. that, given a training set $\iota_{A,S} \in \Omega_{\mathcal{T}}$, produces a NN ϕ accessing at most L training samples of $\iota_{A,S}$ s.t.

$$(3) \qquad \max_{y \in S} \inf_{x^* \in \Xi(A,y)} \|\phi(y) - x^*\|_2 \leq 10^{-(K-2)} \qquad \forall (A,S) \in \Omega(K).$$

- C., Antun, Hansen, *"The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem,"* **Proc. Natl. Acad. Sci. USA**.

Feasibility problem

Given $K \in \mathbb{N}$, $M \in \mathbb{R}$, $A \in \mathbb{R}^{m \times N}$, and $y \in \mathbb{R}^m$ does $\exists x \in \mathbb{R}^N$ s.t.

$$\langle x, c \rangle_K = \lfloor 10^K \langle x, c \rangle \rfloor 10^{-K} \leq M \quad \text{and} \quad Ax = y, x \geq 0?$$

$$\Xi_K(A, y) = \begin{cases} \text{Yes,} & \text{if such an } x \text{ exists} \\ \text{No,} & \text{otherwise.} \end{cases}$$

Theorem: There are infinitely many $M \in \mathbb{R}$ such that the following happens: for any integer $K \geq 3$, $\exists$ class $\Omega(K)$ of inputs s.t. simultaneously

1. $\nexists$ sequence of alg. $\{\Gamma_n\}$ s.t. $\Gamma_n \uparrow \Xi_K$ on Ω. I.e.,$\{\Xi_K, \Omega\} \notin \Sigma_1$.
2. No random alg. exists that solves Ξ_K on all inputs with probability exceeding 1/2.
3. (a) $\exists$ alg. that solves Ξ_{K-1}. However, for any such alg., $T > 0$, fixed input dim. (m, N) , $\exists\iota \in \Omega$ of dim. (m, N) s.t. the runtime on input ι exceeds T.

   (b) For any random alg. Γ, $T > 0$, fixed input dim. (m, N) and $p < 1/2$, $\exists\iota \in \Omega$ of dim. (m, N) s.t.

$$\mathbb{P}(\Gamma(\iota) \neq \Xi_{K-1}(\iota) \text{or run time} > T) > p.$$

4. $\exists$ alg. that solves Ξ_{K-2} over all of Ω s.t. on an input with dim. (m, N) (arbitrary)

   (a) The runtime (and, in the Turing case, the space complexity) is O(poly($m + N$)).

   (b) The number of digits required from the oracle is O(poly(log($m + N$))).

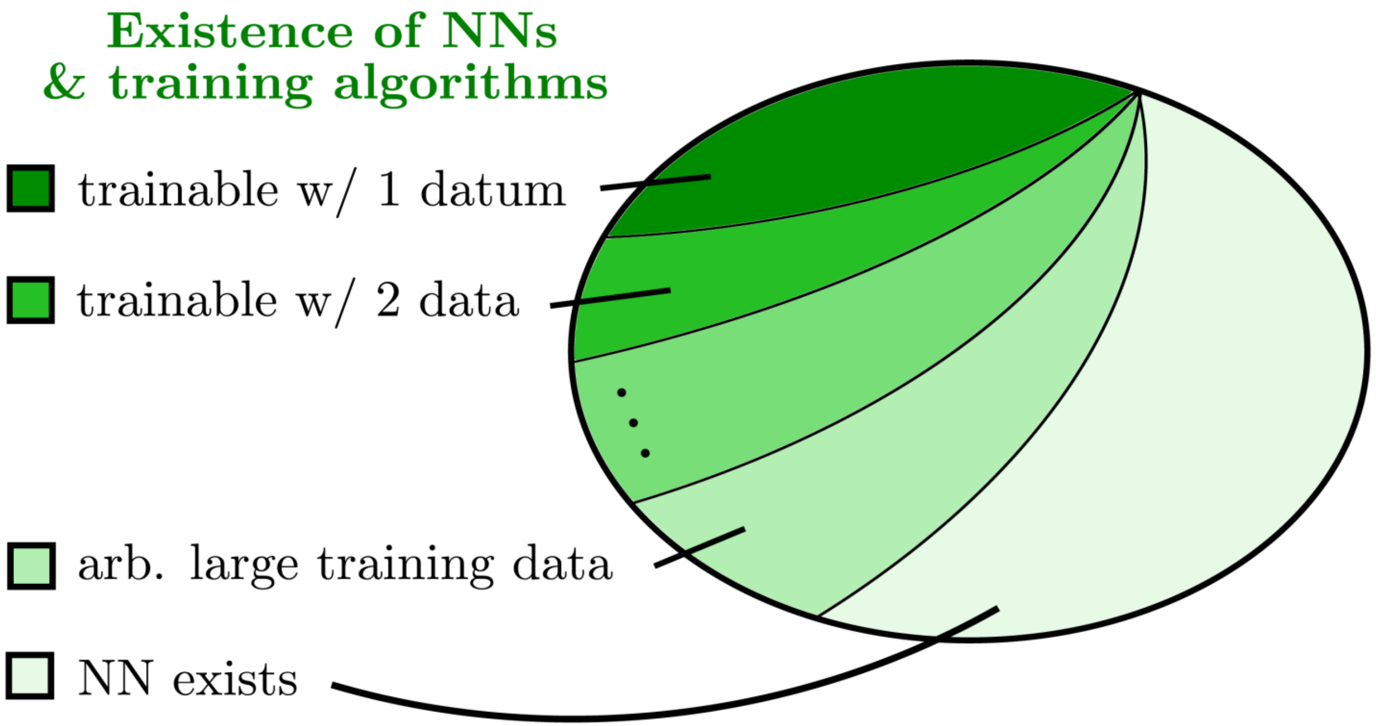
- B., Hansen, Vlačić, "*The extended Smale's 9th problem,*" preprint.

NB: There are classes Ω s.t. infinitely many $\iota \in \Omega$ have $C_{RCC} = \infty$.
However, $\exists$ alg. that solves LP/basis pursuit to K digits on Ω s.t.

(a) Runtime (and, in Turing case, space complexity) is O(poly($m + N + K$)).

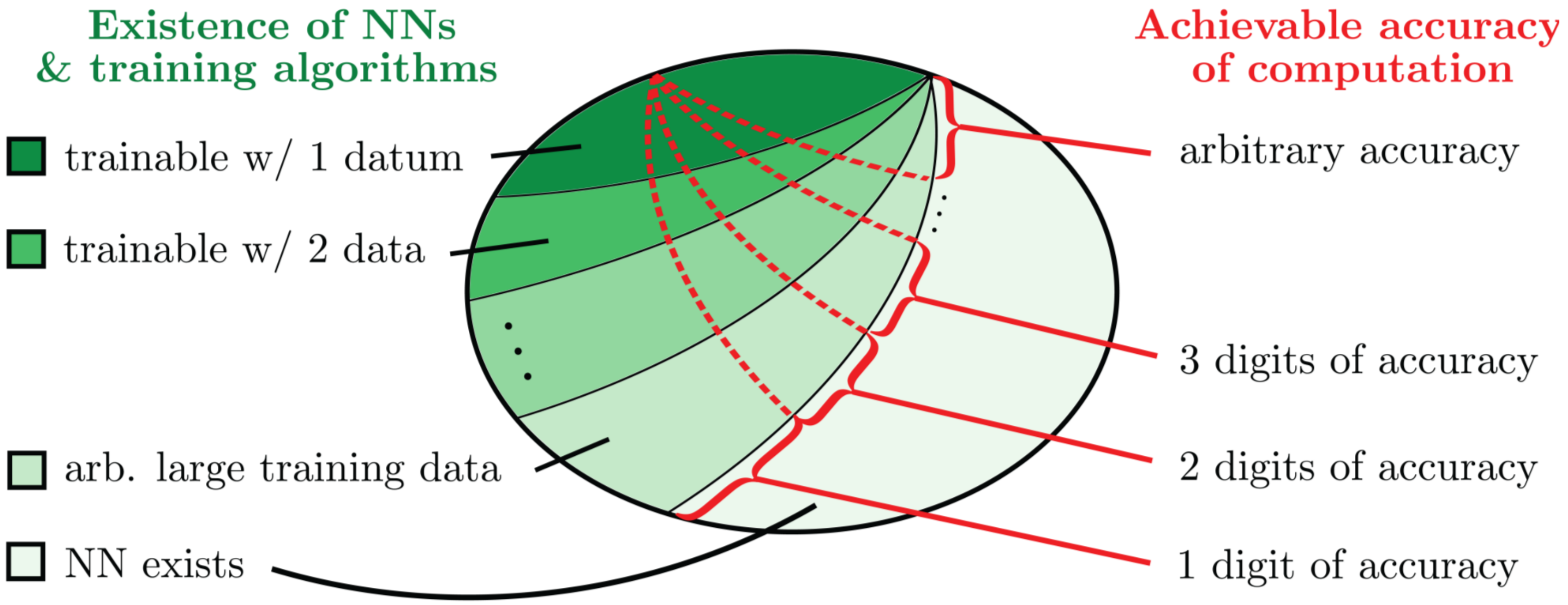
(b) Number of digits required from oracle is O(poly(log($m + N$ + K))).

The world of neural networks



Given a problem and conditions, where does it sit in this diagram?

The world of neural networks



Given a problem and conditions, where does it sit in this diagram?

Example counterpart theorem

Certain conditions: stable neural networks trained with exponential accuracy.
E.g., *approximate Łojasiewicz-type inequality*:

$$(1)\quad \min_{x\in\mathbb{C}^N} f(x) \quad \text{s.t.} \quad \|Ax-y\| \le \varepsilon$$

$$\mathrm{dist}(x, \mathrm{solution}) \le \alpha([f(x)-f^*] + [\|Ax-y\| - \varepsilon] + \delta)$$

Fast **I**terative **RE**started **NET**works (FIRENETs)
(unrolled primal-dual with novel restart scheme)

Theorem: Training algorithm that, under above assumption, produces *stable* neural networks φ_n of width $O(N)$, depth $O(n)$, guaranteed worst bound

$$\mathrm{dist}(\varphi_n(y), \mathrm{solution}) \lesssim e^{-n} + \delta$$

- C., Antun, Hansen, "*The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem,*" **PNAS**, 2022.
- C., "*WARPd: A linearly convergent first-order method for inverse problems with approximate sharpness conditions,*" **SIAM J. Imaging Sci.**, 2022.

Numerical example of GHA

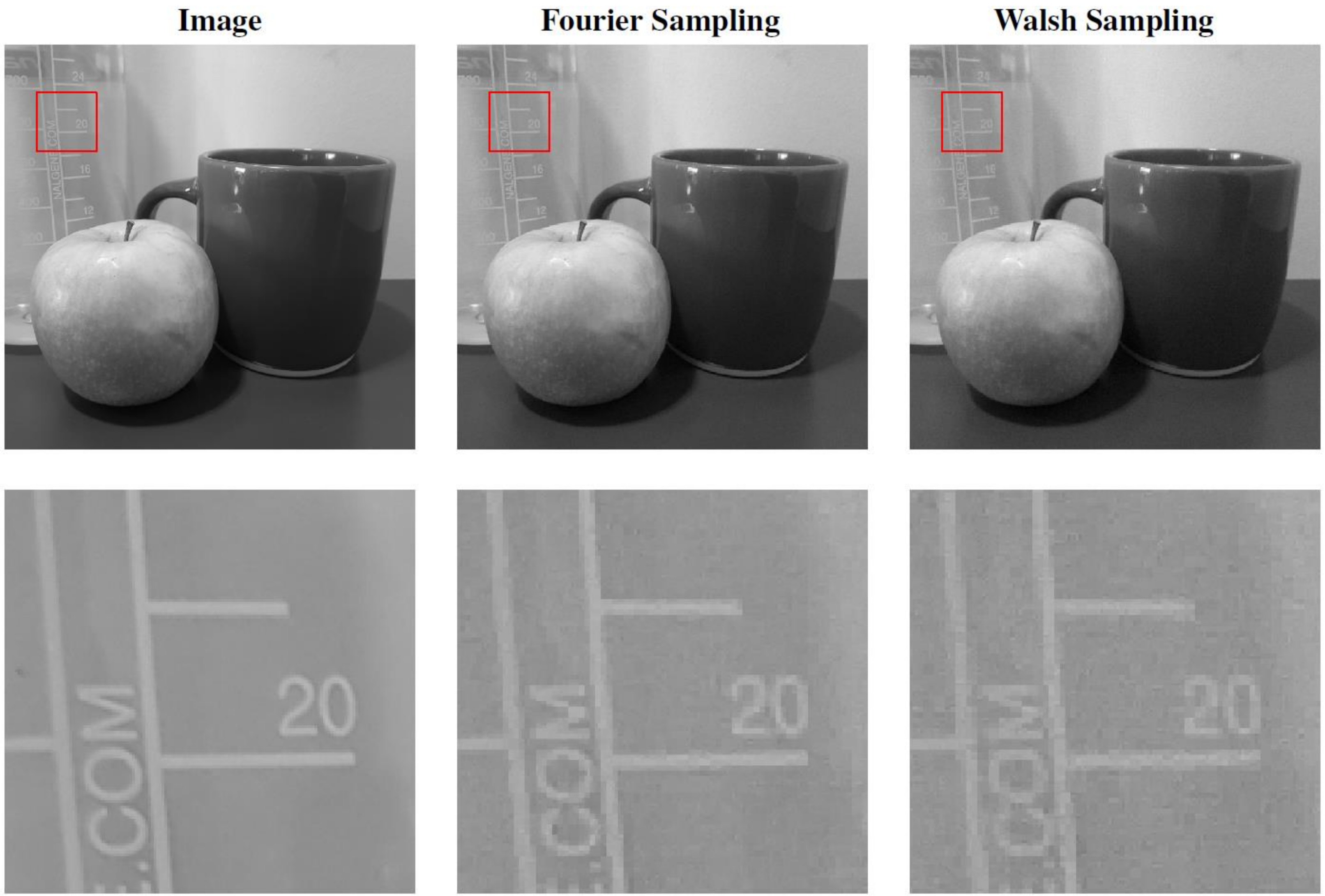
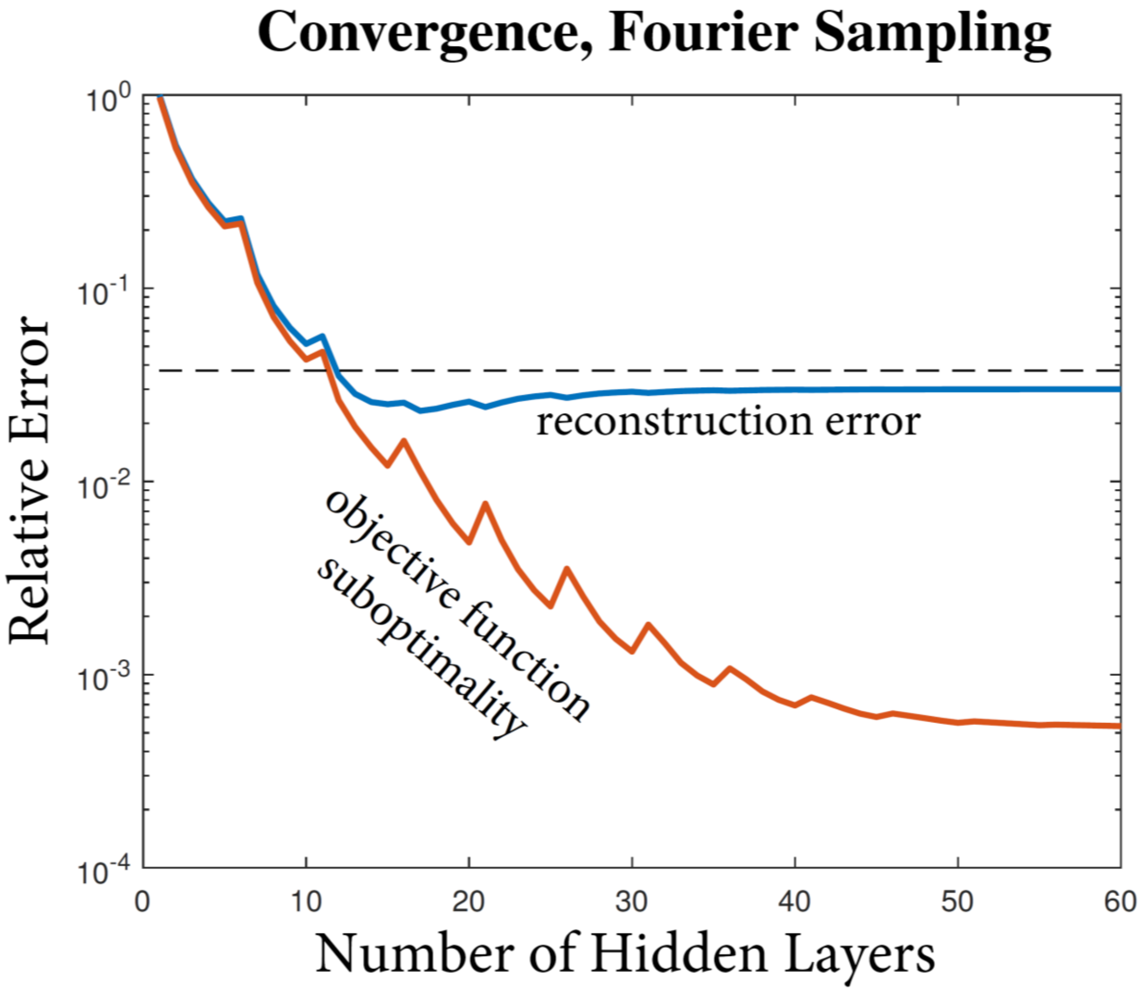
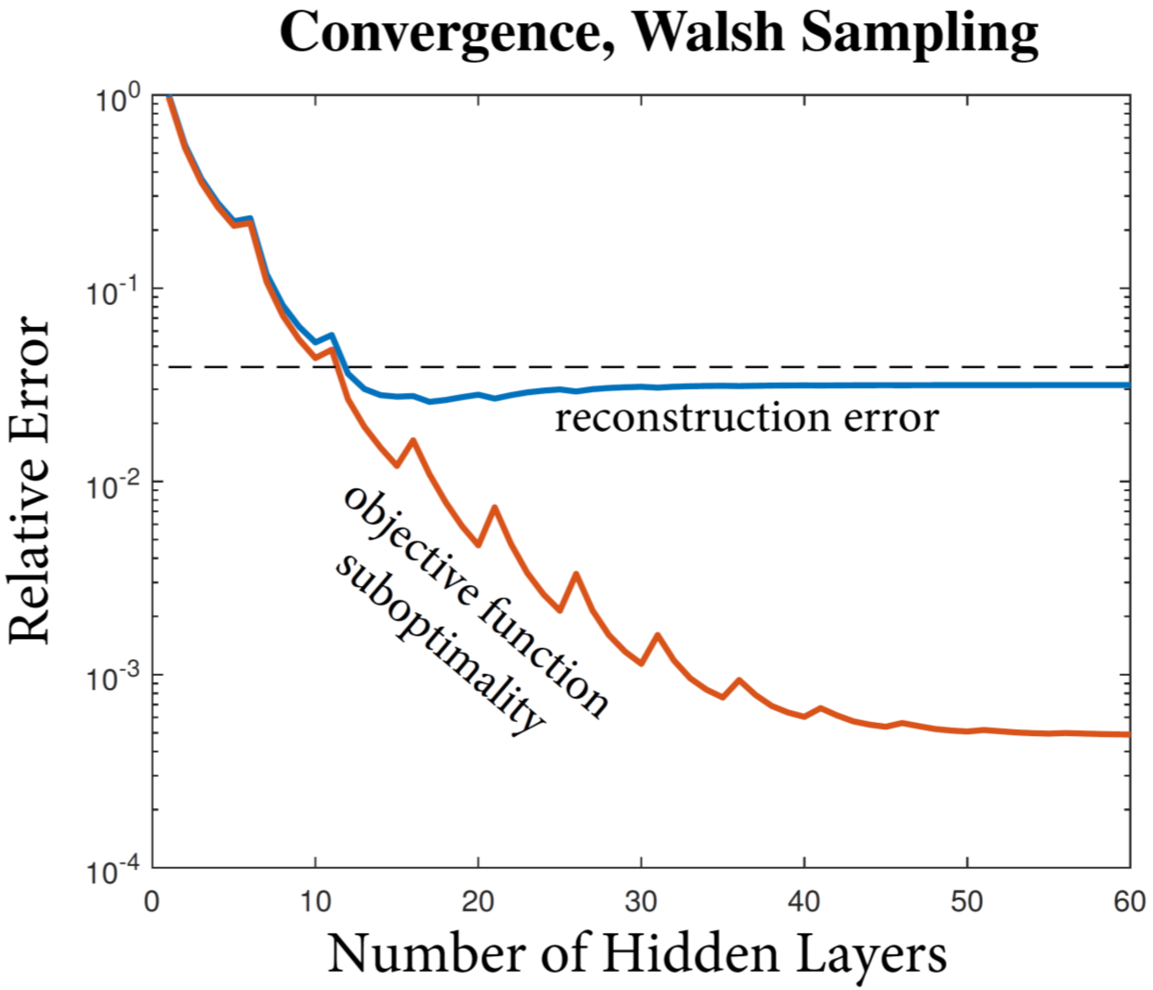


Figure: Images corrupted with 2% Gaussian noise and reconstructed using 15% sampling.

- C., Antun, Hansen, *“The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale’s 18th problem,”* **Proc. Natl. Acad. Sci. USA**.

Numerical example of GHA



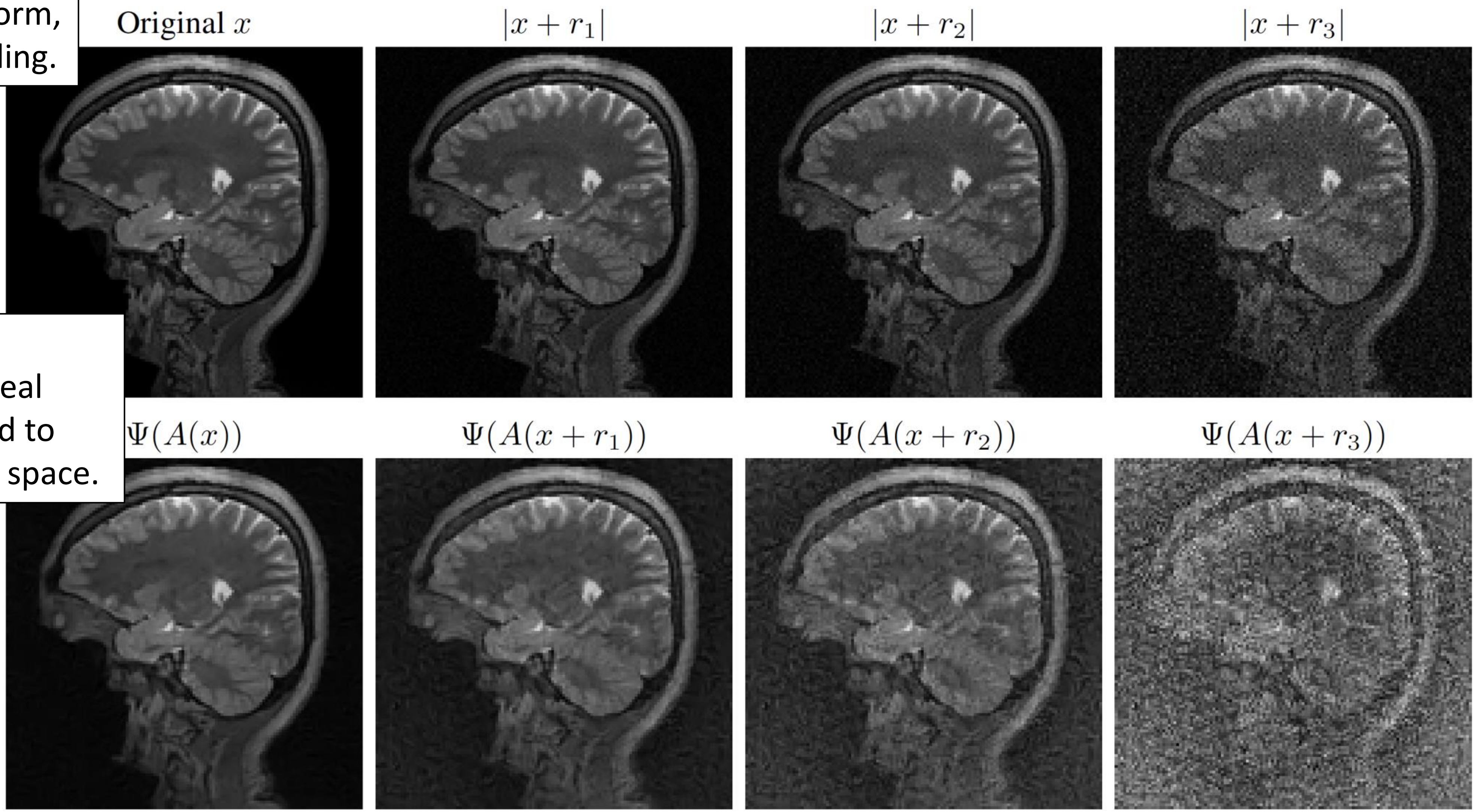


- C., Antun, Hansen, *"The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem,"* **Proc. Natl. Acad. Sci. USA**.

Example of severe instability

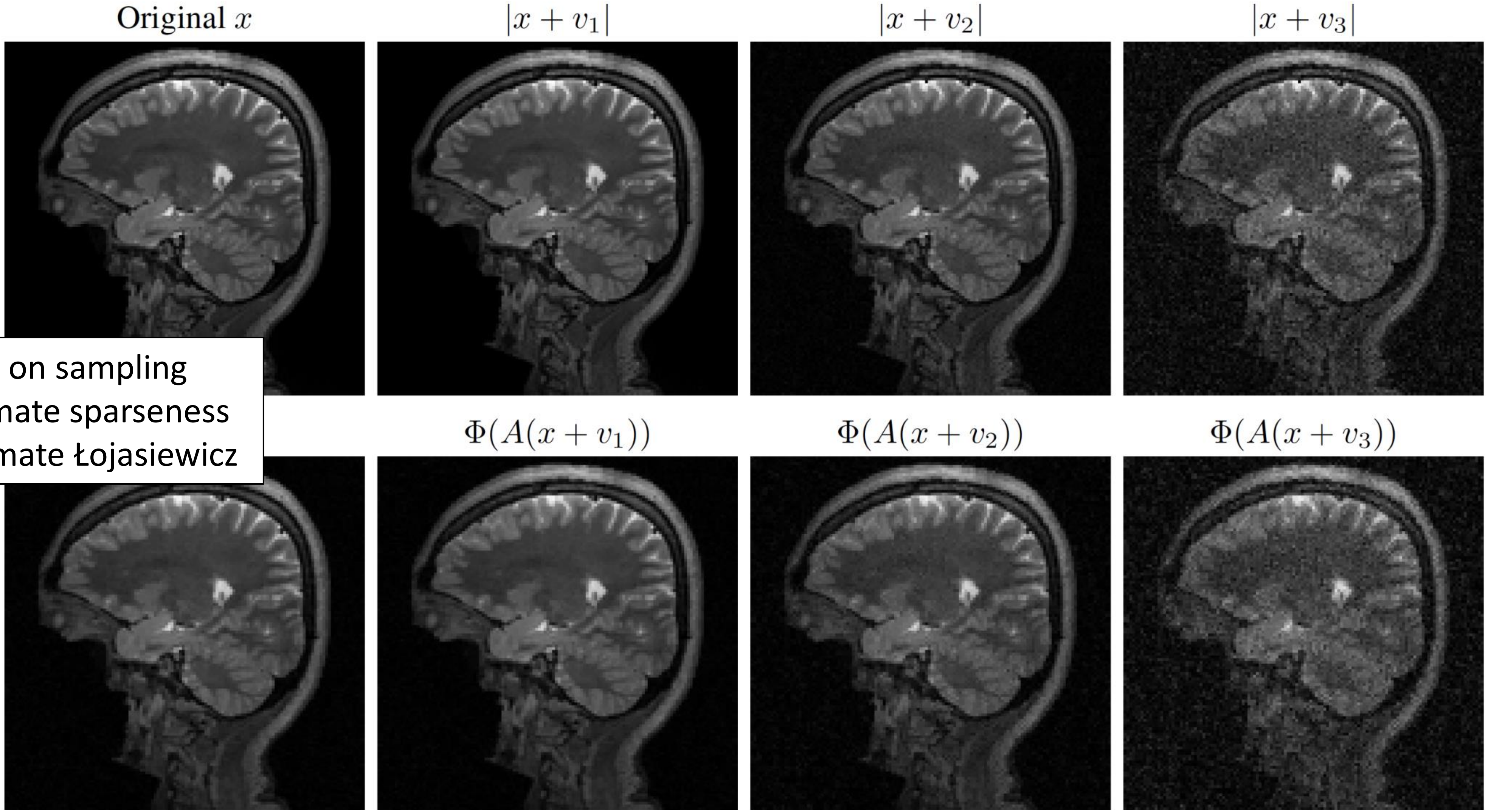
MRI: discrete 2D Fourier transform, 60% subsampling.

Perturbations computed in real space, mapped to measurement space.



- Zhu et al., *“Image reconstruction by domain-transform manifold learning,”* **Nature,** 2018.
- Antun et al., *“On instabilities of deep learning in image reconstruction and the potential costs of AI,”* **PNAS**, 2020.

FIRENET: provably stable (even to adversarial examples) and accurate



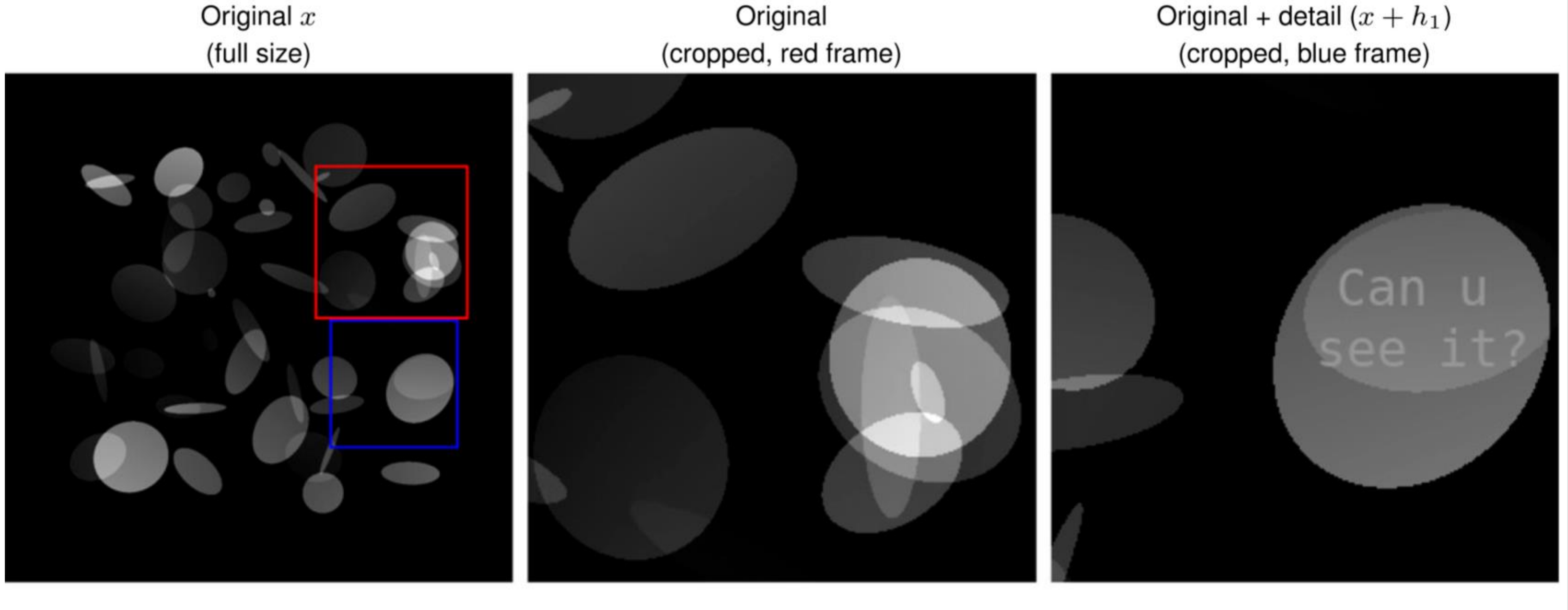
Assumptions on sampling and approximate sparseness give approximate Łojasiewicz

- C., Antun, Hansen, "*The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem,*" **PNAS**, 2022.

Key pillars: stability and accuracy

MRI: discrete 2D Fourier transform, 15% subsampling.

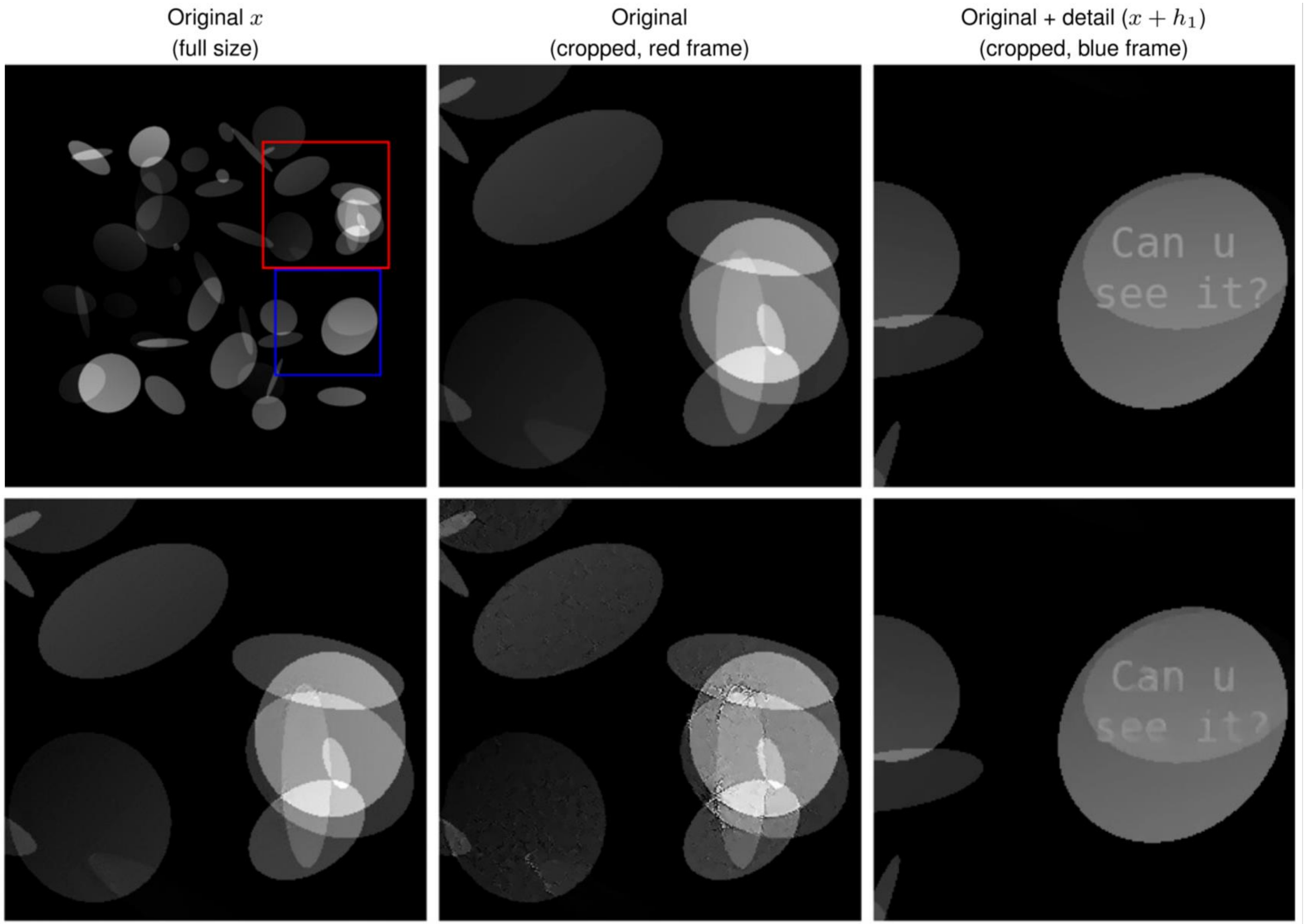
All networks trained on 5000 images of ellipses



- C., Antun, Hansen, "*The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem,*" **PNAS**, 2022.

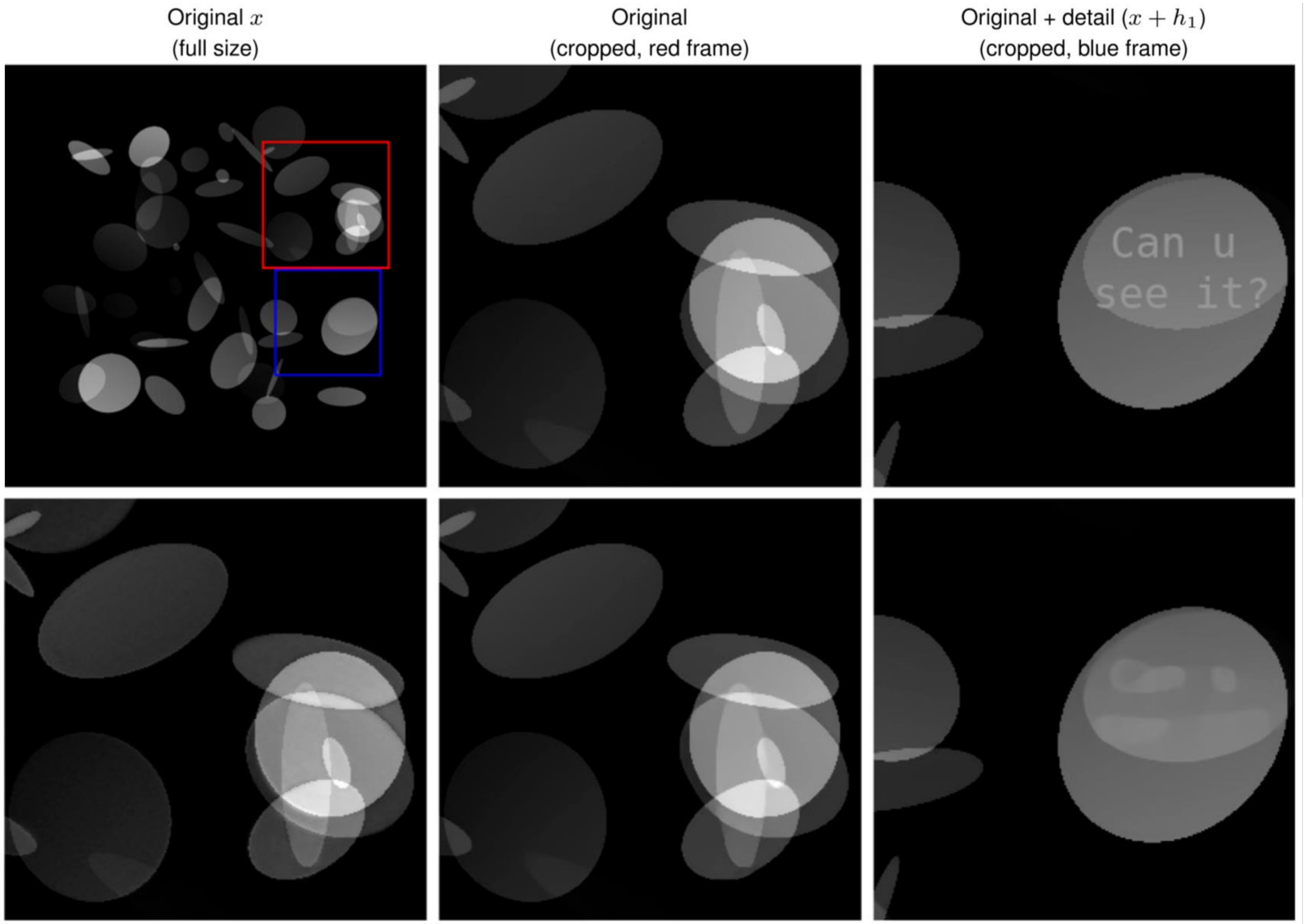
U-Net with no noise: accurate but unstable

U-Net: standard neural network architecture for imaging. Approx 4 million parameters.



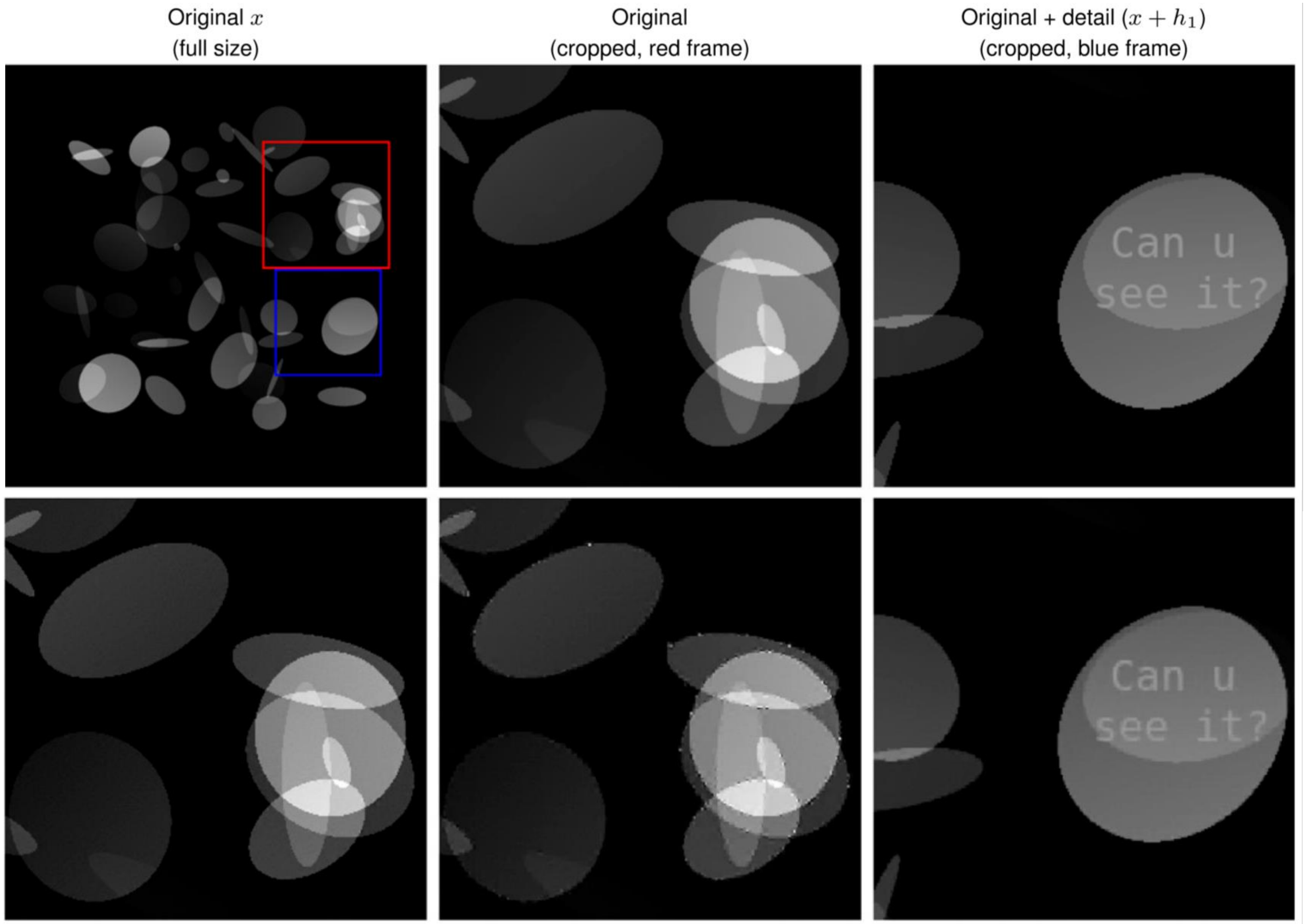
- C., Antun, Hansen, "*The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem,*" **PNAS**, 2022.

U-Net with noise: stable but inaccurate



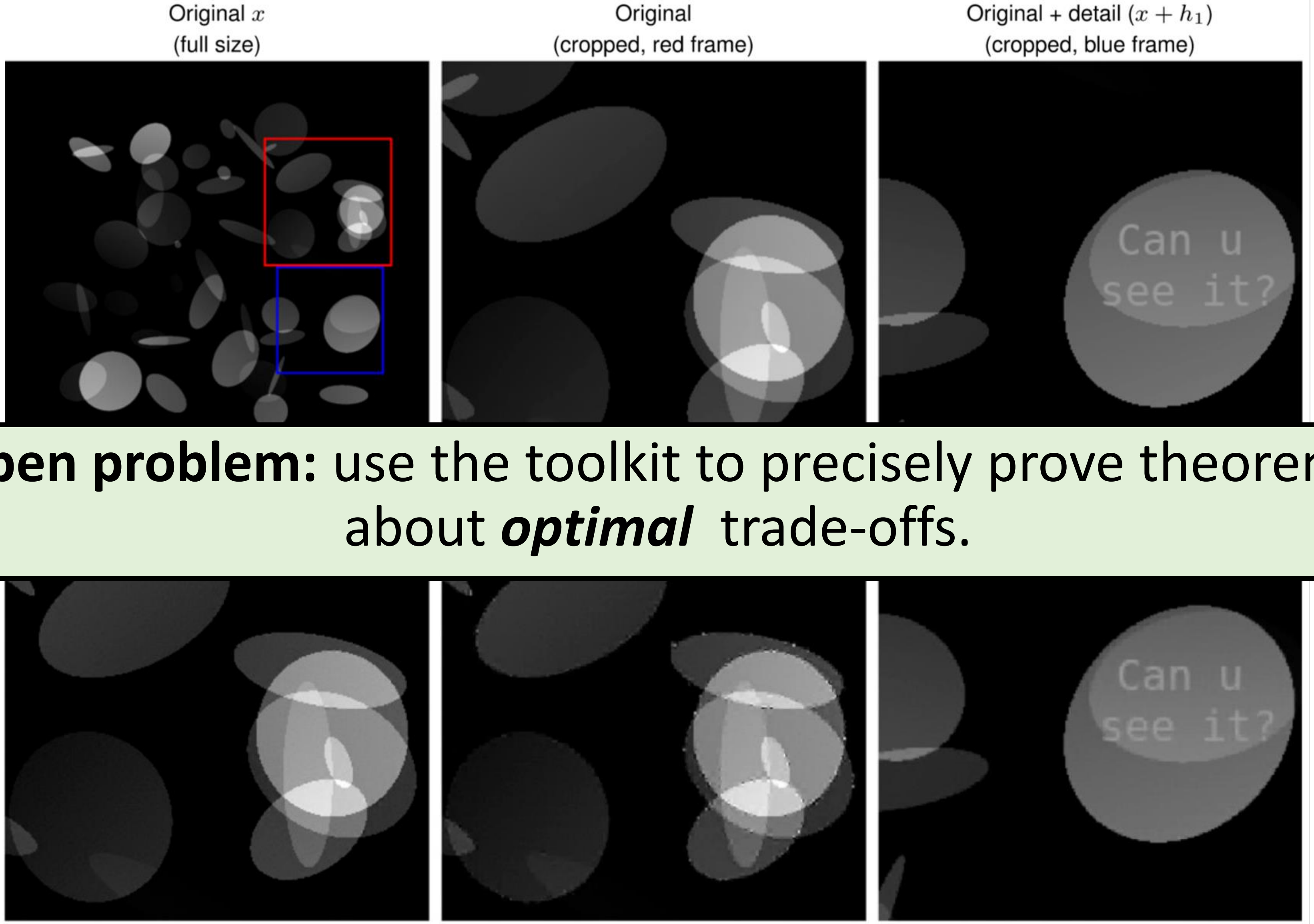
- C., Antun, Hansen, "*The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem,*" **PNAS**, 2022.

FIRENET: balances stability and accuracy?



- C., Antun, Hansen, "*The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem,*" **PNAS**, 2022.

FIRENET: balances stability and accuracy?



Open problem: use the toolkit to precisely prove theorems about ***optimal*** trade-offs.

- C., Antun, Hansen, "*The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem,*" **PNAS**, 2022.

Stabilising unstable neural networks