Clustered Integer 3SUM via Additive Combinatorics

Timothy Chan

(U. of Waterloo)

Moshe Lewenstein

(Bar-Ilan U.)

The Big Questions I

- Problem: All-Pairs Shortest Paths (APSP)

  Given weighted graph with n vertices, find shortest paths between all pairs, in $n^{3-\Omega(1)}$ time?

- Equiv. Problem: (min,+) Matrix Multiplication

  Given $n \times n$ matrices $\{a_{ij}\}, \{b_{ij}\}$, compute $s_{ij} = \min_k(a_{ik} + b_{kj})\ \forall i, j$, in $n^{3-\Omega(1)}$ time?

- Special Case: (min,+) Convolution

  Given sequences $\{a_i\}, \{b_i\}$ of length n, compute $s_i = \min_k(a_k + b_{i-k})\ \forall i$, in $n^{2-\Omega(1)}$ time?

The Big Questions II

- Problem: 3SUM

  Given sets A, B, S of n elements, decide if
  $\exists a \in A,\ b \in B,\ s \in S$ with $a + b = s$,
  in $n^{2-\Omega(1)}$ time?

- Equiv. Problem: 3SUM$^+$

  Given sets A, B, S of n elements, decide for every
  $s \in S$, if $\exists a \in A,\ b \in B$ with $a + b = s$,
  in $n^{2-\Omega(1)}$ time?

Conjecture: no to these questions

But I like positive results. . .

"Easy" Special Cases

- Small Int. APSP

  $c^{O(1)}n^{2.373}$ time [Alon&Galil&Margalit'91/Seidel'92] (undirected) or $c^{O(1)}n^{2.58}$ time [Zwick'98] (directed) if weights are in $[c]$

- Small Int. (min,+) Convolution

  $O(cn \log n)$ time by FFT if elements are in $[c]$

- Bounded Int. 3SUM$^+$

  $O(cn \log n)$ time by FFT if elements are in $[cn]$

Open Special Cases

- Problem: Small-Diff. Int. (min,+) Convolution

  Given int. sequence $\{a_i\}, \{b_i\}$ with $|a_{i+1} - a_i|, |b_{i+1} - b_i| \leq c$, compute (min,+) convolution

- Equiv. Problem: Bounded Int. Monotone (min,+) Convolution

  Given monotone increas. sequence $\{a_i\}, \{b_i\}$ in $[cn]$, compute (min,+) convolution

Open Special Cases

- Equiv. Problem: Binary Jumbled Indexing

  Given binary string of length n, compute, for all i,

  $s_i =$ min (or max) # of 1's over all length-i substrings

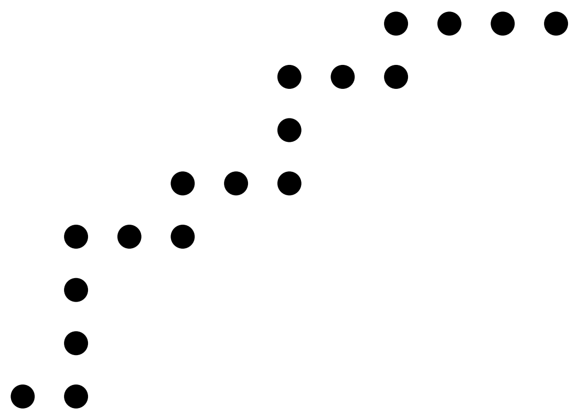
posed by many people [Burcsi&Cicalese&Fici&Lipták'10, Moosa&Rahman'10, Hermelin&Landau&Rabinovich& Weimann'14, . . .]

(let $a_i = i$-th prefix sum $\Rightarrow$ $s_i = \min_k(a_{k+i} - a_k)$)

Open Special Cases

- Equiv. Problem: Bounded Int. Connected Monotone $3SUM^+$ in 2D

  Given $A, B, S \subset [cn]^2$ that form connected xy-monotone sequences, solve $3SUM^+$

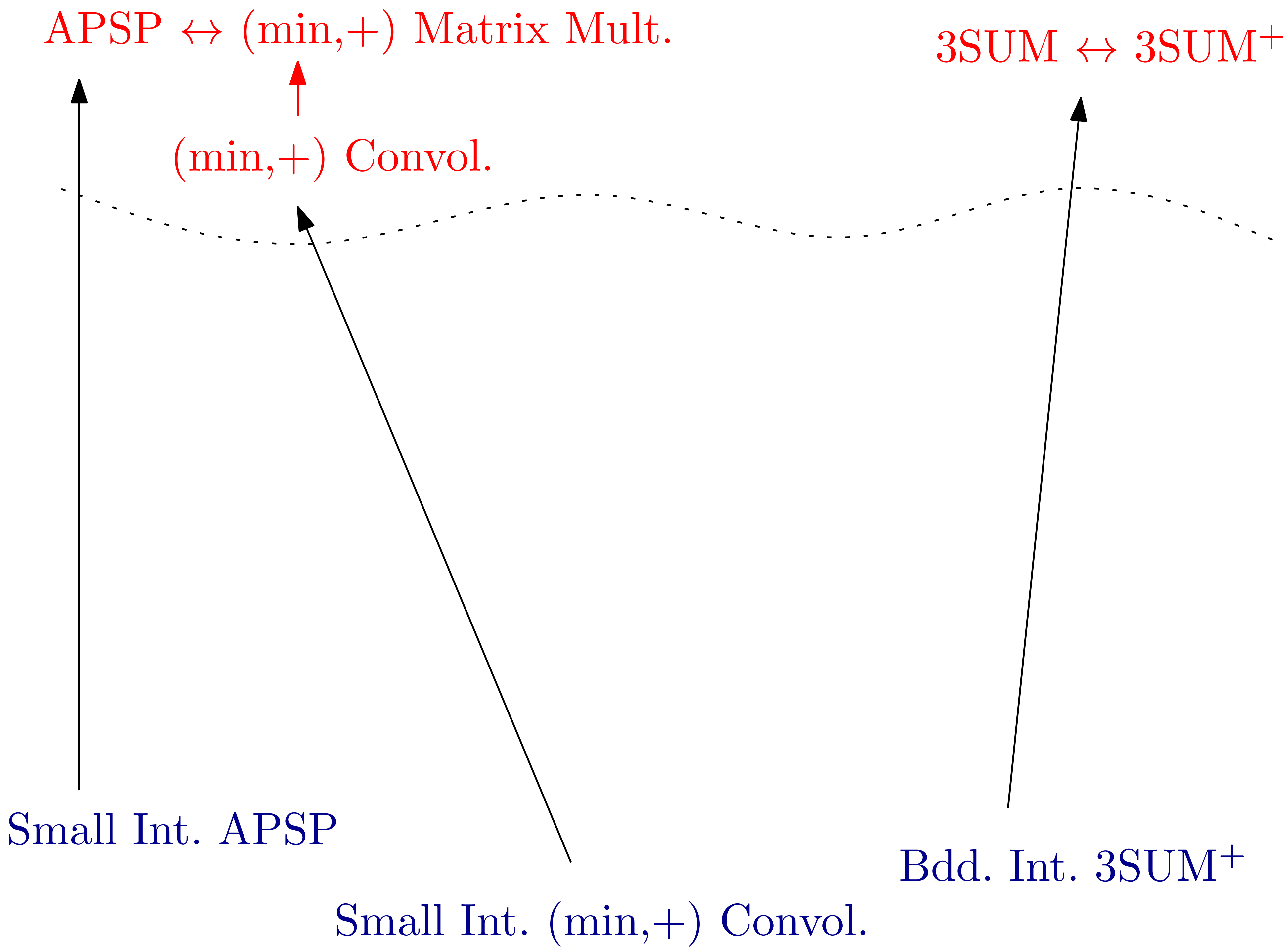


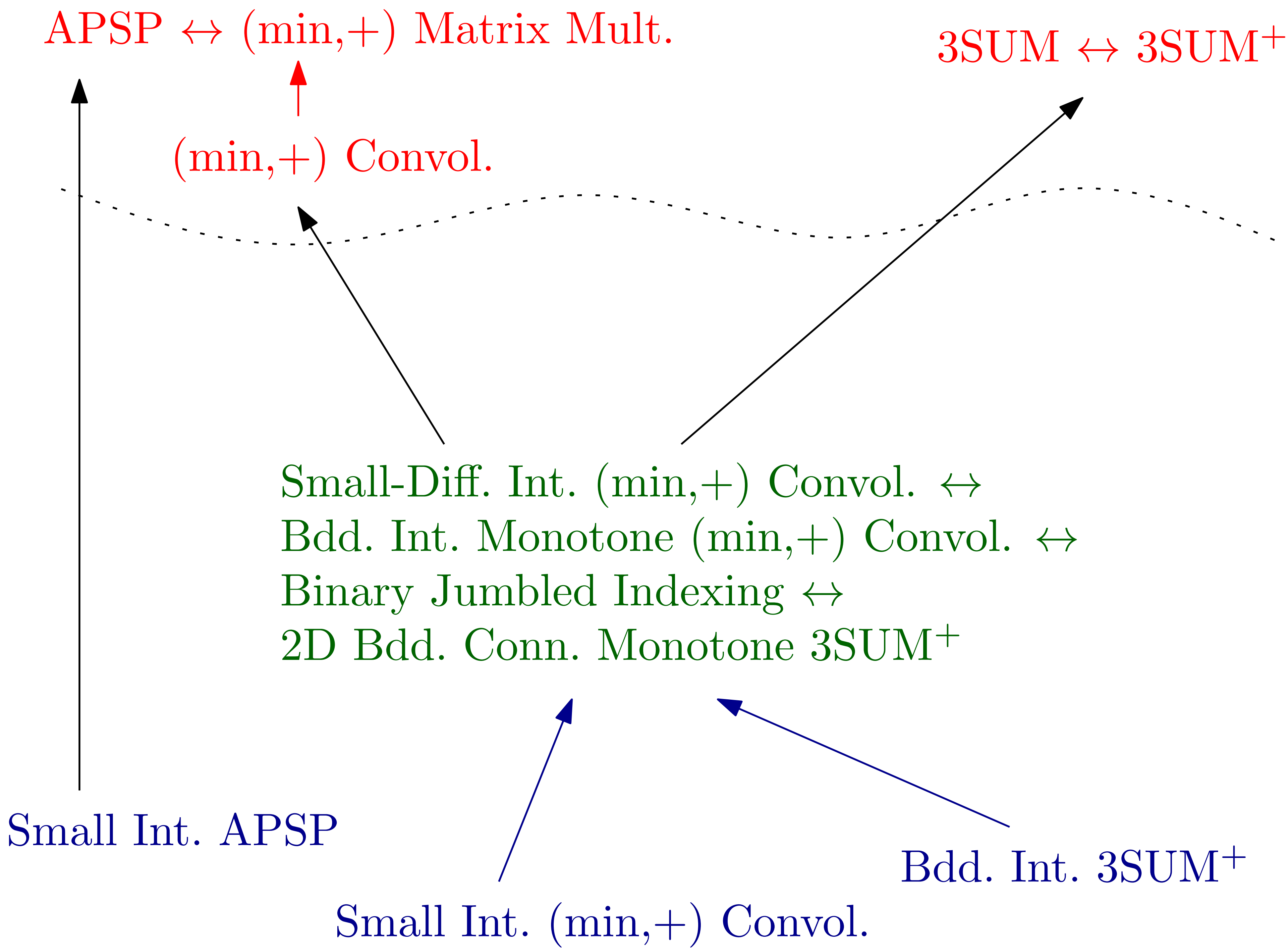
(let x = length of prefix, y = # of 1's in prefix)

New Results

- First truly subquadratic alg'ms for this group of problems!

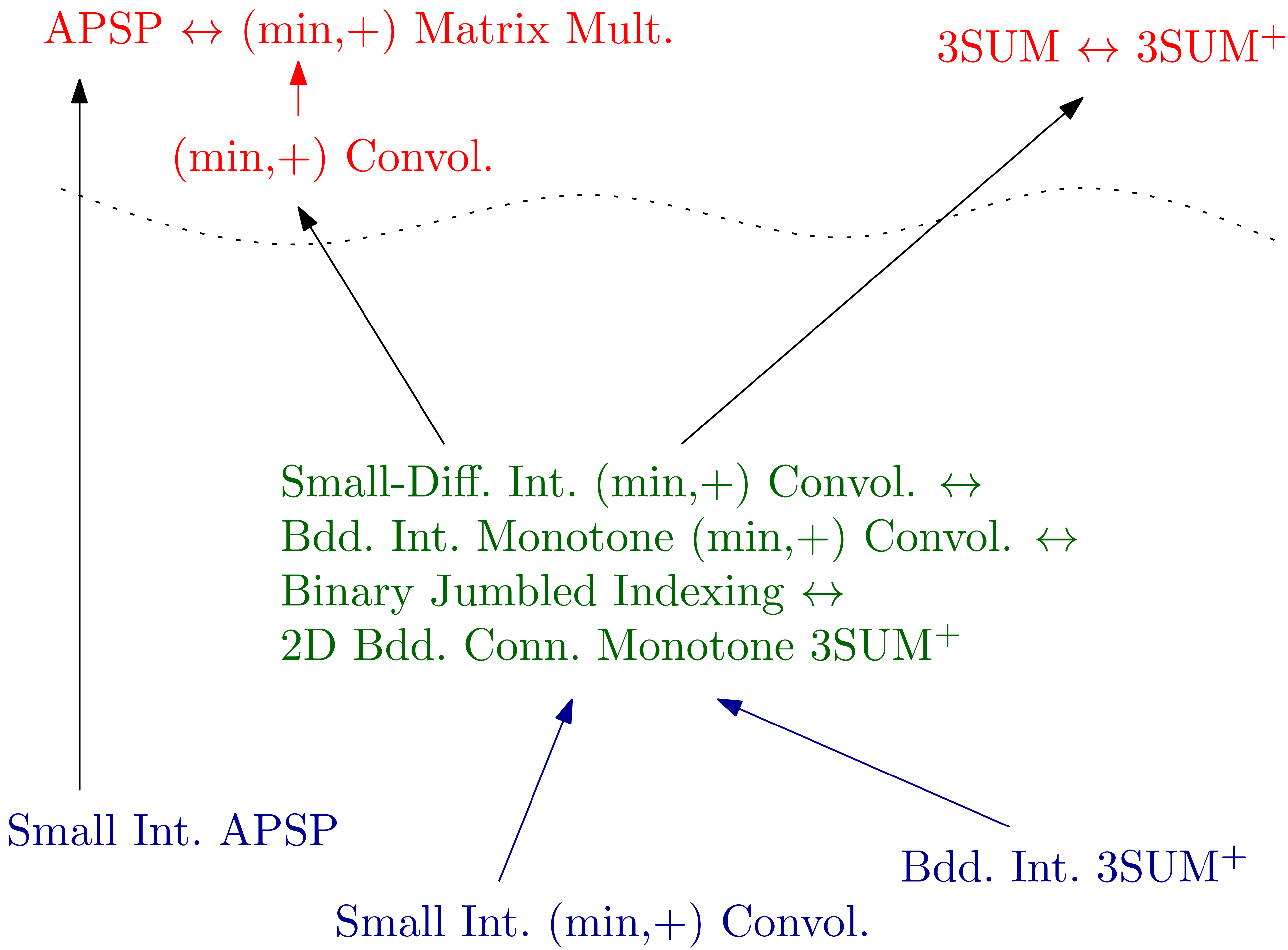
- Randomized time $\widetilde{O}(n^{(9+\sqrt{177})/12}) = \widetilde{O}(n^{1.859})$
- Deterministic time $\widetilde{O}(n^{1.864})$


APSP ↔ (min,+) Matrix Mult.
3SUM ↔ 3SUM+
(min,+) Convol.
Small Int. APSP
Bdd. Int. 3SUM+
Small Int. (min,+) Convol.

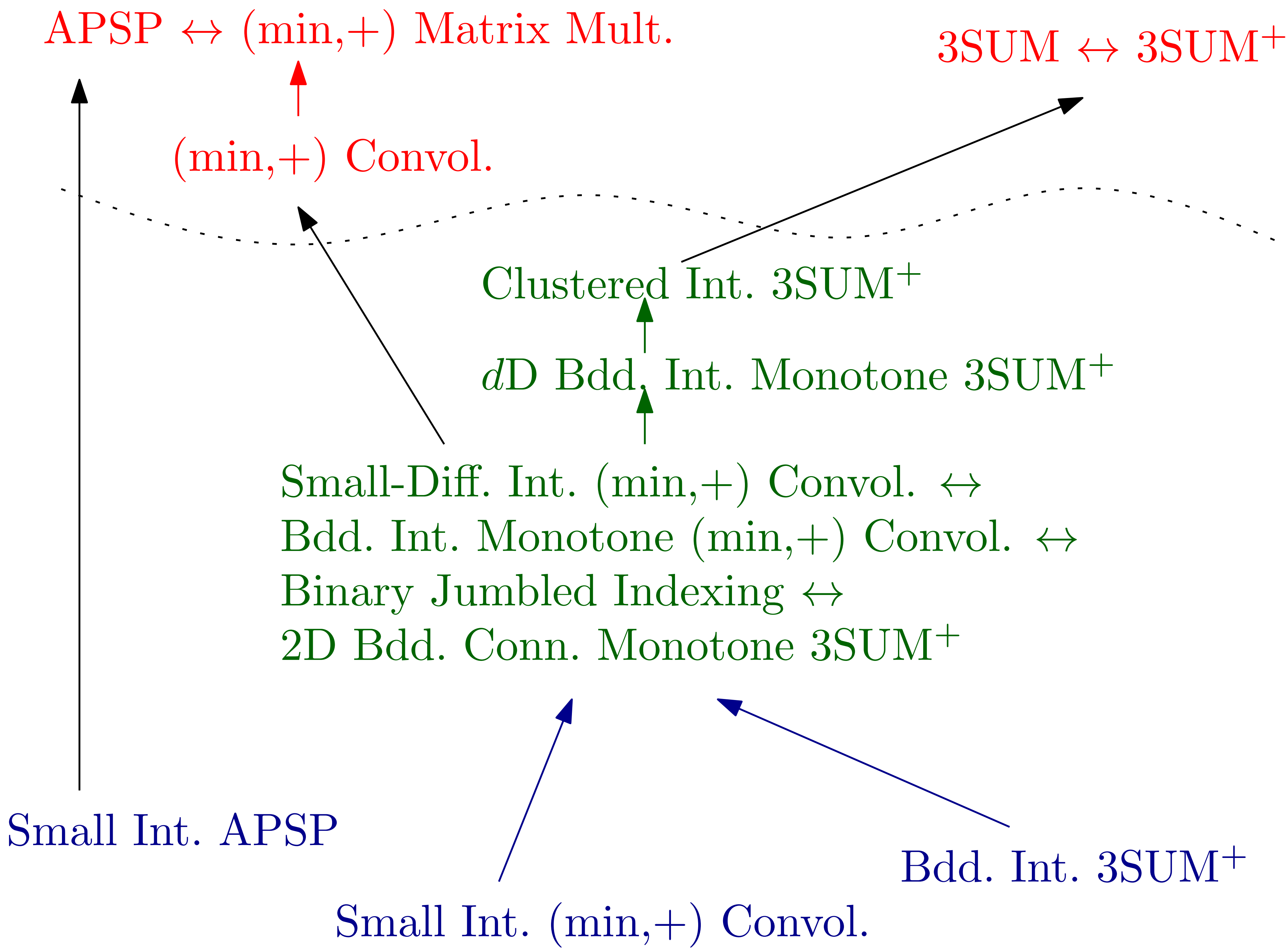

APSP ↔ (min,+) Matrix Mult.
3SUM ↔ 3SUM+
(min,+) Convol.
Small-Diff. Int. (min,+) Convol. ↔
Bdd. Int. Monotone (min,+) Convol. ↔
Binary Jumbled Indexing ↔
2D Bdd. Conn. Monotone 3SUM+
Small Int. APSP
Bdd. Int. 3SUM+
Small Int. (min,+) Convol.

More Results

- Bounded Int. Monotone 3SUM$^+$ in dD in $n^{2-2/(d+O(1))}$ rand. time
- Clustered Int. 3SUM$^+$ in $n^{2-\Omega(\varepsilon)}$ rand. time if input can be covered by $n^{1-\varepsilon}$ intervals of length n



APSP ↔ (min,+) Matrix Mult.
3SUM ↔ 3SUM+
(min,+) Convol.
Small-Diff. Int. (min,+) Convol. ↔
Bdd. Int. Monotone (min,+) Convol. ↔
Binary Jumbled Indexing ↔
2D Bdd. Conn. Monotone 3SUM+
Small Int. APSP
Bdd. Int. 3SUM+
Small Int. (min,+) Convol.


APSP ↔ (min,+) Matrix Mult.
3SUM ↔ 3SUM$^+$
(min,+) Convol.
Clustered Int. 3SUM$^+$
dD Bdd. Int. Monotone 3SUM$^+$
Small-Diff. Int. (min,+) Convol. ↔
Bdd. Int. Monotone (min,+) Convol. ↔
Binary Jumbled Indexing ↔
2D Bdd. Conn. Monotone 3SUM$^+$
Small Int. APSP
Bdd. Int. 3SUM$^+$
Small Int. (min,+) Convol.

Yet More Results. . .

- Data structure version of these problems

  For any query element s, decide if $\exists a \in A,\ b \in B$ with $a + b = s$

  e.g., $n^{2-\Omega(1/d)}$ preprocessing alg'm for d-ary jumbled indexing with $O(n^{2/3+\varepsilon})$ query time

- 3SUM for preprocessed universes:

  After preprocessing A, B, S in $\widetilde{O}(n^2)$ time, can solve 3SUM for any subsets A', B', S' of A, B, S in $\widetilde{O}(n^{13/7})$ time

  this holds for arbitrary integer input!

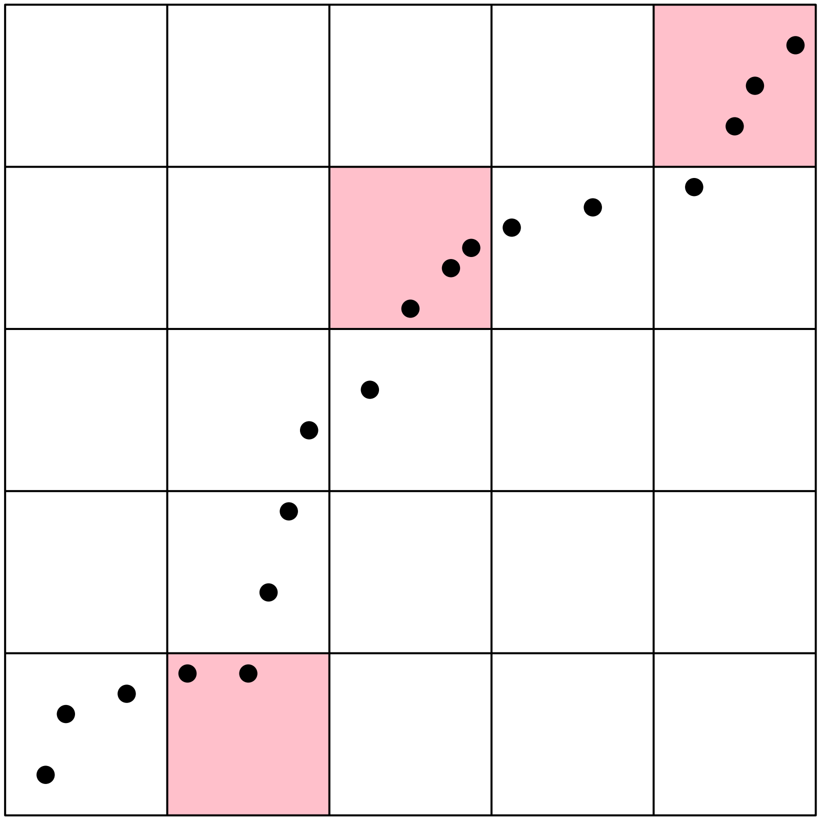
Surprising New Techniques

Besides FFT & fast matrix multiplication,

additive combinatorics. . .

2D Bdd. Int. Monotone 3SUM

Divide&Conquer Alg'm (1st Attempt)



- Given monotone sets $A, B, S \subset [n]^2$
- Divide $[n]^2$ into g^2 $(n/g) \times (n/g)$ grid cells
- Let A^*, B^*, S^* be the nonempty cells of A, B, S $(|A^*|, |B^*|, |S^*| = O(g))$
- For each triple (a^*, b^*, s^*) with $a^* + b^* = s^*$ $(+\{\pm 1\}^2)$, recurse on the points inside cells a^*, b^*, s^*

Divide&Conquer Alg'm (1st Attempt)

- Recurrence:

$$T(n) \ \le \ O(g^2)\, T(n/g) + O(n)$$

$\Rightarrow \ \ T(n) = \widetilde{O}(n^2)$ bad!

- Bad case: when can # of subproblems $\approx \Omega(g^2)$?

  e.g., when A^*, B^*, S^* are all nearly collinear...
  ...but then we can subtract a linear function & make all y values small ints, & solve by FFT

Need a Theorem that Says. . .

If we are in the bad case,
then the points must be nearly collinear??

Balog–Szemerédi–Gowers (BSG) Theorem

Let A, B, S be sets of size N in an abelian group.
If $|\{(a,b) \in A \times B : a+b \in S\}| = \Omega(\alpha N^2)$,
then $\exists\ A' \subset A,\ B' \subset B$ s.t.

- $|A' + B'| = O((1/\alpha)^5 N)$
- $|A'|, |B'| = \Omega(\alpha N)$

[Freiman's theorem says that if $|A' + A'|$ is small, then A' is close to "collinear" in some vague sense...]

[First proof by Balog&Szemerédi'94 required regularity lemma...
Simpler proof by Gowers'01, further refined by Balog'07/ Sudakov& Szemerédi&Vu'05]

Balog–Szemerédi–Gowers (BSG) Theorem: Stronger Version

Let A, B, S be sets of size N in an abelian group.

If $G \subset \{(a,b) \in A \times B : a+b \in S\}$ & $|G| = \Omega(\alpha N^2)$,

then $\exists\ A' \subset A,\ B' \subset B$ s.t.

- $|A' + B'| = O((1/\alpha)^5 N)$
- $|G \cap (A' \times B')| = \Omega(\alpha^2 N^2)$

Corollary to BSG

Let A, B, S be sets of size N in an abelian group.
Then $\exists\ A_1, \ldots, A_k \subset A,\ B_1, \ldots, B_k \subset B$ s.t.

- $R = \{(a,b) \in A \times B : a + b \in S\} \setminus \bigcup_i (A_i \times B_i)$ has size $O(\alpha N^2)$
- $|A_i + B_i| = O((1/\alpha)^5 N)$
- $k = O((1/\alpha)^2)$

Corollary to BSG

Let A, B, S be sets of size N in an abelian group.
Then $\exists\ A_1, \ldots, A_k \subset A,\ B_1, \ldots, B_k \subset B$ s.t.

- $R = \{(a,b) \in A \times B : a + b \in S\} \setminus \bigcup_i (A_i \times B_i)$ has size $O(\alpha N^2)$
- $|A_i + B_i| = O((1/\alpha)^5 N)$
- $k = O(1/\alpha)$

New Divide&Conquer Alg'm

0. Apply BSG Corollary to the sets A^*, B^*, S^* of grid cells
   $\Rightarrow \widetilde{O}(g^2)$ rand. time [see paper]

1. For each $(a^*, b^*) \in R$,
   recurse for the points inside cells $a^*, b^*, a^* + b^*$
   $\Rightarrow O(\alpha g^2)\, T(n/g)$ time

2. For $i = 1, \ldots, k$,
   compute $\{$all points in $A_i^*\} + \{$all points in $B_i^*\}$
   & check if it contains any point in S
   $\Rightarrow$ sumsets have $O(1/\alpha) \cdot O((1/\alpha)^5 g) \cdot (n/g)^2$ total size & can be computed by FFT in rand. time near linear in output size [see paper, or Cole&Hariharan'02]

New Divide&Conquer Alg'm

- Recurrence:

$$T(n) \le O(\alpha g^2)\, T(n/g) + \widetilde{O}(n + g^2) + \widetilde{O}(1/\alpha) \cdot O((1/\alpha)^5 g) \cdot (n/g)^2$$

- Set $g = n^{0.9293}, \;\; 1/\alpha = n^{0.1313}$

  $\Rightarrow \;\; T(n) = \boxed{O(n^{1.859})}$

Final Remarks

- Other results also follow from BSG Corollary
- Open: further improve the exponents,
  e.g., by improving the α-dependencies in BSG?
- Could additive combinatorics help for kSUM?
  Bdd. Monotone (min,+) Matrix Multiplication?
  Subset Sum?? General Int. 3SUM???
  General Int. APSP???