The Complexity of Gradient Descent: $CLS = PPAD \cap PLS$

ALEXANDROS HOLLENDER

JOINT WORK WITH JOHN FEARNLEY, PAUL GOLDBERG AND RAHUL SAVANI



NASH: Find a mixed Nash equilibrium of a game.

NASH:

Find a mixed Nash equilibrium of a game.

FACTORING:

Find a prime factor of a number $n \ge 2$.

NASH:

Find a mixed Nash equilibrium of a game.

FACTORING:

Find a prime factor of a number $n \ge 2$.

BROUWER:

Find a fixpoint of a continuous function $f: [0,1]^3 \rightarrow [0,1]^3$.

NASH:

Find a mixed Nash equilibrium of a game.

FACTORING: Find a prime factor of a number $n \ge 2$.

BROUWER:

Find a fixpoint of a continuous function $f: [0,1]^3 \rightarrow [0,1]^3$.

CONTRACTION:

Find the unique fixpoint of a contraction $f: [0,1]^n \rightarrow [0,1]^n$.

NASH:

Find a mixed Nash equilibrium of a game.

FACTORING: Find a prime factor of a number $n \ge 2$.

BROUWER:

Find a fixpoint of a continuous function $f: [0,1]^3 \rightarrow [0,1]^3$.

CONTRACTION:

Find the unique fixpoint of a contraction $f: [0,1]^n \rightarrow [0,1]^n$.

PURE-CONGESTION:

Find a pure Nash equilibrium of a congestion game.

NASH:

Find a mixed Nash equilibrium of a game.

FACTORING:

Find a prime factor of a number $n \ge 2$.

BROUWER:

Find a fixpoint of a continuous function $f: [0,1]^3 \rightarrow [0,1]^3$.

CONTRACTION:

Find the unique fixpoint of a contraction $f: [0,1]^n \rightarrow [0,1]^n$.

PURE-CONGESTION:

Find a pure Nash equilibrium of a congestion game.

What do these problems have in common?

NASH:

Find a mixed Nash equilibrium of a game.

FACTORING: Find a prime factor of a number $n \ge 2$.

BROUWER:

Find a fixpoint of a continuous function $f: [0,1]^3 \rightarrow [0,1]^3$.

CONTRACTION:

Find the unique fixpoint of a contraction $f: [0,1]^n \rightarrow [0,1]^n$.

PURE-CONGESTION:

Find a pure Nash equilibrium of a congestion game.

What do these problems have in common?

They are NP Total Search (TFNP) problems!

- Total: there is always a solution
- NP: it is easy to verify solutions

NASH:

Find a mixed Nash equilibrium of a game.

FACTORING: Find a prime factor of a number $n \ge 2$.

BROUWER:

Find a fixpoint of a continuous function $f: [0,1]^3 \rightarrow [0,1]^3$.

CONTRACTION:

Find the unique fixpoint of a contraction $f: [0,1]^n \rightarrow [0,1]^n$.

PURE-CONGESTION:

Find a pure Nash equilibrium of a congestion game.

What do these problems have in common?

They are NP Total Search (TFNP) problems!

- Total: there is always a solution
- NP: it is easy to verify solutions

Can a TFNP problem be NP-hard?

NASH:

Find a mixed Nash equilibrium of a game.

FACTORING: Find a prime factor of a number $n \ge 2$.

BROUWER:

Find a fixpoint of a continuous function $f: [0,1]^3 \rightarrow [0,1]^3$.

CONTRACTION:

Find the unique fixpoint of a contraction $f: [0,1]^n \rightarrow [0,1]^n$.

PURE-CONGESTION:

Find a pure Nash equilibrium of a congestion game.



Total NP search problems:

- "search" : looking for a solution, not just YES or NO
- "NP": any solution can be checked efficiently
- "total": there always exists at least one solution

Total NP search problems:

- "search" : looking for a solution, not just YES or NO
- "NP": any solution can be checked efficiently
- "total": there always exists at least one solution

TFNP lies between P and NP (search versions)

Total NP search problems:

- "search" : looking for a solution, not just YES or NO
- "NP": any solution can be checked efficiently
- "total": there always exists at least one solution

How do we show that a TFNP-problem is hard:

Total NP search problems:

- "search" : looking for a solution, not just YES or NO
- "NP": any solution can be checked efficiently

• "total": there always exists at least one solution

How do we show that a TFNP-problem is hard:

No TFNP-problem can be NP-hard, unless NP = coNP...

Total NP search problems:

- "search" : looking for a solution, not just YES or NO
- "NP": any solution can be checked efficiently

• "total": there always exists at least one solution

How do we show that a TFNP-problem is hard:

```
No TFNP-problem can be NP-hard, unless NP = coNP...
```

```
3-SAT \leq NASH \Rightarrow certificate for unsatisfiable 3-SAT formulas
```

Total NP search problems:

- "search" : looking for a solution, not just YES or NO
- "NP": any solution can be checked efficiently

• "total": there always exists at least one solution

How do we show that a TFNP-problem is hard:

No TFNP-problem can be NP-hard, unless NP = coNP...

Total NP search problems:

- "search" : looking for a solution, not just YES or NO
- "NP": any solution can be checked efficiently
- "total": there always exists at least one solution

How do we show that a TFNP-problem is hard:

- No TFNP-problem can be NP-hard, unless NP = coNP...
- Believed that no TFNP-complete problems exists...













What reasons are there to believe that PPAD \neq P, PLS \neq P, etc?

What reasons are there to believe that PPAD \neq P, PLS \neq P, etc?

• many seemingly hard problems lie in PPAD, PLS etc...

What reasons are there to believe that PPAD \neq P, PLS \neq P, etc?

- many seemingly hard problems lie in PPAD, PLS etc...
- oracle separations between the classes (in particular PPAD ≠ PLS)

What reasons are there to believe that PPAD \neq P, PLS \neq P, etc?

- many seemingly hard problems lie in PPAD, PLS etc...
- oracle separations between the classes (in particular PPAD \neq PLS)
- hard under cryptographic assumptions















PPAD ∩ PLS seems unnatural...
Problem A : PPAD-complete

Problem *B* : PLS-complete

Problem A : PPAD-complete

Problem *B* : PLS-complete

EITHER-SOLUTION(*A*,*B*):

Input: instance I_A of A, instance I_B of B*Goal:* find a solution of I_A , or a solution of I_B

Problem A : PPAD-complete

Problem *B* : PLS-complete

EITHER-SOLUTION(A,B): *Input:* instance I_A of A, instance I_B of B*Goal:* find a solution of I_A , or a solution of I_B

 \rightarrow EITHER-SOLUTION(*A*,*B*) is (PPAD \cap PLS)-complete!

BROUWER:

Input: a continuous function $f: [0,1]^n \rightarrow [0,1]^n$ Goal: find a fixpoint x

f(x) = x

BROUWER:

Input: a continuous function $f: [0,1]^n \rightarrow [0,1]^n$, precision $\varepsilon > 0$ *Goal:* find an approximate fixpoint x

 $\|f(x) - x\| \le \varepsilon$

BROUWER:

Input: a continuous function $f: [0,1]^n \rightarrow [0,1]^n$, precision $\varepsilon > 0$ Goal: find an approximate fixpoint x

 $\|f(x) - x\| \le \varepsilon$

REAL-LOCAL-OPT:

Input:

- a continuous function $p: [0,1]^n \rightarrow [0,1]$
- a (possibly non-continuous) function $g: [0,1]^n \rightarrow [0,1]^n$

BROUWER:

Input: a continuous function $f: [0,1]^n \rightarrow [0,1]^n$, precision $\varepsilon > 0$ Goal: find an approximate fixpoint x

 $\|f(x) - x\| \le \varepsilon$

REAL-LOCAL-OPT:

Input:

- a continuous function $p: [0,1]^n \rightarrow [0,1]$
- a (possibly non-continuous) function $g: [0,1]^n \rightarrow [0,1]^n$

Goal: find a local minimum of p with respect to g

 $p(g(x)) \ge p(x)$

BROUWER:

Input: a continuous function $f: [0,1]^n \rightarrow [0,1]^n$, precision $\varepsilon > 0$ Goal: find an approximate fixpoint x

 $\|f(x) - x\| \le \varepsilon$

REAL-LOCAL-OPT:

Input:

- a continuous function $p: [0,1]^n \rightarrow [0,1]$
- a (possibly non-continuous) function $g: [0,1]^n \rightarrow [0,1]^n$

Goal: find a local minimum of p with respect to g

 $p(g(x)) \ge p(x) - \varepsilon$

BROUWER:

Input: a continuous function $f: [0,1]^n \rightarrow [0,1]^n$, precision $\varepsilon > 0$ Goal: find an approximate fixpoint x

 $\|f(x) - x\| \le \varepsilon$

REAL-LOCAL-OPT:

Input:

- a continuous function $p: [0,1]^n \rightarrow [0,1]$
- a (possibly non-continuous) function $g: [0,1]^n \rightarrow [0,1]^n$

Goal: find a local minimum of *p* with respect to *g*

 $p(g(x)) \ge p(x) - \varepsilon$

 \rightarrow EITHER-SOLUTION(BROUWER,LOCAL-OPT) is (PPAD \cap PLS)-complete.

Continuous Local Search

But EITHER-SOLUTION(BROUWER,LOCAL-OPT) is not very natural...

Continuous Local Search

But EITHER-SOLUTION(BROUWER,LOCAL-OPT) is not very natural...

CONTINUOUS-LOCAL-OPT:

Input: continuous functions $g: [0,1]^n \rightarrow [0,1]^n$ and $p: [0,1]^n \rightarrow [0,1]$ *Goal:* find x such that

 $p(g(x)) \ge p(x) - \varepsilon$

Continuous Local Search

But EITHER-SOLUTION(BROUWER,LOCAL-OPT) is not very natural...

CONTINUOUS-LOCAL-OPT:

Input: continuous functions $g: [0,1]^n \rightarrow [0,1]^n$ and $p: [0,1]^n \rightarrow [0,1]$ *Goal:* find x such that

 $p(g(x)) \ge p(x) - \varepsilon$

→ class Continuous Local Search (CLS)

[Daskalakis-Papadimitriou, 2011]









PPAD: "all problems that can be solved by a path-following algorithm" (Lemke-Howson algorithm for NASH)

PPAD: "all problems that can be solved by a path-following algorithm" (Lemke-Howson algorithm for NASH)

PLS: "all problems that can be solved by a local search algorithm"

PPAD: "all problems that can be solved by a path-following algorithm" (Lemke-Howson algorithm for NASH)

- PLS: "all problems that can be solved by a local search algorithm"
- **CLS**: "all problems that can be solved by a *continuous* local search algorithm"

PPAD: "all problems that can be solved by a path-following algorithm" (Lemke-Howson algorithm for NASH)

PLS: "all problems that can be solved by a local search algorithm"

CLS: "all problems that can be solved by a *continuous* local search algorithm"

GD: "all problems that can be solved by gradient descent"

GD: "all problems that can be solved by gradient descent"

GD: "all problems that can be solved by gradient descent"

Input: C^1 -function $f: [0,1]^n \rightarrow [0,1]$, step size $\eta > 0$, precision $\varepsilon > 0$

GD: "all problems that can be solved by gradient descent"

Input: C^1 -function $f: [0,1]^n \rightarrow [0,1]$, step size $\eta > 0$, precision $\varepsilon > 0$

 $x_{k+1} \leftarrow x_k - \eta \nabla f(x_k)$

GD: "all problems that can be solved by gradient descent"

Input: C^1 -function $f: [0,1]^n \rightarrow [0,1]$, step size $\eta > 0$, precision $\varepsilon > 0$

$$x_{k+1} \leftarrow x_k - \eta \nabla f(x_k)$$

Goal: find a point where gradient descent terminates

GD: "all problems that can be solved by gradient descent"

Input: C^1 -function $f: [0,1]^n \rightarrow [0,1]$, step size $\eta > 0$, precision $\varepsilon > 0$

$$x_{k+1} \leftarrow x_k - \eta \nabla f(x_k)$$

Goal: find a point where gradient descent terminates

$$[x' \coloneqq x - \eta \nabla f(x)]$$

GD-Local-Search: Goal: find x such that $f(x') \ge f(x) - \varepsilon$

(the next iterate decreases f by at most ε)

GD: "all problems that can be solved by gradient descent"

Input: C^1 -function $f: [0,1]^n \rightarrow [0,1]$, step size $\eta > 0$, precision $\varepsilon > 0$

$$x_{k+1} \leftarrow x_k - \eta \nabla f(x_k)$$

Goal: find a point where gradient descent terminates

$$[x' \coloneqq x - \eta \nabla f(x)]$$

GD-Local-Search: Goal: find x such that $f(x') \ge f(x) - \varepsilon$

(the next iterate decreases f by at most ε)

 \rightarrow in CLS: $p(x) \coloneqq f(x)$ and $g(x) \coloneqq x - \eta \nabla f(x)$

GD: "all problems that can be solved by gradient descent"

Input: C^1 -function $f: [0,1]^n \rightarrow [0,1]$, step size $\eta > 0$, precision $\varepsilon > 0$

$$x_{k+1} \leftarrow x_k - \eta \nabla f(x_k)$$

Goal: find a point where gradient descent terminates

$$[x' \coloneqq x - \eta \nabla f(x)]$$

GD-Local-Search: Goal: find x such that $f(x') \ge f(x) - \varepsilon$

(the next iterate decreases f by at most ε)

GD: "all problems that can be solved by gradient descent"

Input: C^1 -function $f: [0,1]^n \rightarrow [0,1]$, step size $\eta > 0$, precision $\varepsilon > 0$

 $x_{k+1} \leftarrow x_k - \eta \nabla f(x_k)$

Goal: find a point where gradient descent terminates

 $[x' \coloneqq x - \eta \nabla f(x)]$

GD-Local-Search: Goal: find x such that $f(x') \ge f(x) - \varepsilon$

(the next iterate decreases f by at most ε)

GD-Fixed-Point: Goal: find x such that $||x' - x|| \le \varepsilon$

(the next iterate is ε -close)

GD: "all problems that can be solved by gradient descent"

Input: C^1 -function $f: [0,1]^n \rightarrow [0,1]$, step size $\eta > 0$, precision $\varepsilon > 0$

 $x_{k+1} \leftarrow x_k - \eta \nabla f(x_k)$

Goal: find a point where gradient descent terminates

 $[x' \coloneqq x - \eta \nabla f(x)]$

GD-Local-Search: Goal: find x such that $f(x') \ge f(x) - \varepsilon$

(the next iterate decreases f by at most ε)

GD-Fixed-Point: Goal: find x such that $||x' - x|| \le \varepsilon$

 \rightarrow polynomial-time equivalent!

(the next iterate is ε -close)











PPAD \cap **PLS** = **CLS** = **GD**

EITHER-SOLUTION(*A*, *B*) CONTINUOUS-LOCAL-OPT BANACH 2D-GD-FIXED-POINT






• PPAD \cap PLS is an interesting class!

Consequences

• PPAD \cap PLS is an interesting class!

• It captures continuous local search, and even gradient descent

Consequences

• PPAD \cap PLS is an interesting class!

• It captures continuous local search, and even gradient descent

• CLS and GD are robust with respect to:

- > dimension
- > domain
- > arithmetic circuits

> ...

Proof Sketch

Canonical complete problem: **END-OF-LINE**

Canonical complete problem: **END-OF-LINE**

Input: directed graph of paths and cycles, and a source

Canonical complete problem: **END-OF-LINE**

Input: directed graph of paths and cycles, and a source



Canonical complete problem: **END-OF-LINE**

Input: directed graph of paths and cycles, and a source



Canonical complete problem: **END-OF-LINE**

Input: directed graph of paths and cycles, and a source

Goal: find a sink, or another source



END-OF-LINE Canonical complete problem:

Input: directed graph of paths and cycles, and a source

Goal: find a sink, or another source



END-OF-LINE Canonical complete problem: Input: directed graph of paths and cycles, and a source Goal: find a sink, or another source The catch: the graph is given *implicitly* Vertex set $\{0,1\}^n$ Boolean circuits S and P successor circuit *S*: $\{0,1\}^n \rightarrow \{0,1\}^n$ predecessor circuit *P*: $\{0,1\}^n \rightarrow \{0,1\}^n$

Goal: reduction from EITHER-SOLUTION(END-OF-LINE, LOCAL-OPT) to 2D-GD-FIXED-POINT

Goal: reduction from EITHER-SOLUTION(END-OF-LINE, LOCAL-OPT) to 2D-GD-FIXED-POINT

Goal: reduction from EITHER-SOLUTION(END-OF-LINE, LOCAL-OPT) to 2D-GD-FIXED-POINT



Goal: reduction from EITHER-SOLUTION(END-OF-LINE, LOCAL-OPT) to 2D-GD-FIXED-POINT



Goal: reduction from EITHER-SOLUTION(END-OF-LINE, LOCAL-OPT) to 2D-GD-FIXED-POINT



Goal: reduction from EITHER-SOLUTION(END-OF-LINE, LOCAL-OPT) to 2D-GD-FIXED-POINT











Special case of END-OF-LINE: No backward edges allowed!



Special case of END-OF-LINE: No backward edges allowed!



Special case of END-OF-LINE: No backward edges allowed!





































Locally computable!

[Hubáček-Yogev, 2017] for CLS

Back to standard End-of-Line
Back to standard End-of-Line



Back to standard End-of-Line



Back to standard End-of-Line



green edges: forward red edges: backwards









































 \rightarrow to find a gradient descent fixed point, we have to solve the PPAD problem or the PLS problem

←● 10 $\xrightarrow{}_{7}$ \ddagger_{46} \ddagger_{46} \ddagger_{46} \mathbf{I}_{46} $_{46}$ \mathbf{I}_{46} \mathbf{I}_{46} $\begin{smallmatrix} 46 \\ 46 \end{smallmatrix} \begin{smallmatrix} 46 \\ 46 \end{smallmatrix} –$ $\begin{smallmatrix} 46 \\ 46 \end{smallmatrix} \begin{array}{c} \mathbf{1} \\ 46 \end{smallmatrix} \begin{array}{c} \mathbf{1} \\ 46 \end{array} \end{array}$ **↓** ¶ ₄₆ ¶ $_{46}$ m I**\$** 46 Ţ ⊶ ←• **J** \rightarrow ²⁴ 1²⁴ .24 ²⁴ 1 ²⁴ . 24 . 24 . 24 **J** . 97 ←• ←• **J** ↔ ↔ ← ¶ ₄₆ ¶ Ĵ ←• ←• ↔ ↔ 1 1 1 1 1 70 75 82 1 1 1 1 ↔ ↔ ↔ ↔ ↔ ↔ ←• ¶ ₄₆ ¶ • ↔ ↔ 1 1 1 1 1 1 70 75 82 . 22 1 1 . 22 -→ -→ -→ -↔ ↔ ↔ ↔ **↓** ⇔ ← ← $\mathbf{1}_{46}$ ←● **J** - ₄₆ ♥ Ĵ **↓** ↓ ↔ **\$** 1 ... 75 . 82 . 22 Ţ ↔ ←• **\$** $\mathbf{1}_{46}$ ←• ←• ←• ← ← \leftrightarrow \leftrightarrow \leftarrow ←• **\$** ← 73 73 73 73 73 73 91 75 . 82 . 22 73 73 73 73 73 69 73 73 - 99 $\bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet \rightarrow \bullet \rightarrow$ ←● ←● ↔ ←• $\mathbf{1}_{46}$ 1 1 1 1 1 1 1 1 1 1 1 70 75 82 22 1 1 € ← ← ← ← ←→ ↔→ ↔→ ↔→ ↔→ **\$** $_{46}$ m I← ⁷⁰ ⁷⁵ ↓ ⁸² ₂₈ ‡ \mathbf{I}_{46} \mathbf{I}_{10} ←• ←• ↔ ←• ←• **\$** $_{46}$ \clubsuit **I 94 96 3**5 1 1 ↔ ↔ • • • • • • • • $\mathbf{1}_{46}$ ↔ ⇔ **\$** ⇔ ⇔ ← Ĵ 1 1 1 1 1 1 1 1 1 1 • • • • • • • • • • 1 ₄₆ • T.

• are there other intersections of classes that are interesting?

- are there other intersections of classes that are interesting?
- candidates for (PPAD ∩ PLS)-completeness:
 - > CONTRACTION
 - > TARSKI
 - POLYNOMIAL-KKT
 - MIXED-CONGESTION

- are there other intersections of classes that are interesting?
- candidates for (PPAD ∩ PLS)-completeness:
 - CONTRACTION
 - > TARSKI

 - POLYNOMIAL-KKT
 MIXED-CONGESTION

Solved!

- are there other intersections of classes that are interesting?
- candidates for (PPAD ∩ PLS)-completeness:
 - CONTRACTION
 - > TARSKI

Solved!

POLYNOMIAL-KKTMIXED-CONGESTION

```
[Babichenko-Rubinstein, 2020]
2D-GD-FIXED-POINT \leq MIXED-CONGESTION \leq POLYNOMIAL-KKT
```

Thank You!