Parallelizing Greedy for Submodular Function Maximization

Kent Quanrud, UIUC

joint work w/ Chandra Chekuri, UIUC

Microsoft Research

February 5, 2019

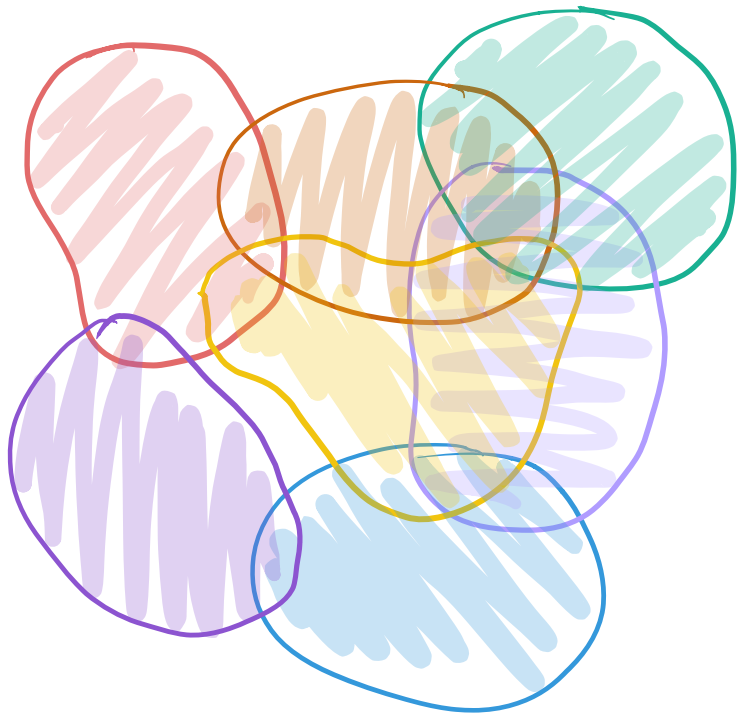
e.g. Maximum coverage

Input: n sets $S_1, S_2, \ldots, S_n$, cardinality k

Goal: choose k sets $S_{i_1}, S_{i_2}, \ldots, S_{i_k}$

maximizing their union $\left| \bigcup_{j=1}^{k} S_{i_j} \right|$

- NP-hard
- $1 - \frac{1}{e}$ APX-hardness
- $1 - \frac{1}{e}$ APX by greedy
- parallel?



Given:

- $f: 2^N \to \mathbb{R}_{\geq 0}$ submodular function
- $k \in \mathbb{N}$ cardinality constraint

Goal

maximize $f(I)$ s.t. $|I| \leq k$

(and in parallel?)

submodular functions

set function $f: 2^N \to \mathbb{R}_{\geq 0}$

- $f(\phi) = 0$
- [decreasing marginal returns] denote $f_S(U) = f(S \cup U) - f(S)$
  $=$ "marginal value of U to S"

$$S \subseteq T \Rightarrow f_S(U) \geq f_T(U)$$

- Oracle: given S, returns $f(S)$

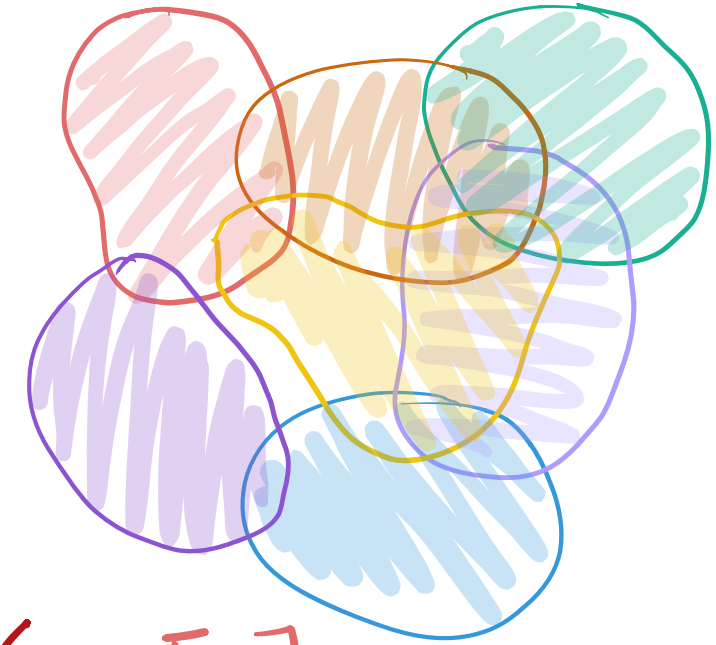
monotone f

$$S \subseteq T \Rightarrow f(S) \leq f(T)$$

Examples of submodular $f: 2^N \to \mathbb{R}_{\geq 0}$

Maximum coverage

Input: n sets $S_1, S_2, \ldots, S_n$

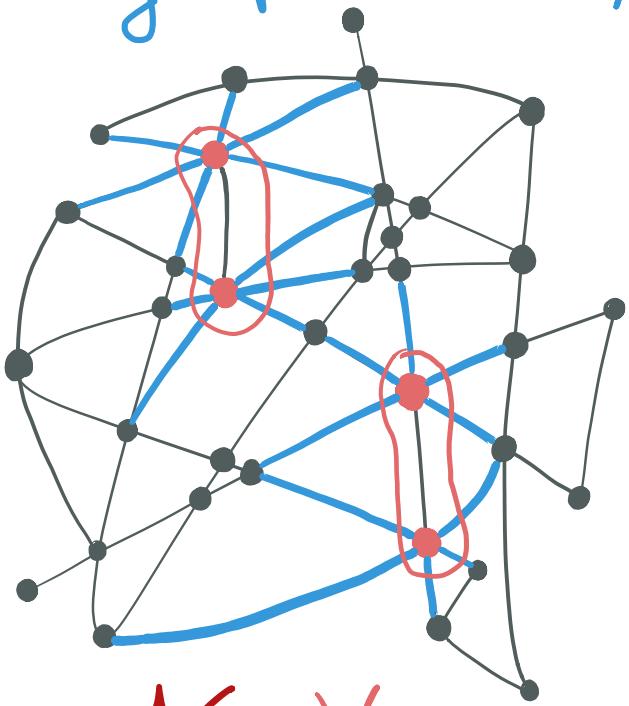


- $N = [n]$
- $f(I) = \left| \bigcup_{i \in I} S_i \right|$

Monotone

Max cut

Input: graph $G = (V, E)$



$N = V$

$f(S) = |\partial(S)|$

Nonnegative [not monotone]



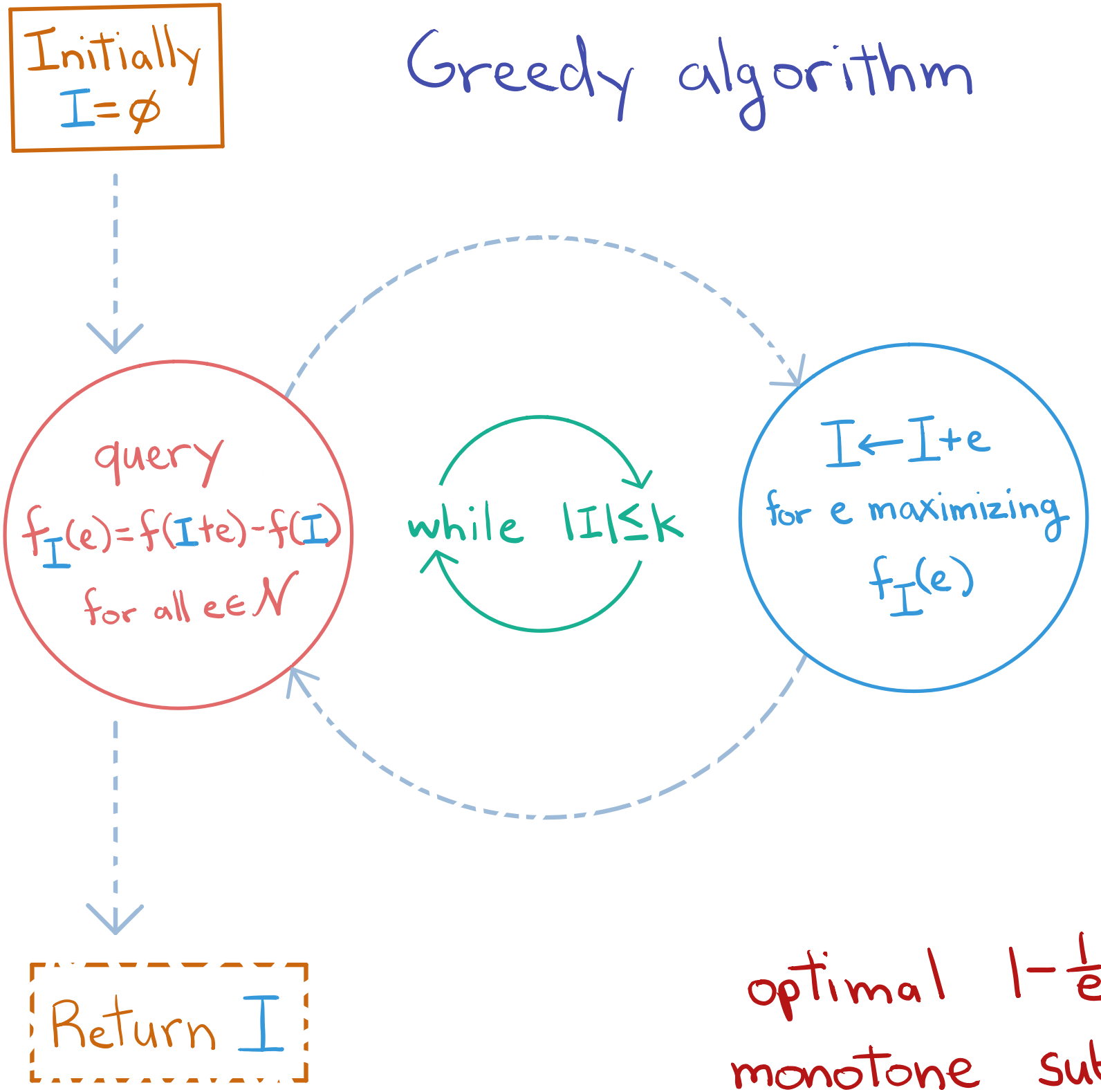
Greedy algorithm

[Nemhauser, Wolsey, Fisher '78]

optimal $1 - \frac{1}{e}$ APX for monotone submodular f

Greedy algorithm

[Nemhauser, Wolsey, Fischer '78]



optimal $1 - \frac{1}{e}$ APX for monotone submodular f

[1978: Nemhauser, Wolsey, Fisher] greedy analysis for monotone f

$1 - \frac{1}{e}$ for cardinality, $\frac{1}{k+1}$ for k matroid intersection

[1978: Nemhauser, Wolsey, Fisher] greedy analysis for monotone f

$1-\frac{1}{e}$ for cardinality, $\frac{1}{k+1}$ for k matroid intersection

...30 years later...

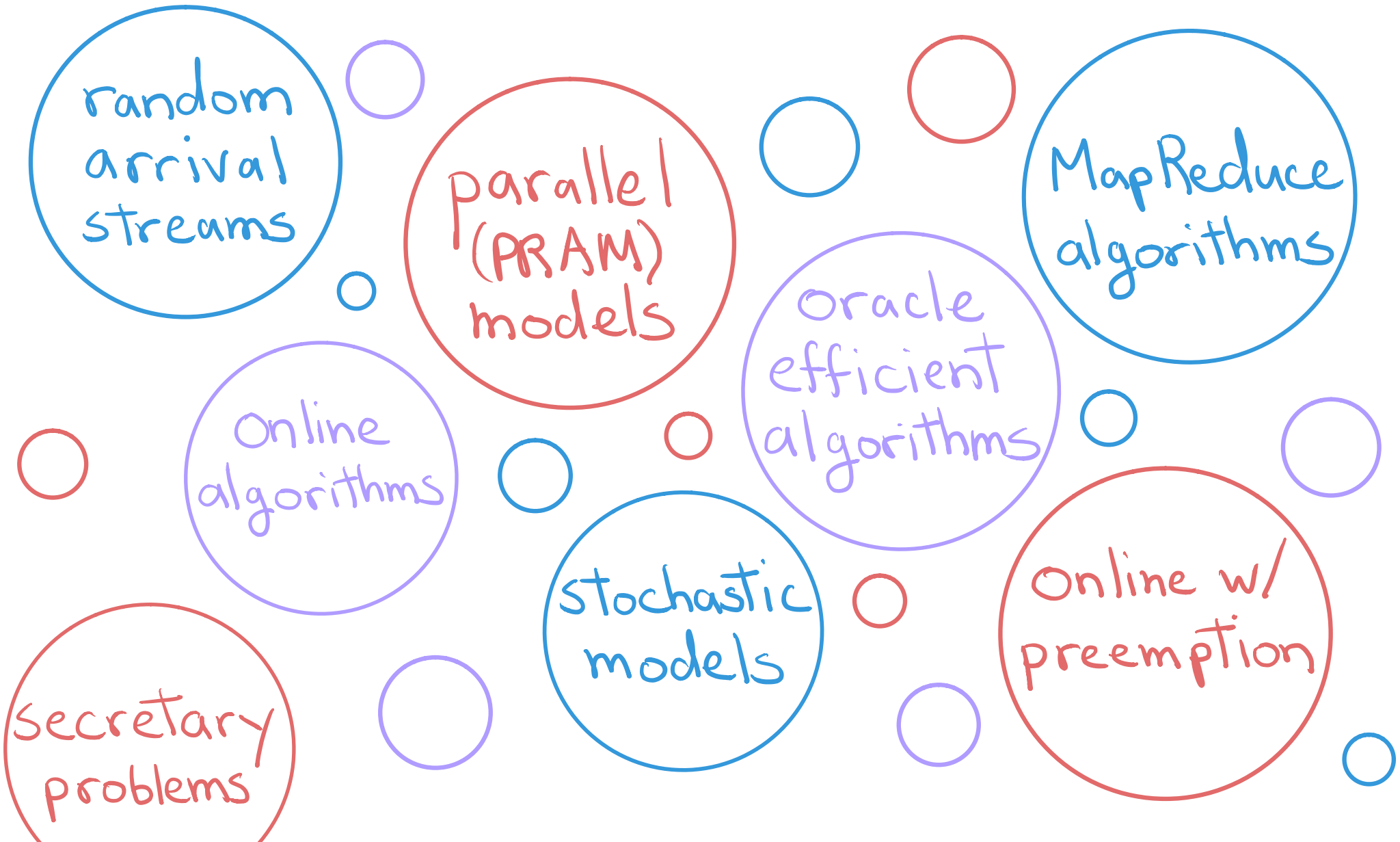
2000's: explosion of interest

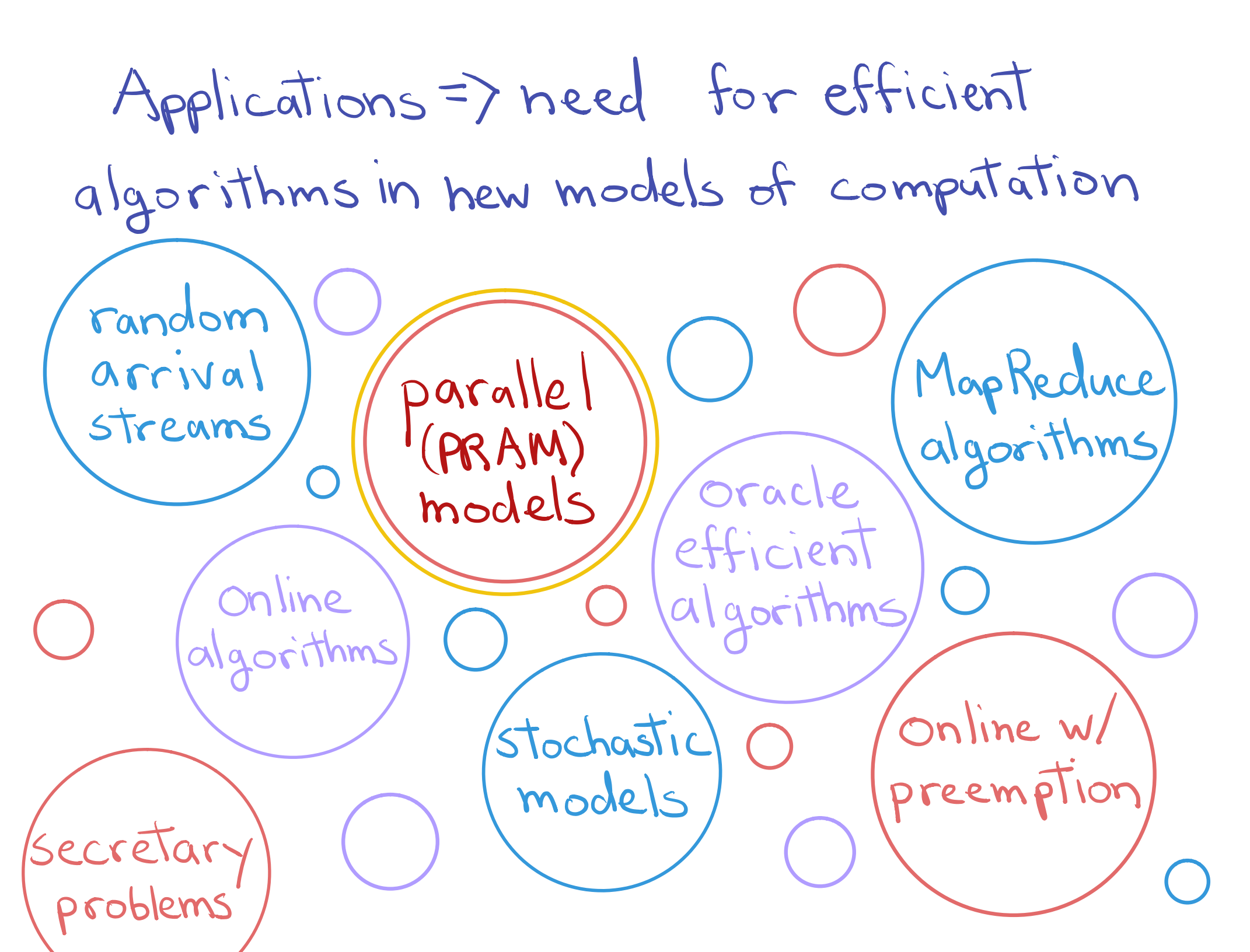
Improved bounds:
- matroids, packing constraints
- nonnegative f, unconstrained, ≥

New techniques:
- continuous relaxations, swap/pipage rounding,
- randomized techniques, ≥

growing applications:
- Machine learning, data mining,
- new connections in Comb. OPT

Applications => need for efficient algorithms in new models of computation

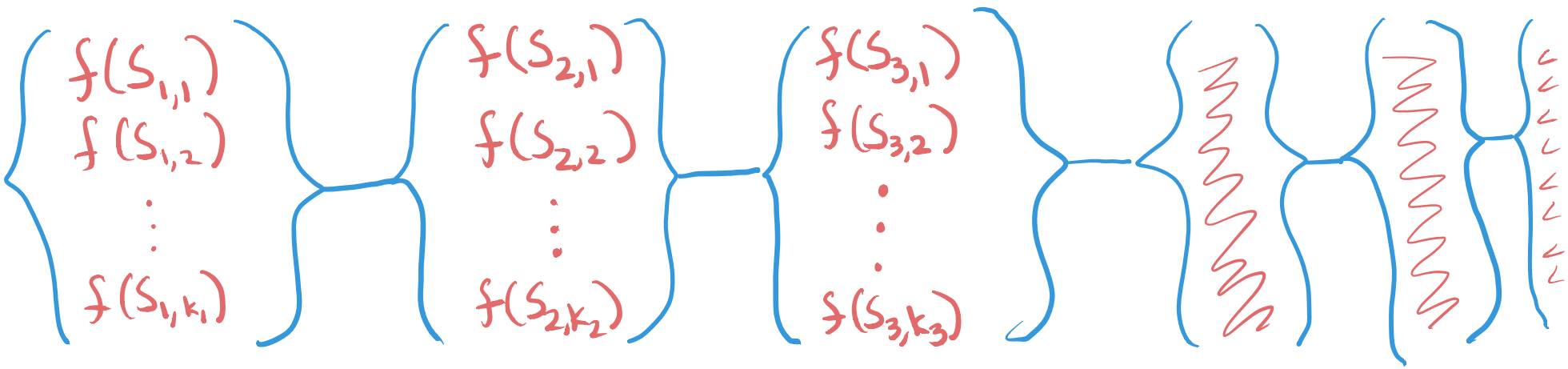



Applications => need for efficient algorithms in new models of computation
random arrival streams
parallel (PRAM) models
MapReduce algorithms
oracle efficient algorithms
online algorithms
stochastic models
online w/ preemption
secretary problems

Adaptivity and parallelization

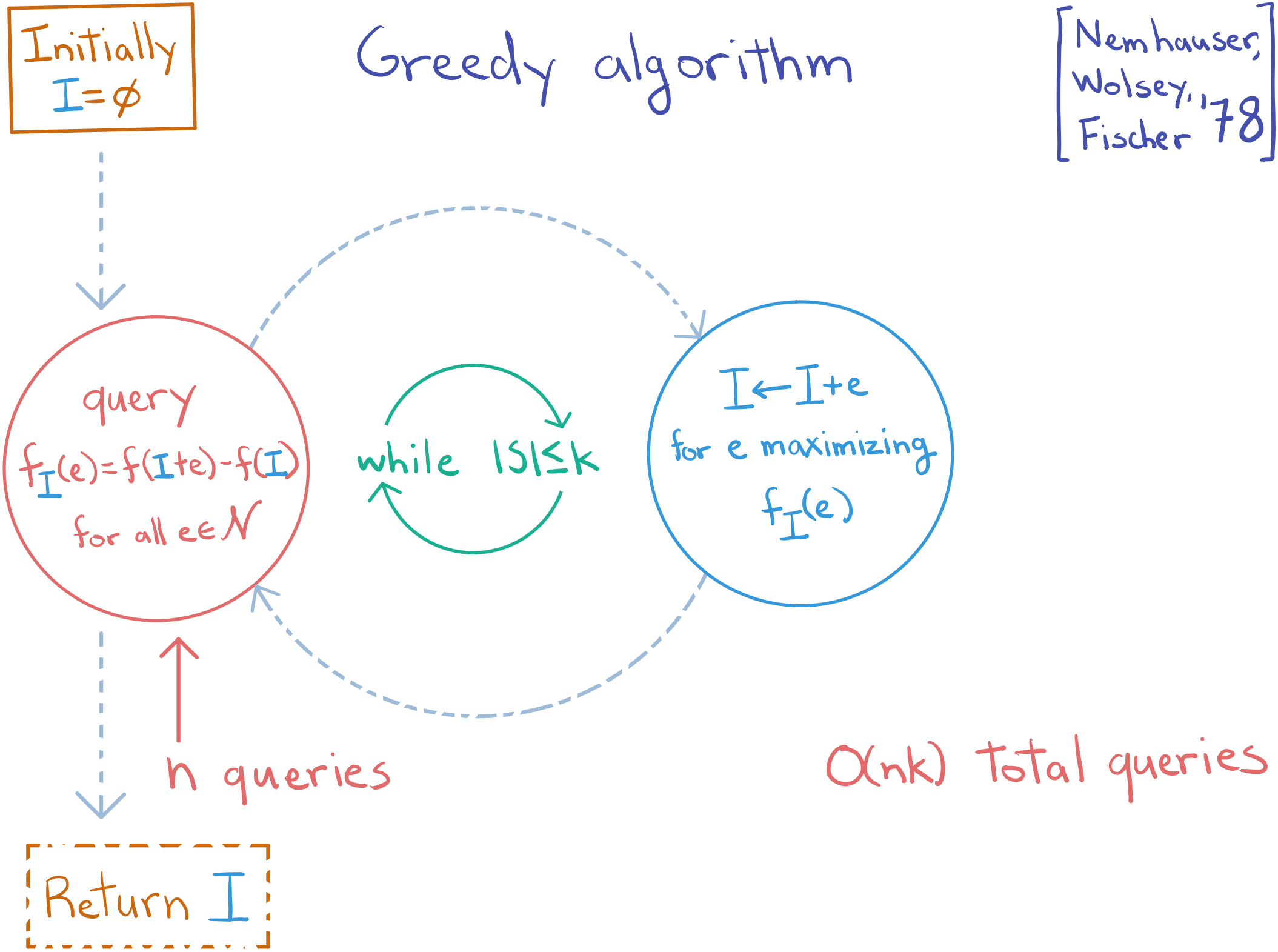
[Balkanski, Singer]

- (essentially) PRAM oracle for f
- Queries to f divided into "adaptive rounds"
- Choice of queries can only depend on queries to f of previous rounds



Greedy algorithm

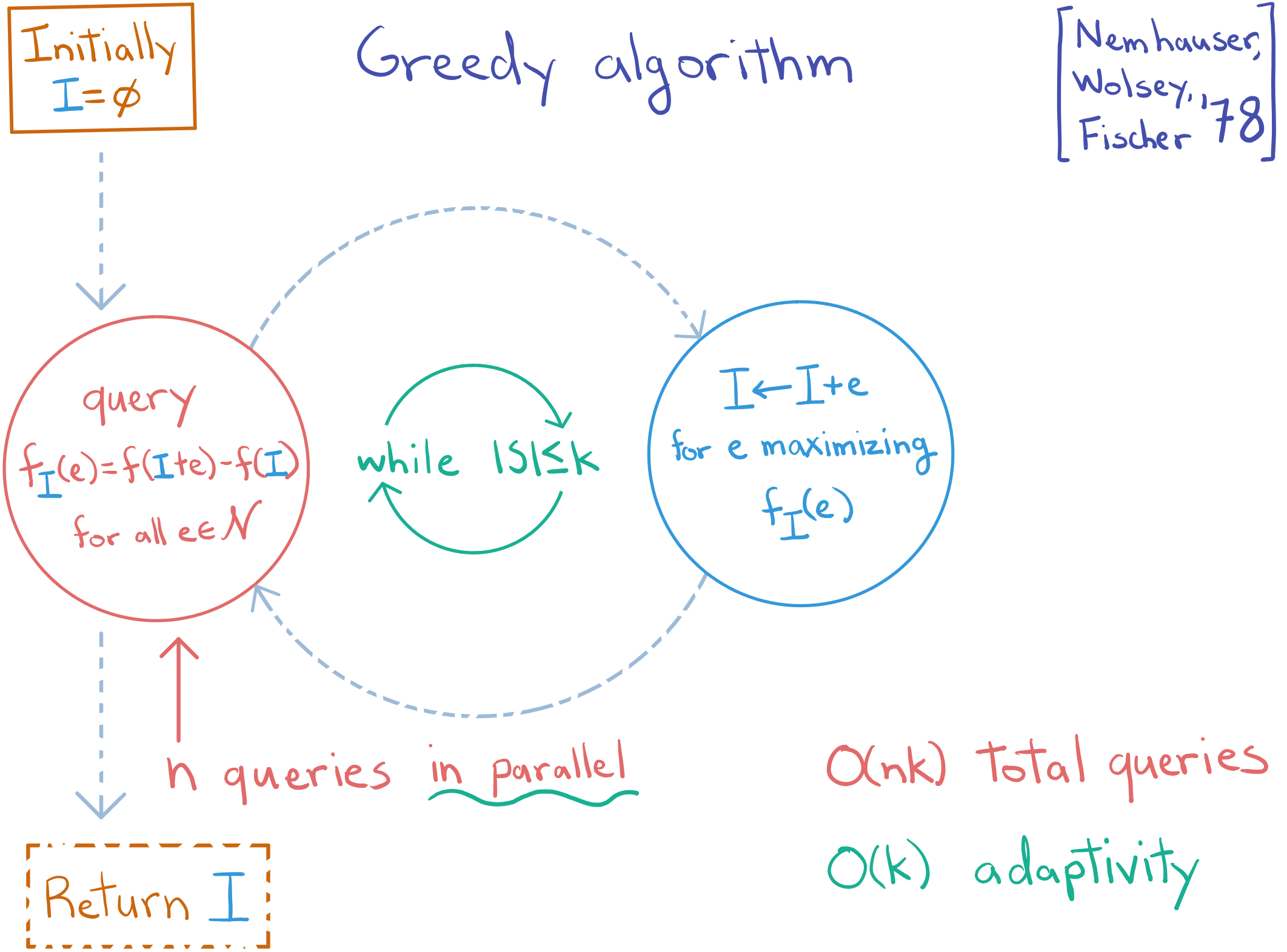
[Nemhauser, Wolsey, Fischer '78]



$O(nk)$ total queries

Greedy algorithm

[Nemhauser, Wolsey, Fischer '78]



$O(nk)$ total queries

$O(k)$ adaptivity

Given:

- $f: 2^N \to \mathbb{R}_{\geq 0}$ submodular function
- $k \in \mathbb{N}$ cardinality constraint

Goal

maximize $f(I)$ s.t. $|I| \leq k$

in parallel in the oracle model

Recent interest in parallel submodular max for cardinality constraints

[Balkanski, Singer] (STOC '18)

lower bound of $\Omega\left(\frac{\log n}{\log\log n}\right)$ rounds for $\Omega(1)$-APX

$\frac{1}{3}$-APX in $O(\log n)$ rounds

[Balkanski, Rubinstein, Singer] [Ene, Nguyen] (SODA '19)

$(1-\frac{1}{e}-\varepsilon)$-APX in $O\left(\frac{\log n}{\varepsilon^2}\right)$ rounds

Our results

Packing constraints

(SODA '19)

[Chekuri, Quanrud]

$(1-\frac{1}{e}-\varepsilon)$-APX in $\text{poly}(\log n, \frac{1}{\varepsilon})$ rounds

takes continuous viewpoint

$Ax \leq b$ captures many constraints

Matroids

[Chekuri, Quanrud]

matches sequential bounds in $\text{poly}(\log n, \frac{1}{\varepsilon})$ rounds

monotone and generally nonnegative f

extends to matroid intersection, matchoids

Further interest in parallel submodular max:
different constraints, nonnegative setting,
better depth, better oracle complexity

[Balkanski
Singer]

[Ene
Nguyen]

[Balkanski
Rubinstein
Singer]

[Fahrbach
Mirrokni
Zadimoghaddam]

[Ene
Nguyen
Vladu]

[Chen
Feldman
Karbasi]

(w/ multiplicities)

Given:

- $f: 2^N \to \mathbb{R}_{\geq 0}$ submodular function
- $k \in \mathbb{N}$ cardinality constraint

Goal

maximize $f(I)$ s.t. $|I| \leq k$

in parallel in the oracle model

Parallelizing greedy for cardinality

$1-\frac{1}{e}$ APX for monotone f in sequential setting

[Fisher, Nemhauser, Wolsey]



[briefly]

$(1-\frac{1}{e}-\varepsilon)$-APX for monotone f w/ $\tilde{O}(\frac{1}{\varepsilon^2})$ depth



[see SODA paper]

$(1-\frac{1}{e}-\varepsilon)$-APX for general packing w/ low depth

Parallelizing greedy for cardinality

$1-\frac{1}{e}$ APX for monotone f in sequential setting

[Fisher, Nemhauser, Wolsey]

[briefly]

$(1-\frac{1}{e}-\varepsilon)$-APX for monotone f w/ $\tilde{O}(\frac{1}{\varepsilon^2})$ depth

This talk

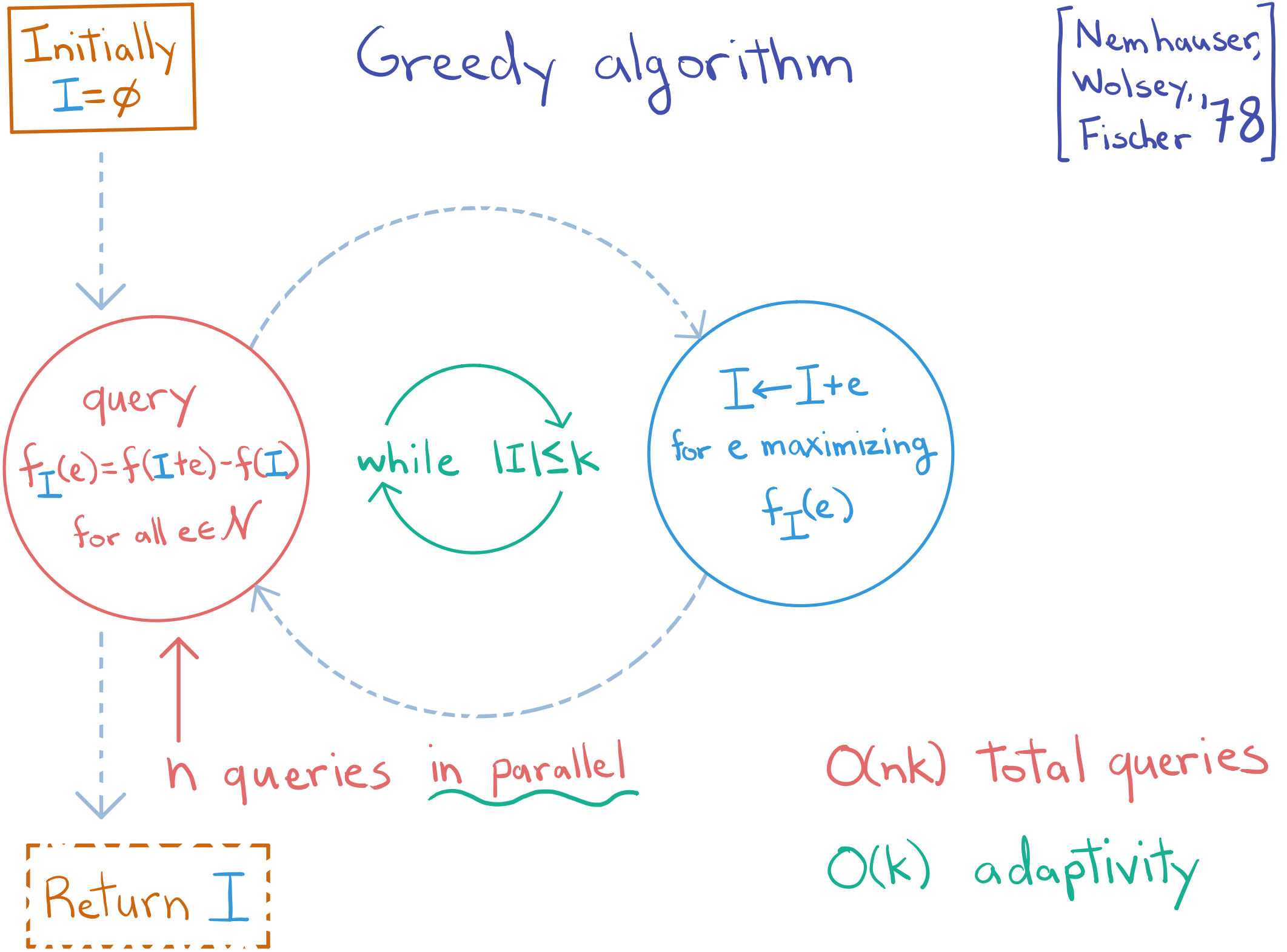
Matroids

[SODA]

$(1-\frac{1}{e}-\varepsilon)$-APX for general packing w/ low depth

Greedy algorithm

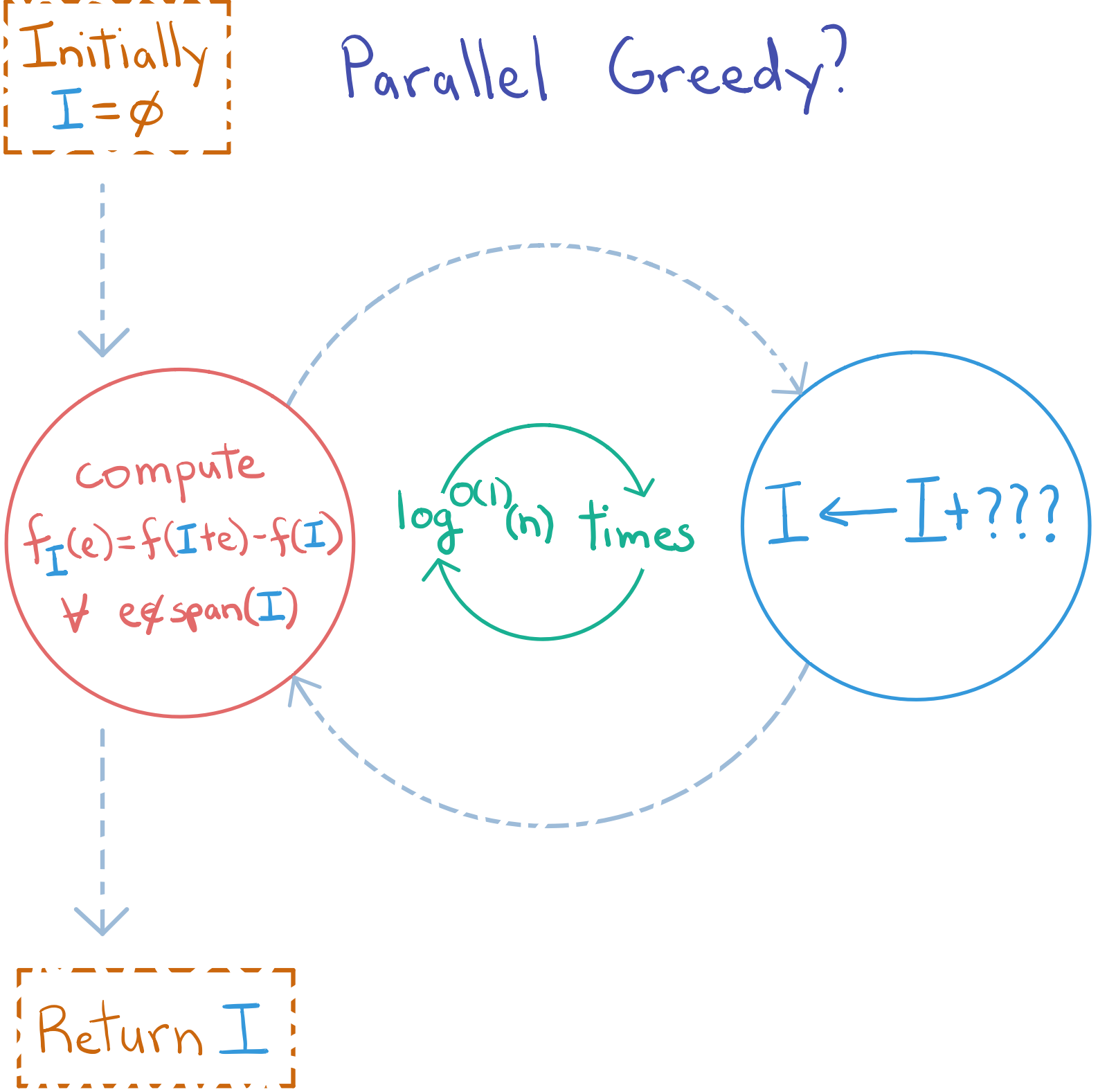
[Nemhauser, Wolsey, Fischer '78]

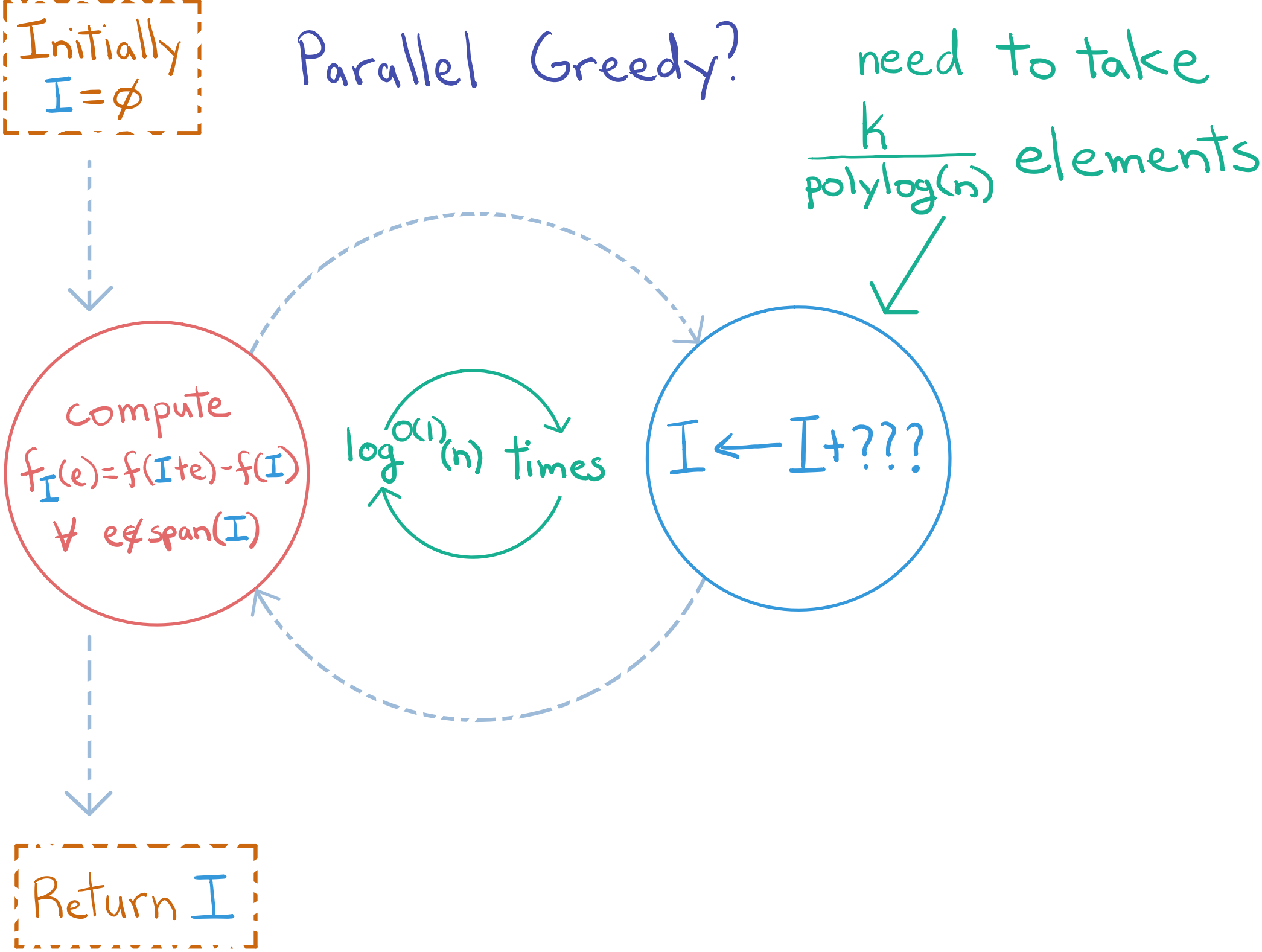


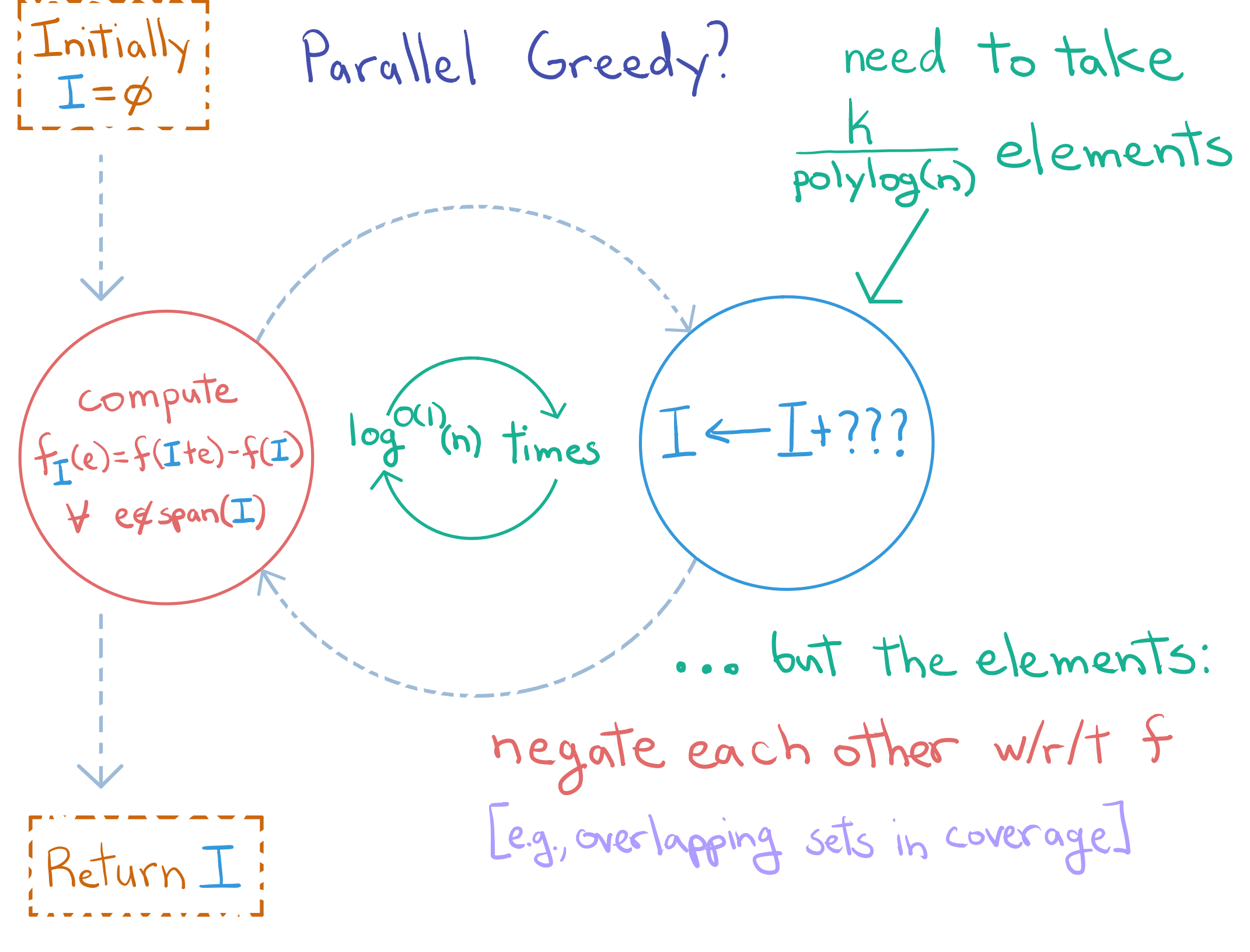
$O(nk)$ total queries

$O(k)$ adaptivity

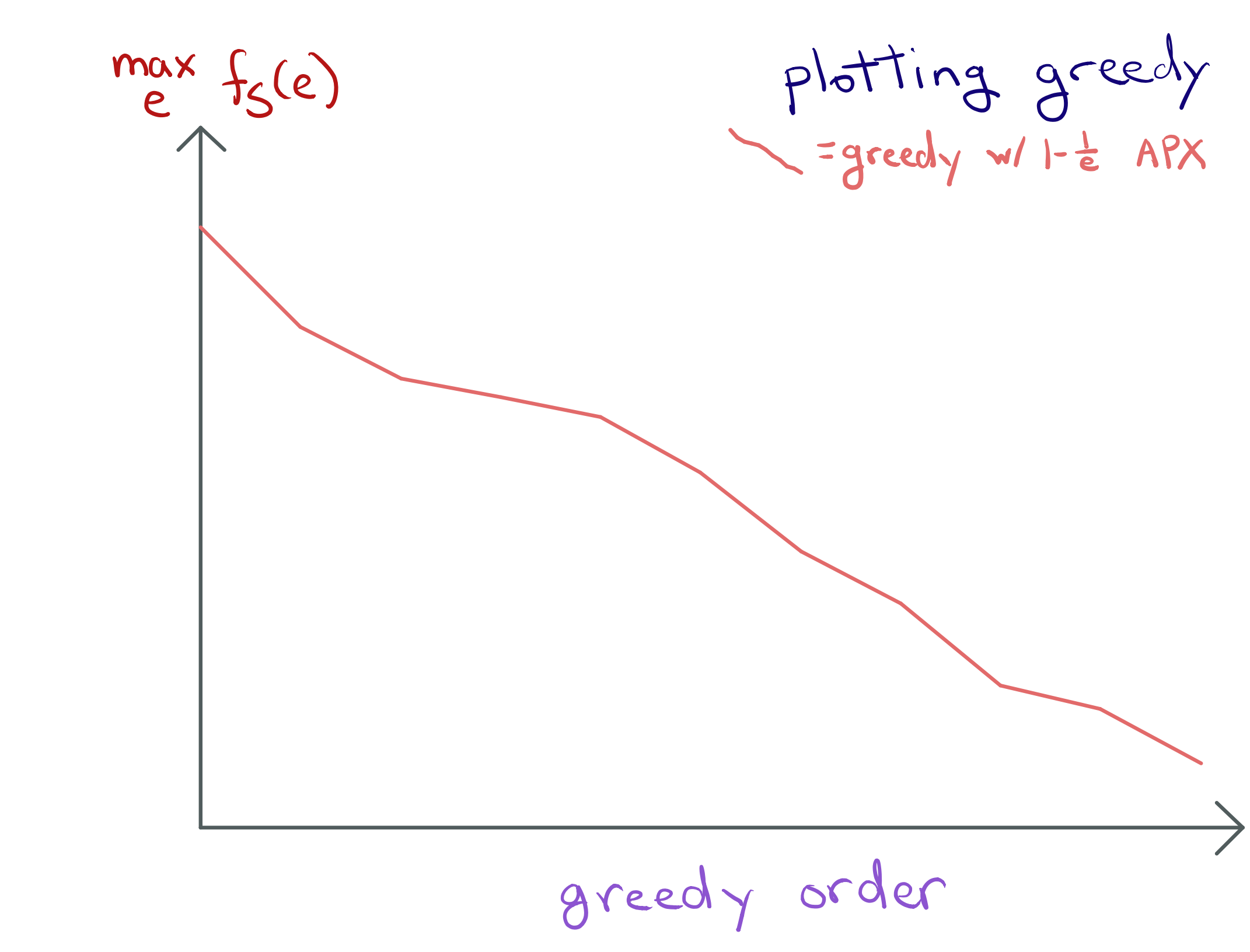
Parallel Greedy?

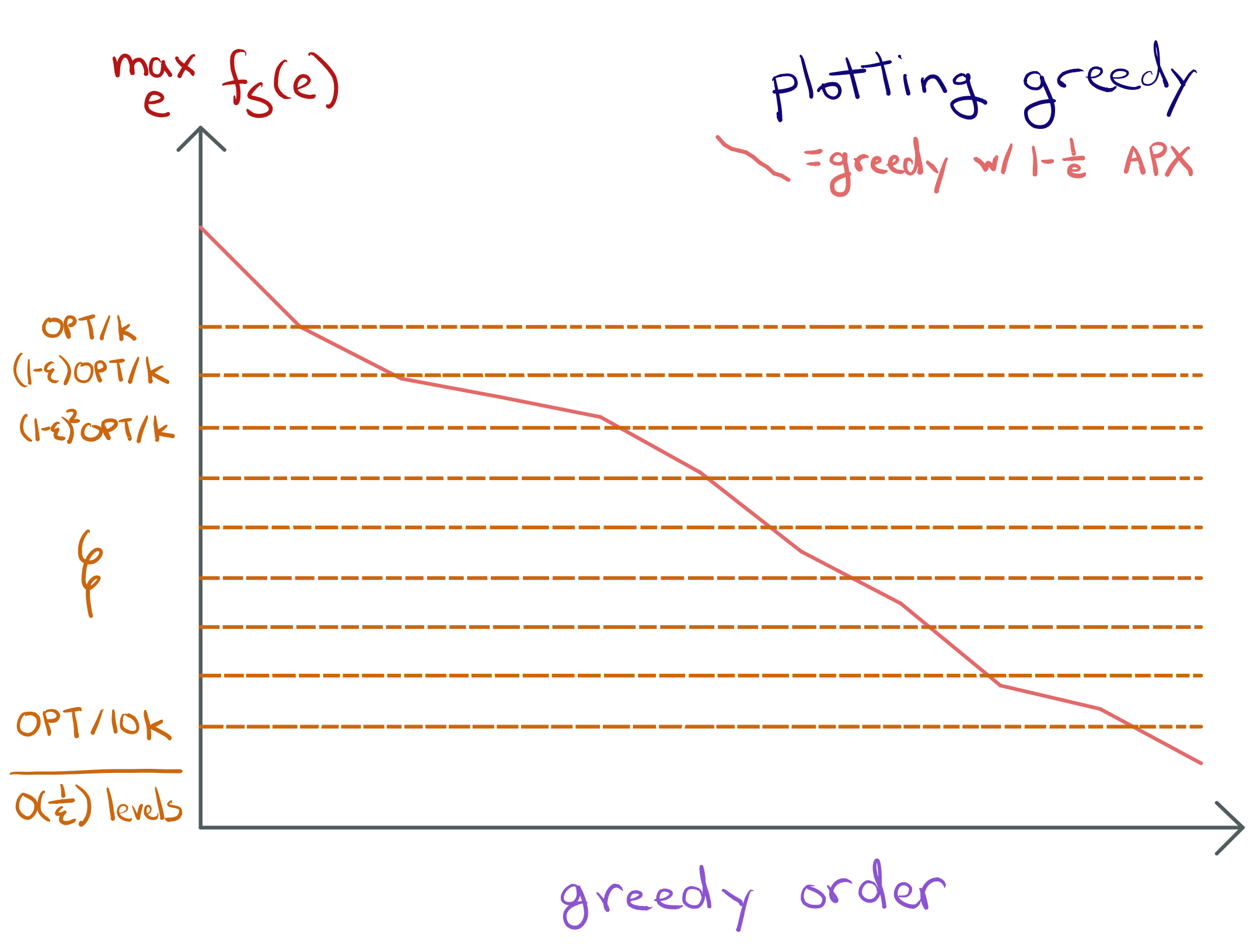


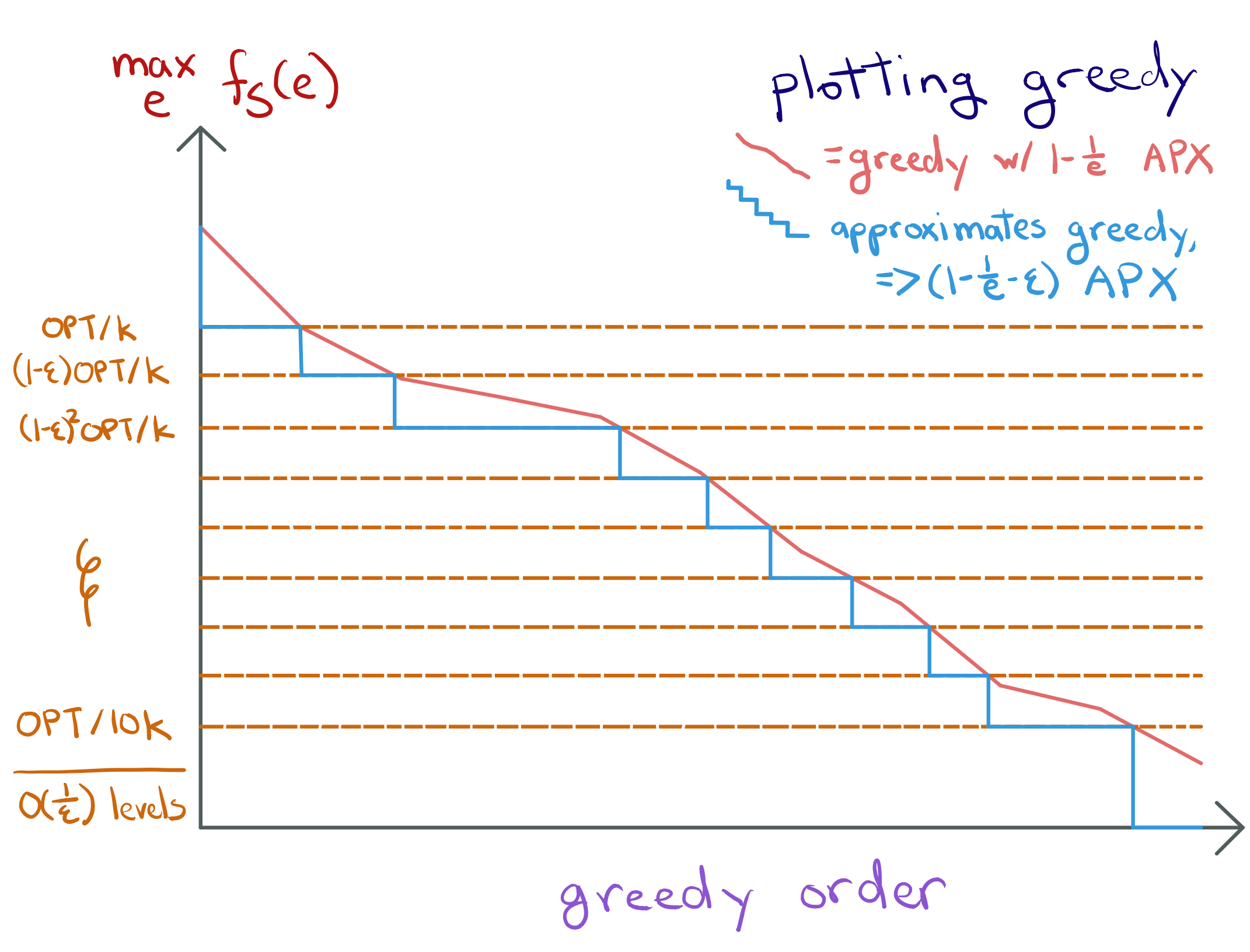

Initially
$I = \emptyset$
Parallel Greedy?
need to take
$\frac{k}{\text{polylog}(n)}$ elements
compute
$f_I(e) = f(I+e) - f(I)$
$\forall\ e \notin \text{span}(I)$
$\log^{O(1)}(n)$ times
$I \leftarrow I + ???$
Return I

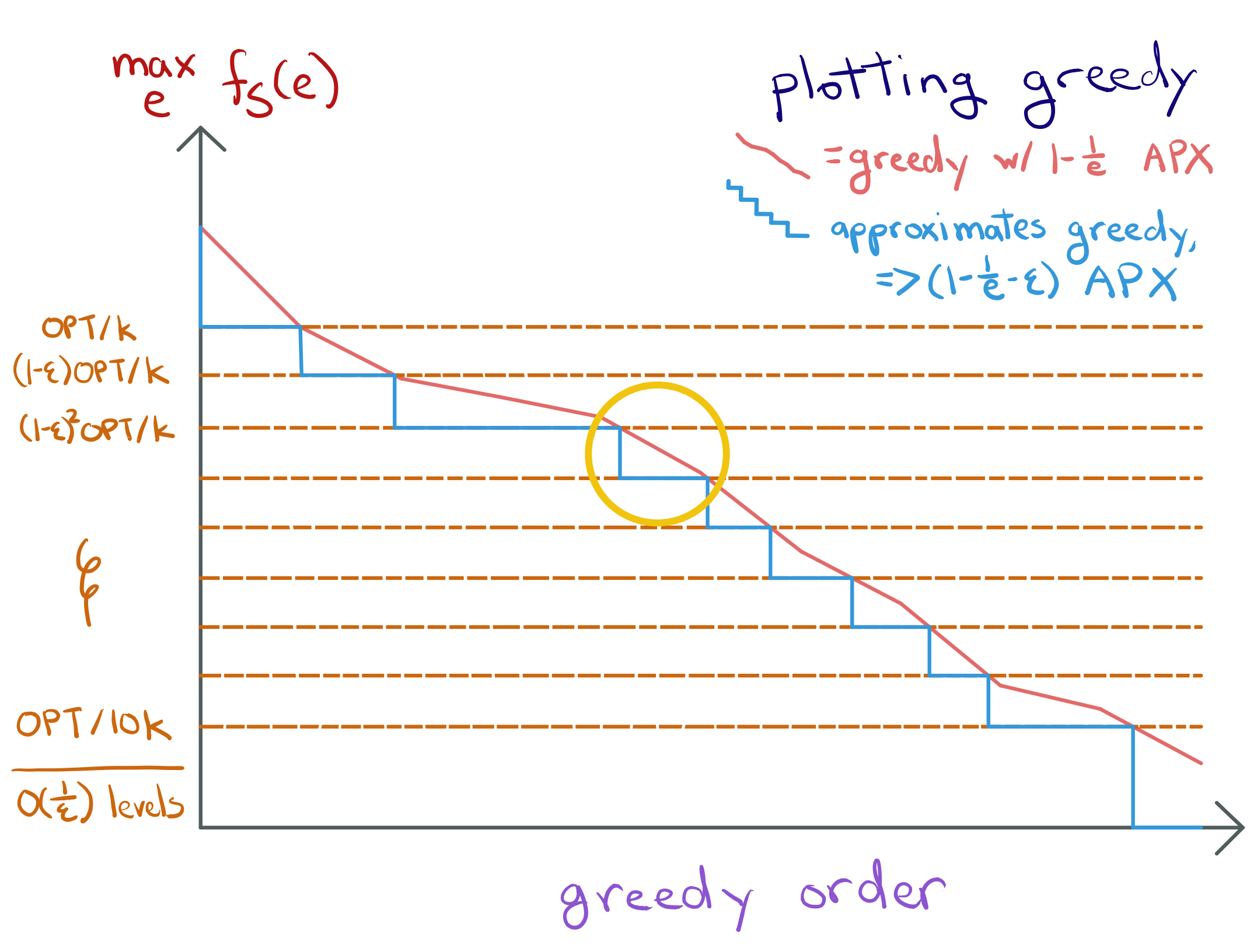

Initially
$I = \emptyset$
Parallel Greedy?
need to take
$\frac{k}{\text{polylog}(n)}$ elements
compute
$f_I(e) = f(I+e) - f(I)$
$\forall\ e \notin \text{span}(I)$
$\log^{O(1)}(n)$ times
$I \leftarrow I + ???$
... but the elements:
negate each other w/r/t f
[e.g., overlapping sets in coverage]
Return I

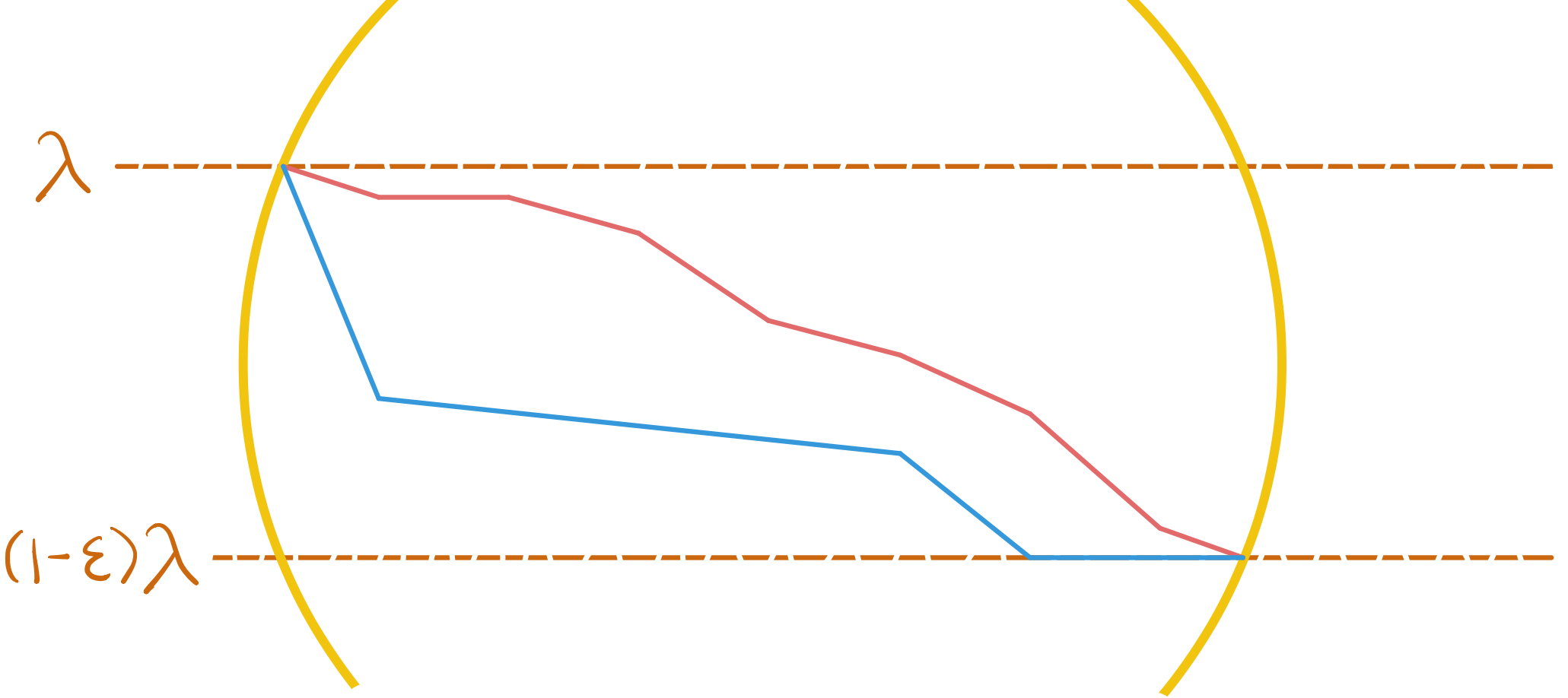
Suppose we plot
the marginal values
of greedy


$\max_e f_S(e)$
plotting greedy
= greedy w/ $1-\frac{1}{e}$ APX
greedy order

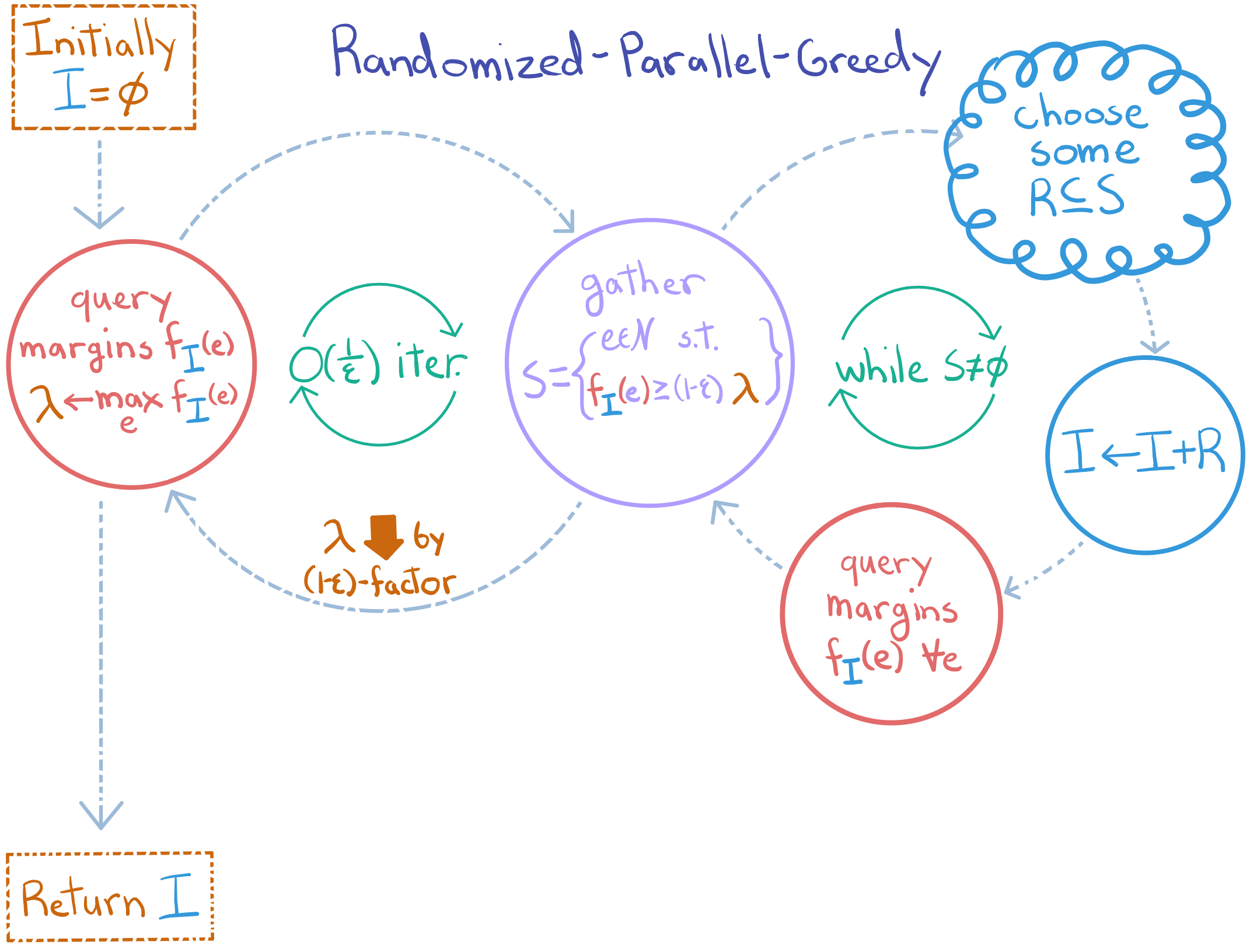

$\max_e f_S(e)$
plotting greedy
= greedy w/ $1-\frac{1}{e}$ APX
OPT/k
$(1-\varepsilon)$OPT/k
$(1-\varepsilon)^2$OPT/k
ε
OPT/10k
$O(\frac{1}{\varepsilon})$ levels
greedy order

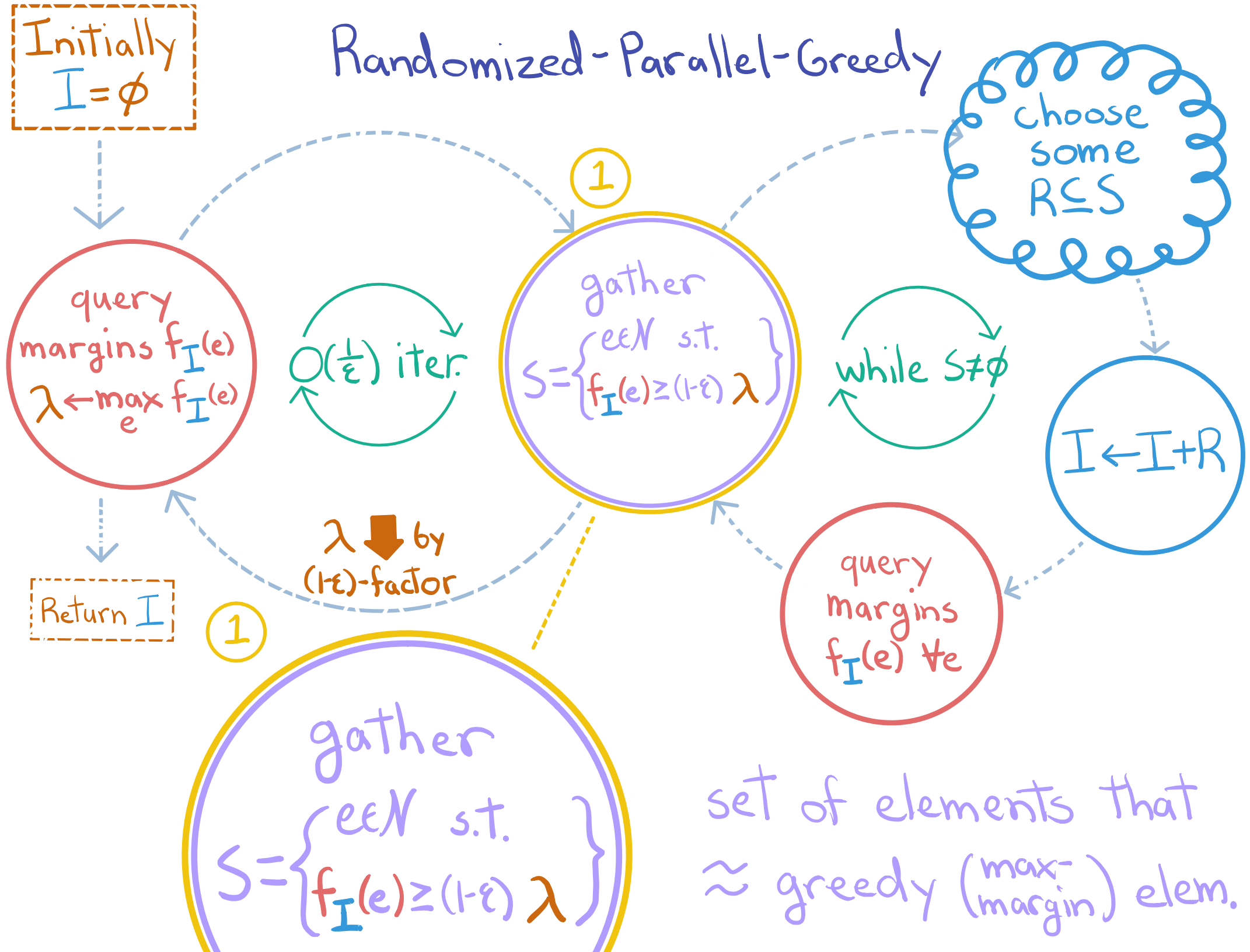

$\max_e f_S(e)$
plotting greedy
= greedy w/ $1-\frac{1}{e}$ APX
approximates greedy,
$\Rightarrow (1-\frac{1}{e}-\varepsilon)$ APX
OPT/k
$(1-\varepsilon)$OPT/k
$(1-\varepsilon)^2$OPT/k
ε
OPT/10k
$O(\frac{1}{\varepsilon})$ levels
greedy order

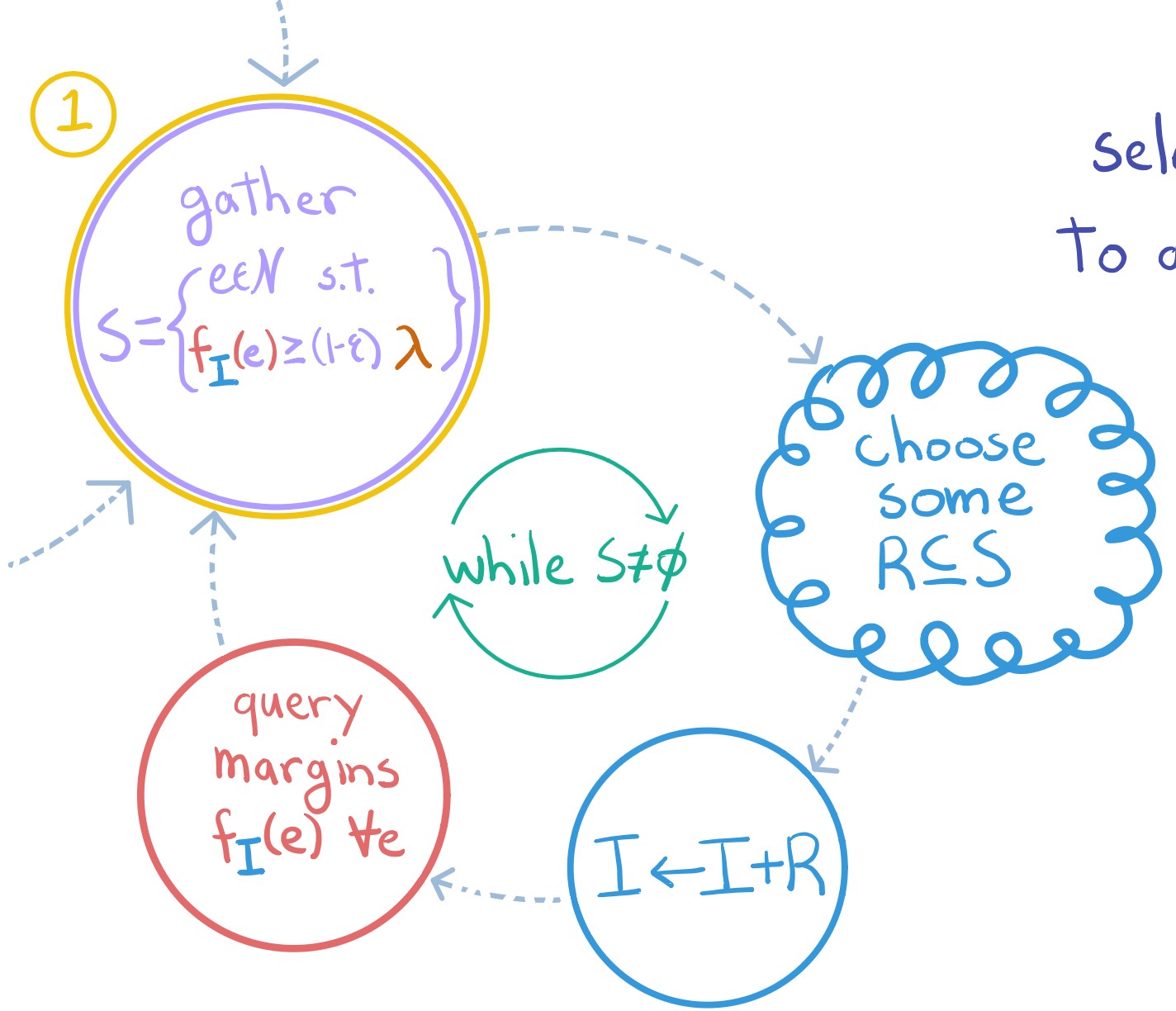

$\max_e f_S(e)$
plotting greedy
= greedy w/ $1-\frac{1}{e}$ APX
approximates greedy,
$\Rightarrow (1-\frac{1}{e}-\varepsilon)$ APX
OPT/k
$(1-\varepsilon)$OPT/k
$(1-\varepsilon)^2$OPT/k
φ
OPT/10k
$O(\frac{1}{\varepsilon})$ levels
greedy order



Goal: drive down max margin to $(1-\varepsilon)\lambda$
while taking elements w/ margin $\approx \lambda$


Randomized-Parallel-Greedy
Initially
$I = \emptyset$
query
margins $f_I(e)$
$\lambda \leftarrow \max_e f_I(e)$
$O(\frac{1}{\varepsilon})$ iter.
gather
$S = \{e \in N \text{ s.t. } f_I(e) \geq (1-\varepsilon)\lambda\}$
while $S \neq \emptyset$
choose
some
$R \subseteq S$
$I \leftarrow I + R$
query
margins
$f_I(e)$ $\forall e$
λ ⬇ by
$(1-\varepsilon)$-factor
Return I


Initially
$I = \emptyset$
Randomized-Parallel-Greedy
choose some $R \subseteq S$
query margins $f_I(e)$
$\lambda \leftarrow \max_e f_I(e)$
$O(\frac{1}{\varepsilon})$ iter.
1
gather
$S = \{e \in N$ s.t. $f_I(e) \geq (1-\varepsilon)\lambda\}$
while $S \neq \emptyset$
$I \leftarrow I + R$
λ ↓ by $(1-\varepsilon)$-factor
query margins $f_I(e)$ $\forall e$
Return I
1
gather
$S = \{e \in N$ s.t. $f_I(e) \geq (1-\varepsilon)\lambda\}$
set of elements that
$\approx$ greedy (max-margin) elem.



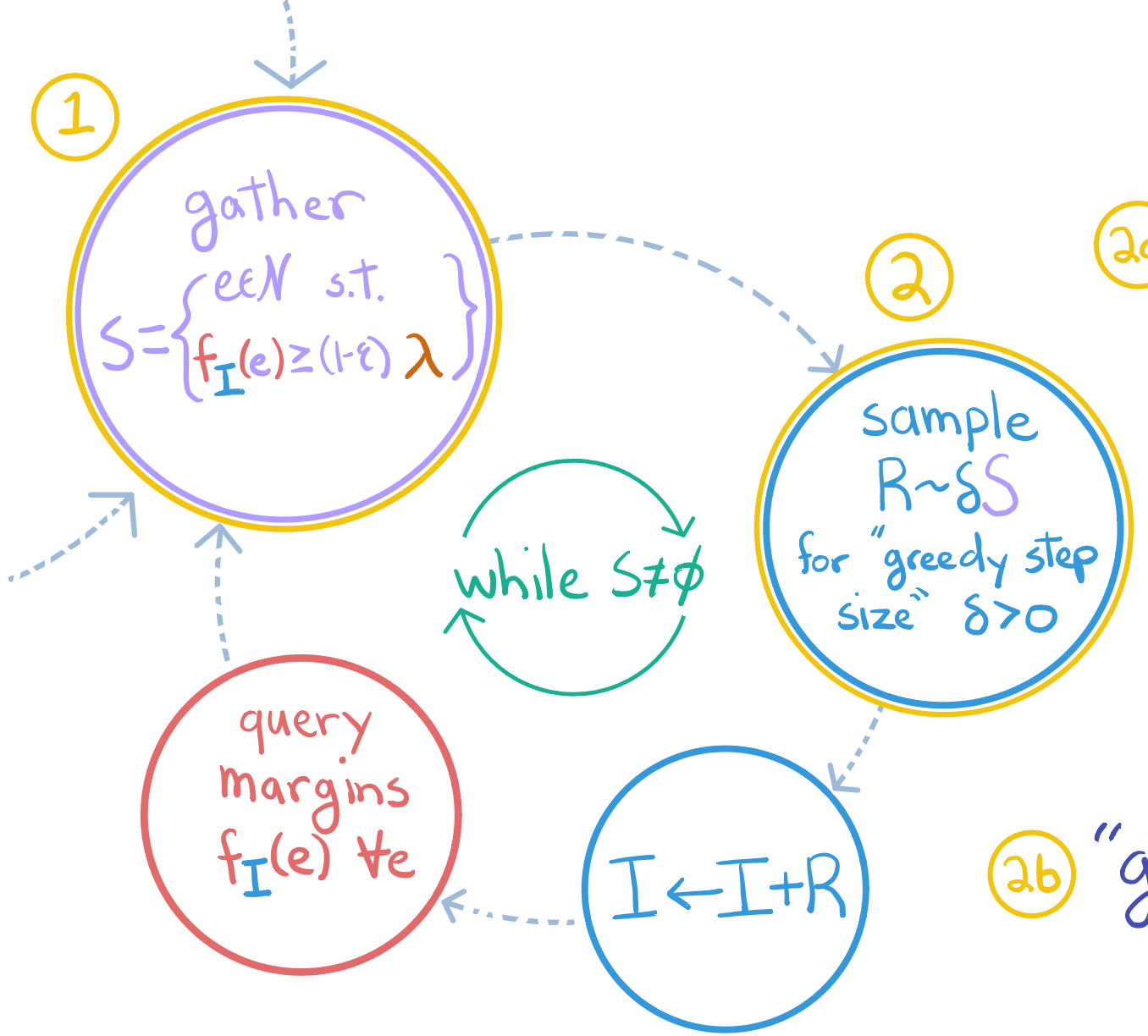
Goal

select a set $R \subseteq S$ to add to I that is both

"good"

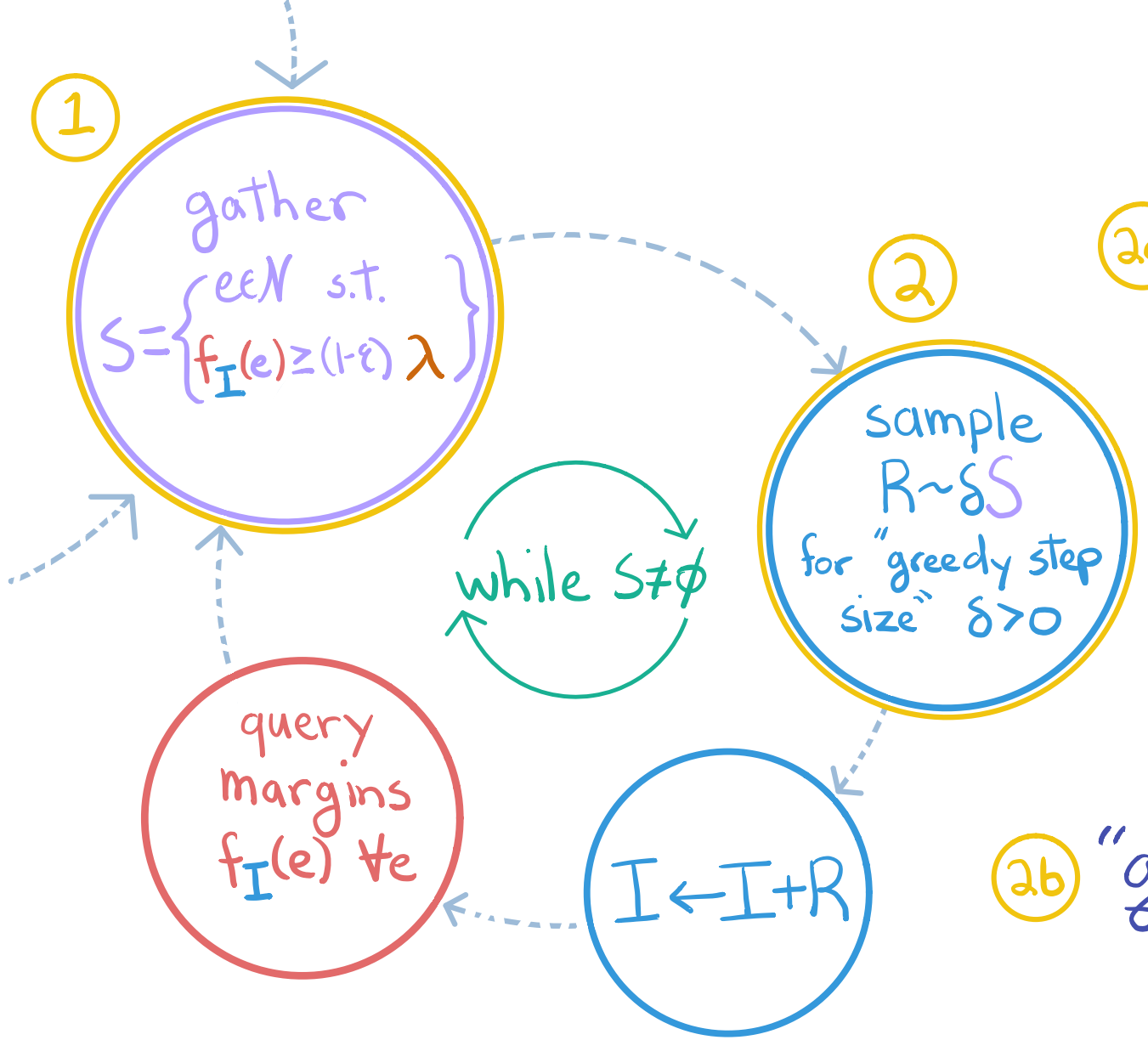
and

"large"



2a sample all $e \in S$ independently w/ some probability δ

2b "greedy step size" δ ?



2a sample all $e \in S$ independently w/ probability δ

2b "greedy step size" δ

$\max \delta > 0$ s.t

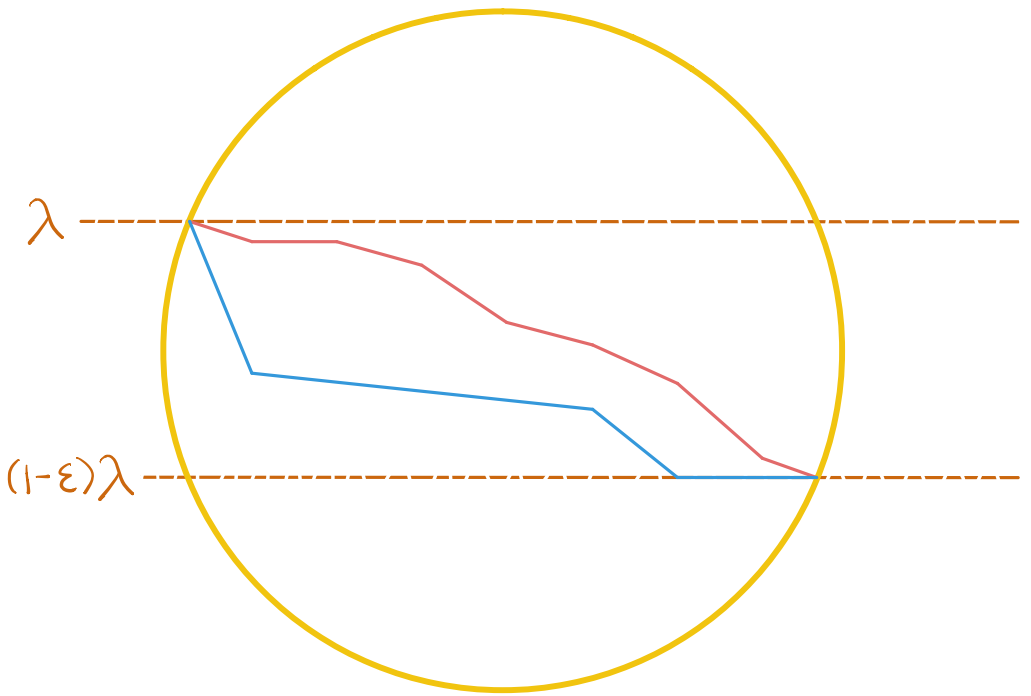
$$\frac{E[\text{bang}]}{E[\text{buck}]} = \frac{E[f_I(R)]}{\delta |S|} \geq (1-O(\varepsilon))\lambda$$

$$\max \delta > 0$$

$$\text{s.t.} \quad \frac{E[b_{ang}]}{E[b_{uck}]} = \frac{E[f_I(R)]}{\delta|S|} \geq (1 - O(\varepsilon))\,\lambda$$

$$\max \delta > 0$$

$$\text{s.t.} \quad \frac{E[b_{ang}]}{E[b_{uck}]} = \frac{E[f_I(R)]}{\delta |S|} \geq (1 - O(\varepsilon))\,\lambda$$



APX-factor follows from that we are essentially following greedy.

what about depth?

$$\max_{\delta > 0} \quad \text{s.t.} \quad \frac{E[\text{bang}]}{E[\text{buck}]} = \frac{E[f_I(R)]}{\delta |S|} \geq (1 - O(\varepsilon))\,\lambda$$

ratio ⬇ as δ ⬆ (b/c submodularity)

$$\min_{\cancel{\max}\ \delta>0} \quad \text{s.t.} \quad \boxed{\frac{E[\text{bang}]}{E[\text{buck}]} = \frac{E[f_I(R)]}{\delta|S|}} \quad \cancel{\geq} \leq (1-O(\varepsilon))\,\lambda$$

ratio ⬇ as δ ⬆ (b/c submodularity)

i.e., average margin of $e \in S$ ⬇ by $(1-\varepsilon)$

min

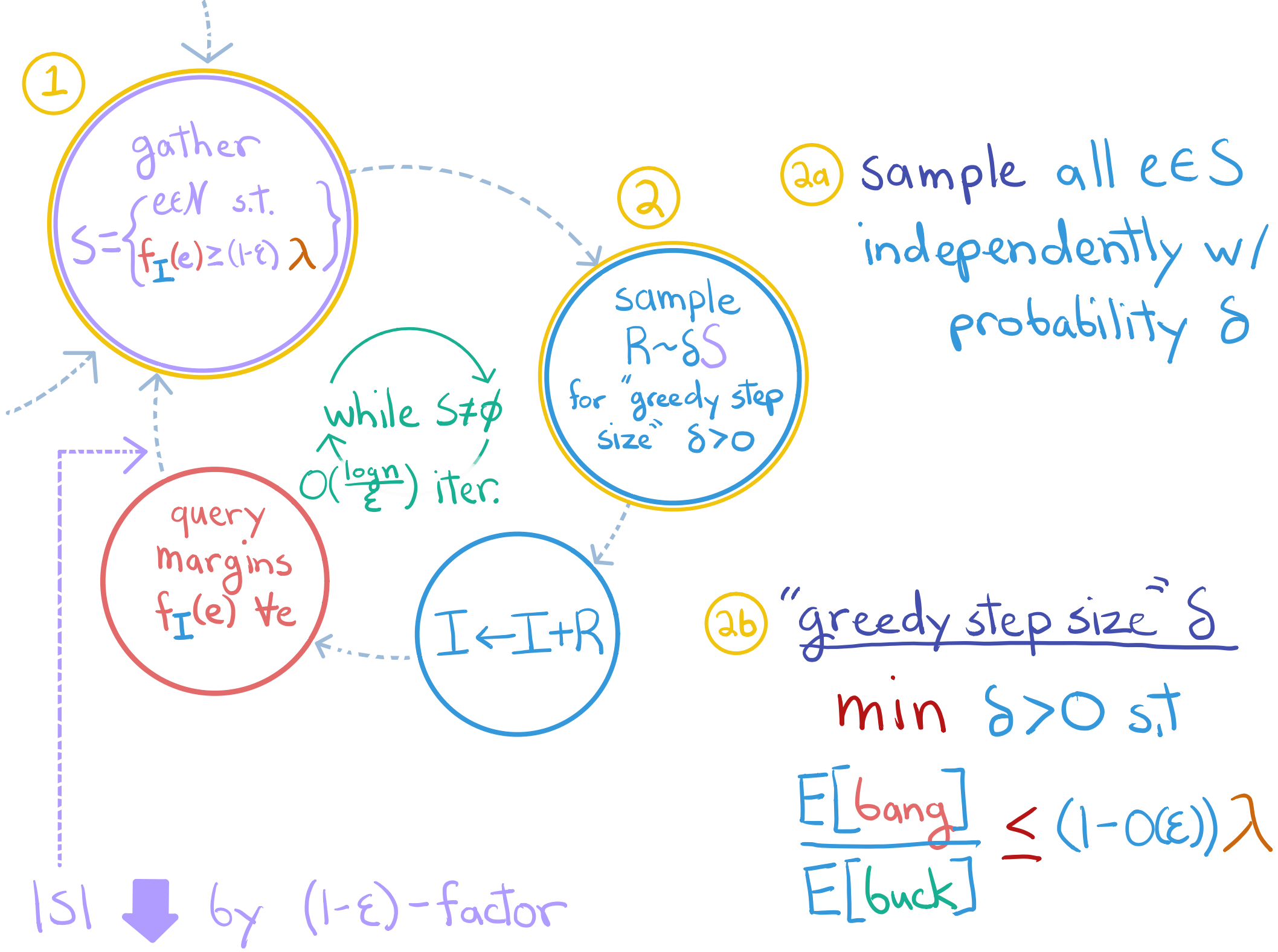
~~max~~ $\delta > 0$

s.t. $\dfrac{E[\text{bang}]}{E[\text{buck}]} = \dfrac{E[f_I(R)]}{\delta |S|}$ $\leq$ ~~$\geq$~~ $(1 - O(\varepsilon))\,\lambda$

ratio ⬇ as δ ⬆ (b/c submodularity)

i.e., average margin of $e \in S$ ⬇ by $(1-\varepsilon)$

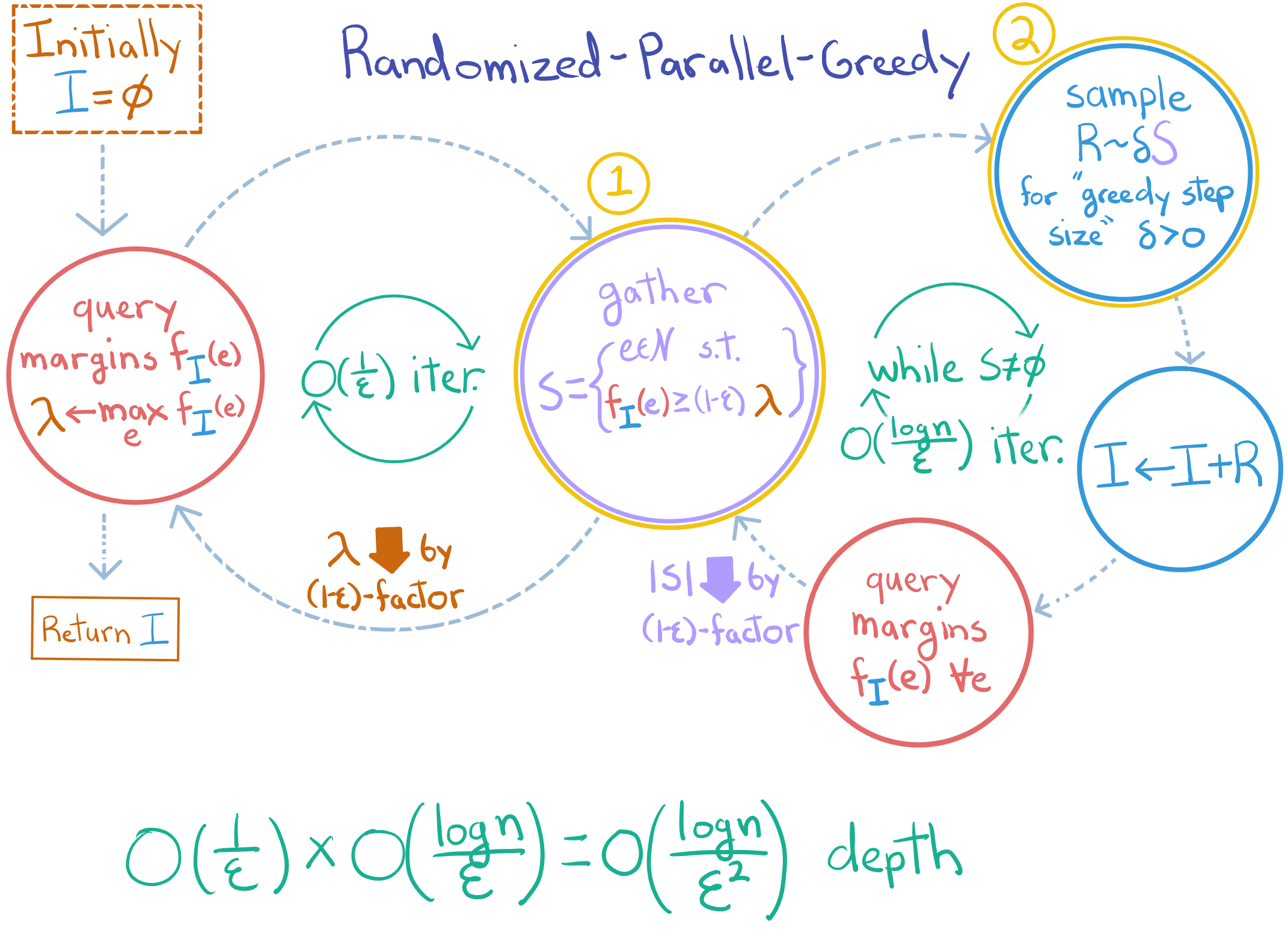
$\Rightarrow \mathcal{R} \Rightarrow$ $|S|$ ⬇ by $(1-\varepsilon)$-factor



2a sample all $e \in S$ independently w/ probability δ

2b "greedy step size" δ

$$\min \delta > 0 \text{ s.t } \frac{E[\text{bang}]}{E[\text{buck}]} \leq (1 - O(\varepsilon))\lambda$$



$$O\left(\frac{1}{\varepsilon}\right) \times O\left(\frac{\log n}{\varepsilon}\right) = O\left(\frac{\log n}{\varepsilon^2}\right) \text{ depth}$$

Parallelizing greedy for cardinality

$\frac{1}{2}$-APX for monotone f in sequential setting [Fisher, Nemhauser, Wolsey]

① thresholding

② sampling by greedy step size δ

$(1-\frac{1}{e}-\varepsilon)$-APX for monotone f w/ $\tilde{O}(\frac{1}{\varepsilon^2})$ depth

[SODA]

$(1-\frac{1}{e}-\varepsilon)$-APX for general packing w/ low depth

This talk

Matroids

Given:

- $f: 2^N \to \mathbb{R}_{\geq 0}$ submodular function
- $M = (N, \mathcal{I})$ matroid set system ($\mathcal{I} \subseteq 2^N$)

Goal

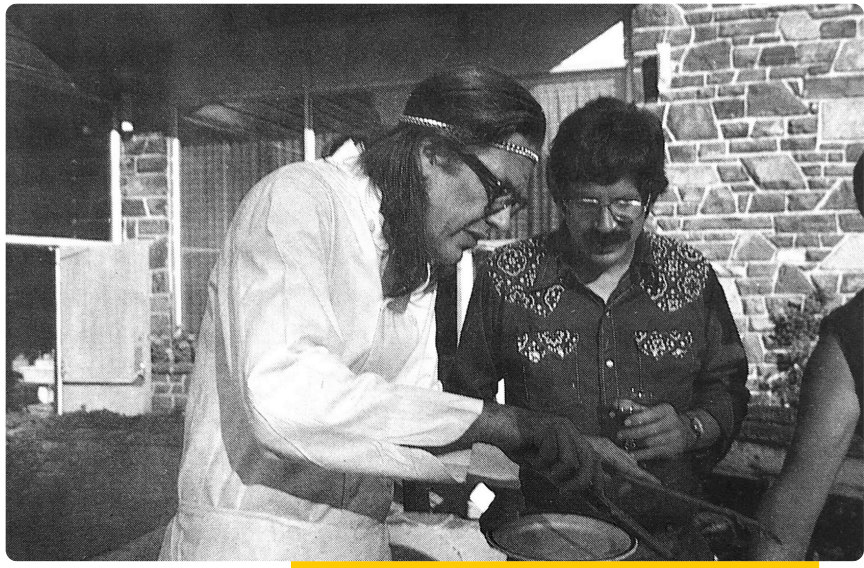
maximize $f(I)$ s.t. $I \in \mathcal{I}$

in parallel in the oracle model

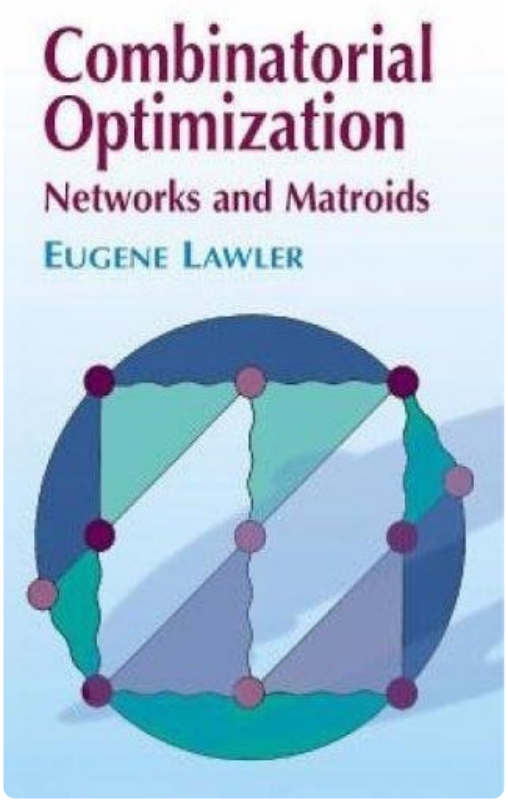
why matroids?

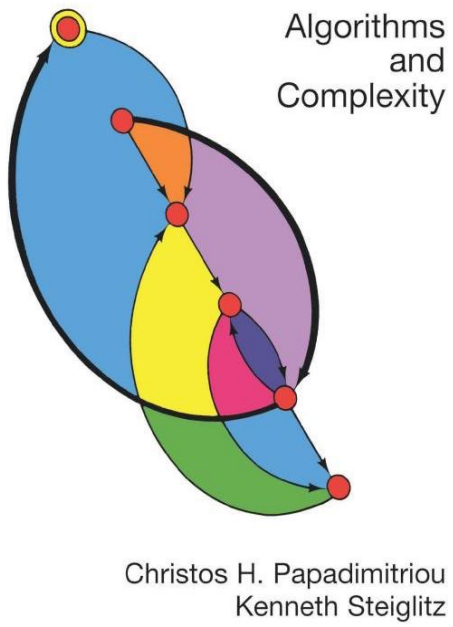
1

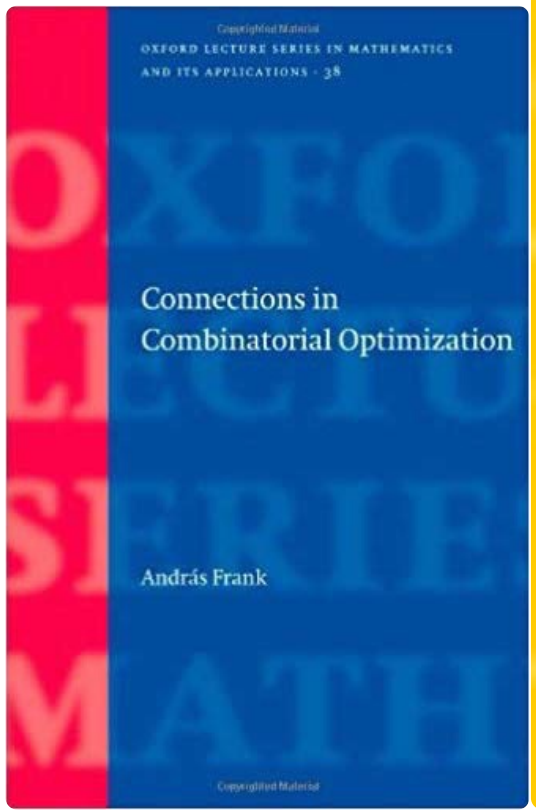
Matroids play a central & unifying role in combinatorial optimization

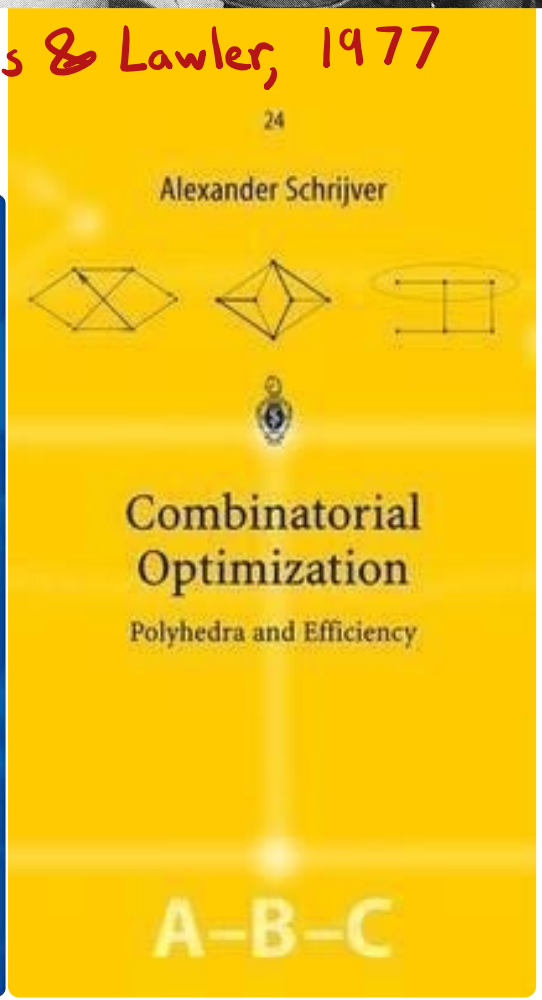


Edmonds & Lawler, 1977







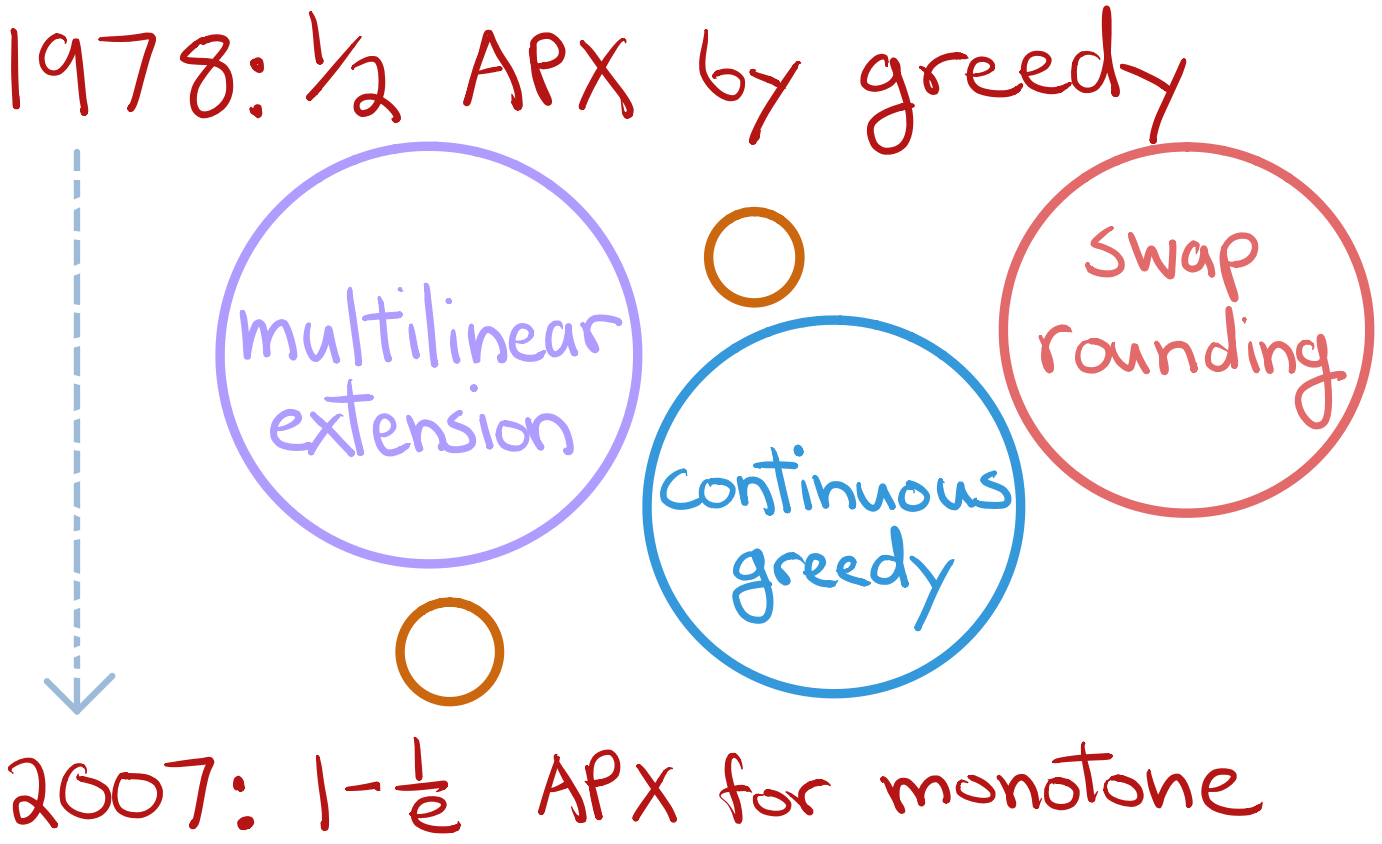


② For theory of submodular maximization, matroid constraints were a technical challenge inspiring many techniques

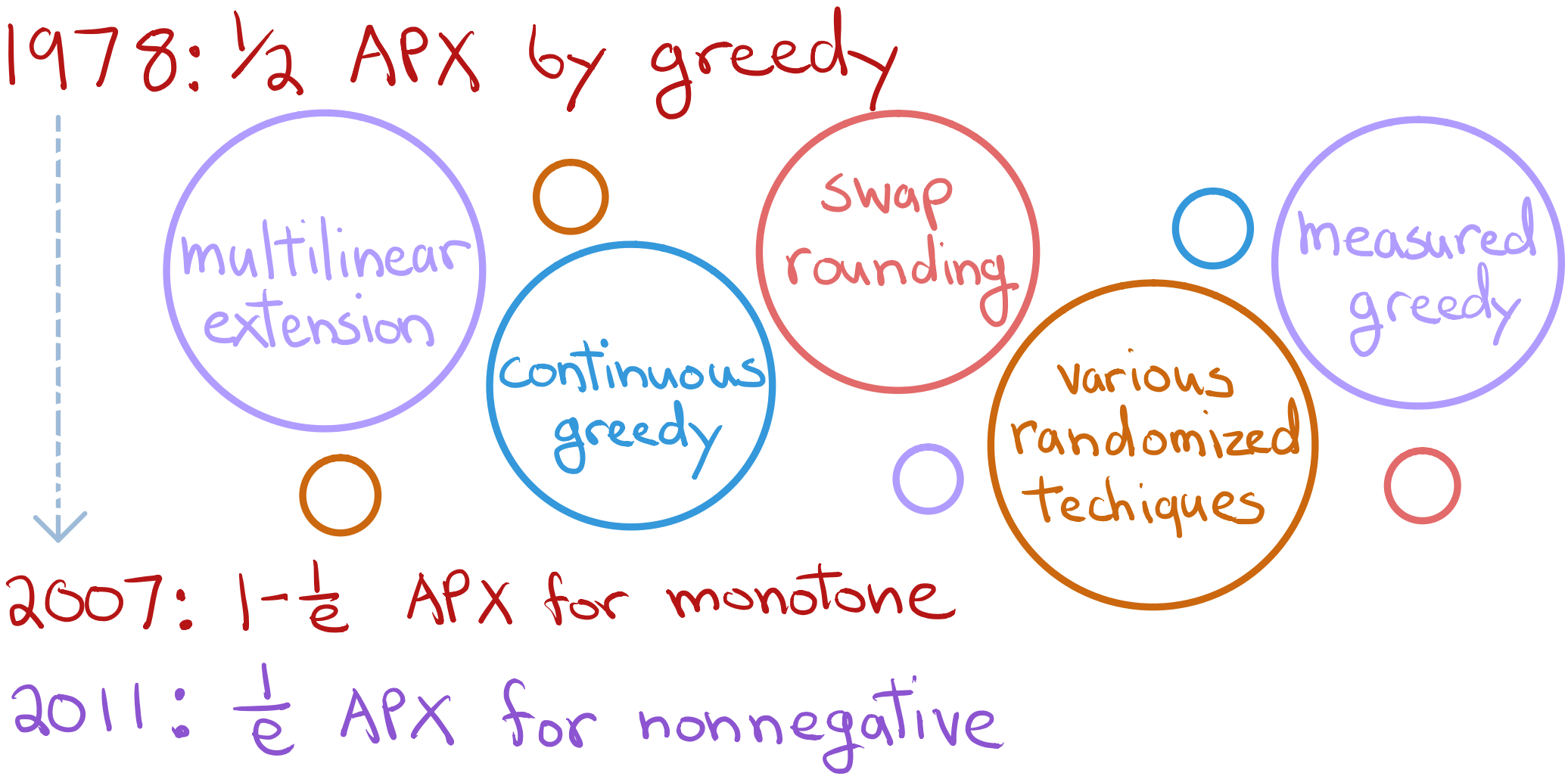
1978: ½ APX by greedy



② For theory of submodular maximization, matroid constraints were a technical challenge inspiring many techniques



② For theory of submodular maximization, matroid constraints were a technical challenge inspiring many techniques





generality

between submodular objective and matroid constraints, we get approximations for many problems at once

Matroid $(N, \mathcal{I})$

N "ground set"

$\mathcal{I} \subseteq 2^N$ "independent sets"

s.t. (1)

(2)

(3)

Matroid $(N, \mathcal{I})$

N "ground set"

$\mathcal{I} \subseteq 2^N$ "independent sets"

s.t. (1) $\emptyset \in \mathcal{I}$

(2)

(3)

Matroid $(N, \mathcal{I})$

N "ground set"

$\mathcal{I} \subseteq 2^N$ "independent sets"

s.t. (1) $\emptyset \in \mathcal{I}$

(2) hereditary:

$S \subseteq T \subseteq N,\ T \in \mathcal{I} \Rightarrow S \in \mathcal{I}$

(3)

Matroid $(N, \mathcal{I})$

N "ground set"

$\mathcal{I} \subseteq 2^N$ "independent sets"

s.t. (1) $\emptyset \in \mathcal{I}$

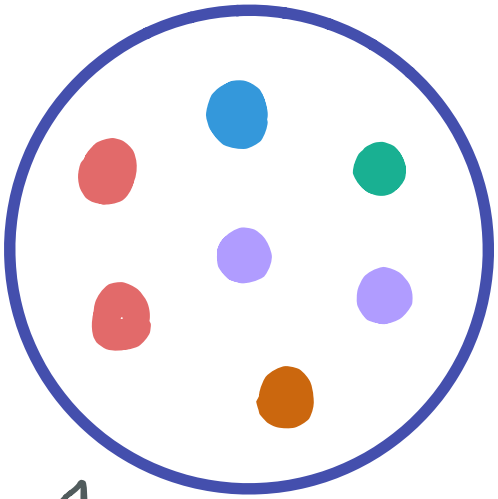
(2) hereditary:
$S \subseteq T \subseteq N,\ T \in \mathcal{I} \Rightarrow S \in \mathcal{I}$

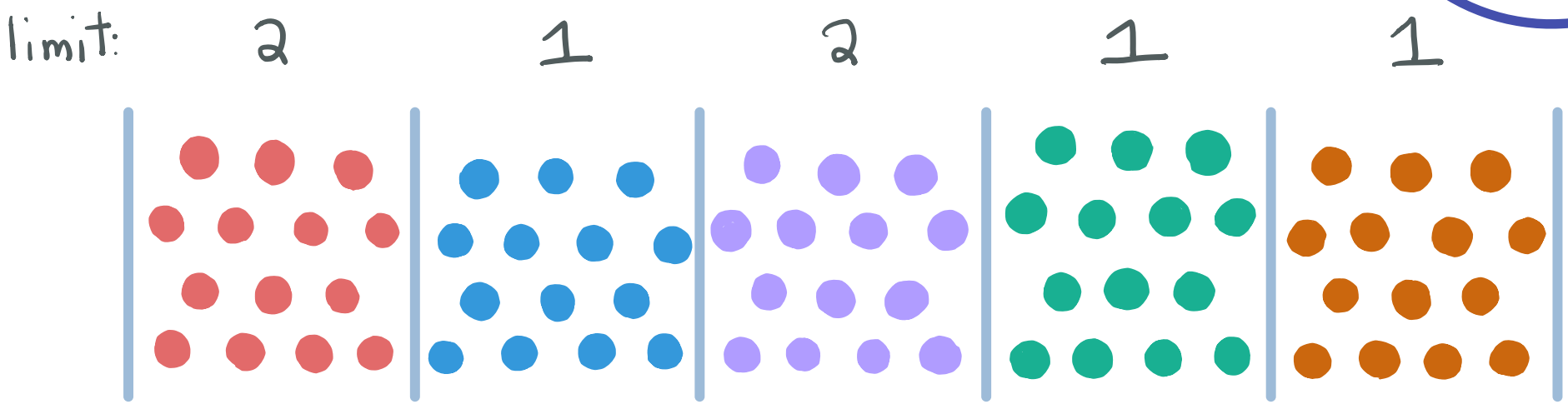
(3) extendability: $S, T \in \mathcal{I},\ |S| < |T|$
$\Rightarrow S + e \in \mathcal{I}$ for some $e \in T \setminus S$

e.g. "partition matroid"

N = set of elements w/ colors

$\mathcal{I}$ = sets w/ limited # of elements of each color

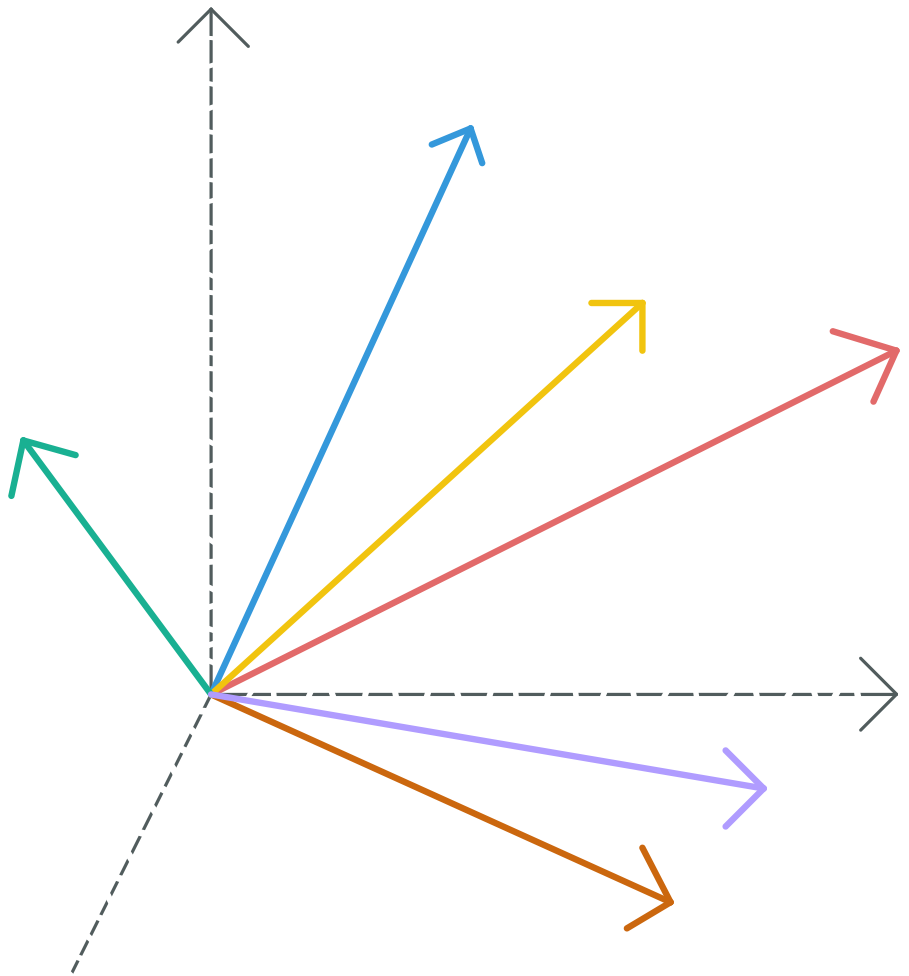




e.g. "linear matroid"

N = set of vectors in vector space

$\mathcal{I}$ = independent sets of vectors



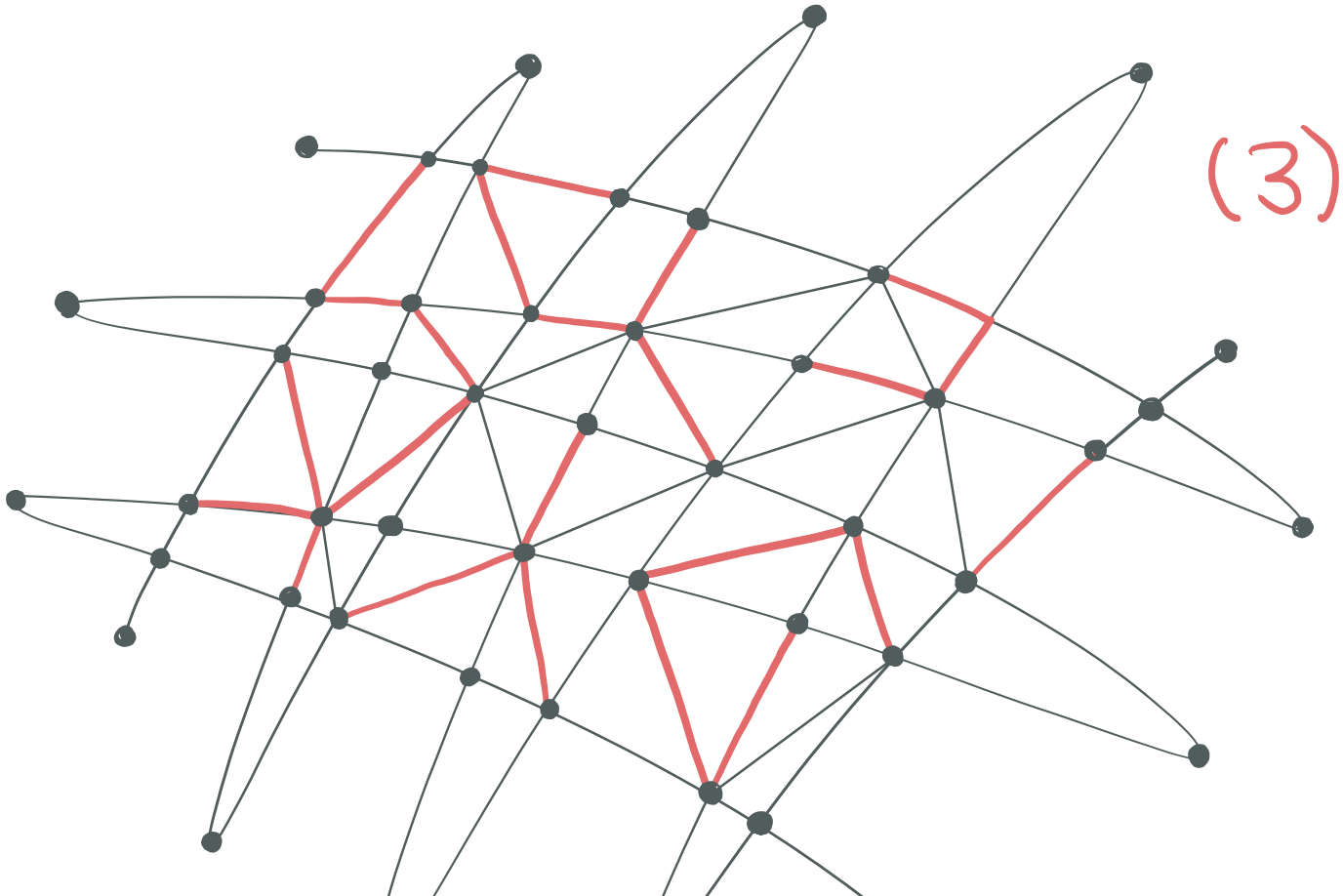
(3) extendability:

one can always extend an independent set of vectors to a base

e.g. "graphic matroid"

N = edges of a fixed graph

$\mathcal{I}$ = acyclic sets of edges (forests)



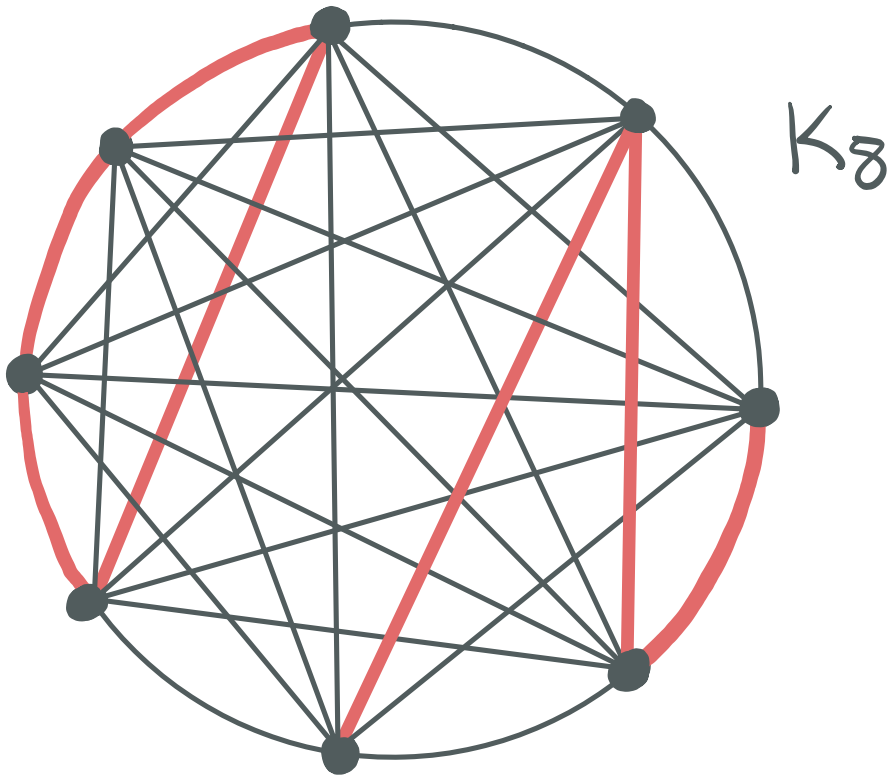
(3) extendability:
every forest in connected graph extends to a spanning tree

Matroid terminology

rank(S) [$S \subseteq N$]

max. cardinality of any independent set in S

e.g. graphic matroid



rank = 8-2=6

rank(S)

= n - # {connected components in S}

Matroid terminology

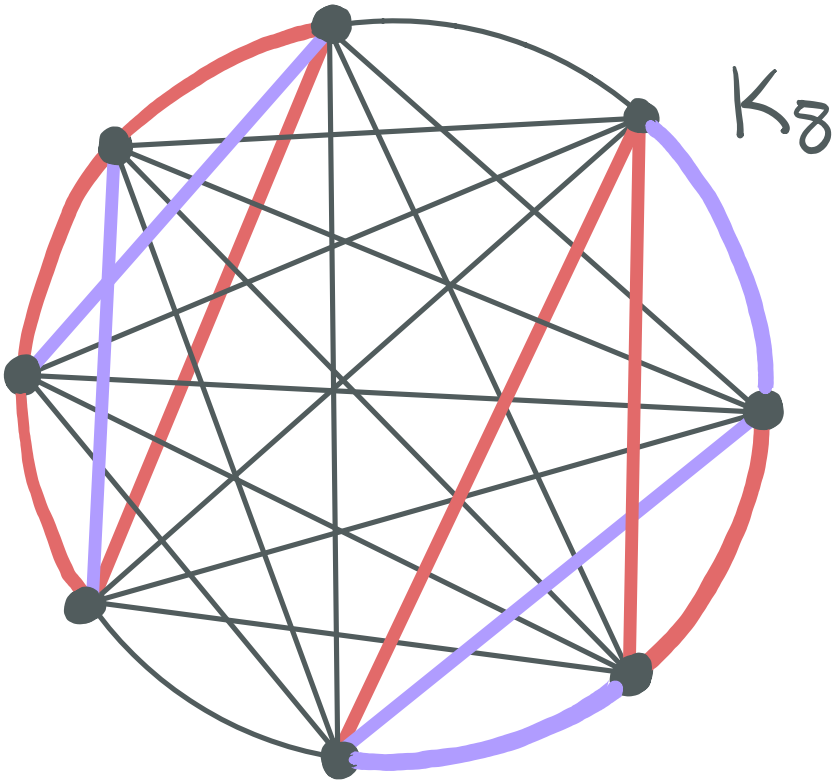
rank(S) — max. cardinality of any independent set in S

span(S) — $\{e: \text{rank}(S+e) = \text{rank}(S)\}$

e.g. graphic matroid

span(S) =

{ edges whose endpoints are connected by S }



Span Oracle

Query: $e \in \text{span}(S)$?

weaker than rank oracle

PRAM rank oracles known for linear matroids, graphic matroids

Given:

- $f: 2^N \to \mathbb{R}_{\geq 0}$ submodular function
- $M = (N, \mathcal{I})$ matroid set system ($\mathcal{I} \subseteq 2^N$)

Goal

Maximize $f(I)$ s.t. $I \in \mathcal{I}$

in parallel in the oracle model

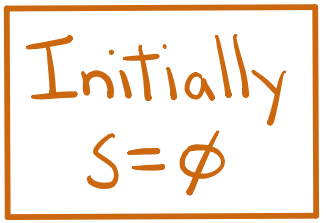
$\frac{1}{2}$-APX for monotone f [Fisher, Nemhauser, Wolsey]

this talk

$\frac{1/2 - \varepsilon}{\text{constant}}$-APX for $\frac{\text{monotone}}{\text{nonnegative}}$ f w/ $\tilde{O}\left(\frac{1}{\varepsilon^2}\right)$ depth

see paper

fractional $\frac{1 - 1/e - \varepsilon}{1/e - \varepsilon}$-APX for $\frac{\text{monotone}}{\text{nonnegative}}$ f w/ $\tilde{O}\left(\frac{1}{\varepsilon^3}\right)$ depth





Greedy algorithm

(for matroids)

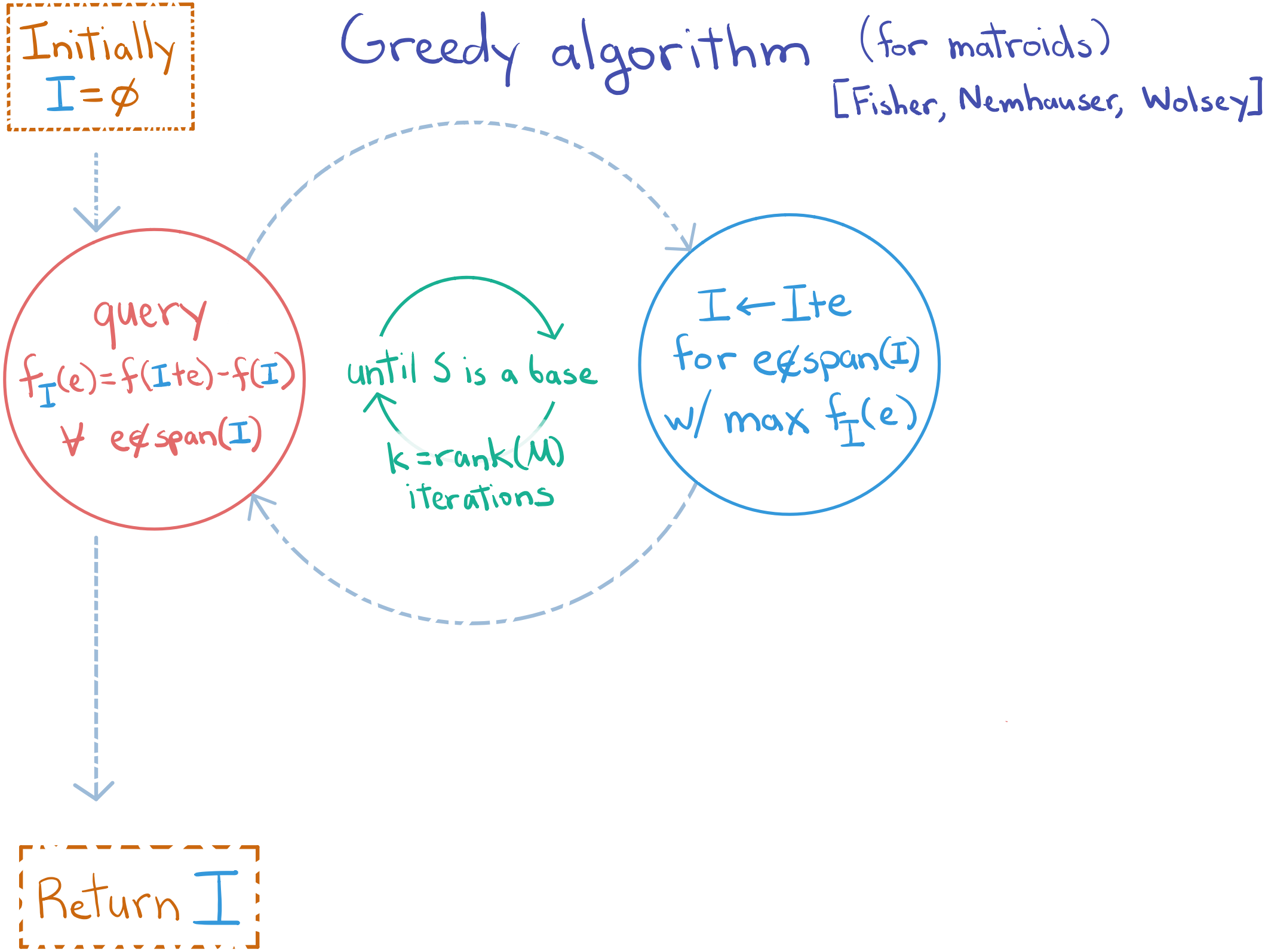
[Fisher, Nemhauser, Wolsey]

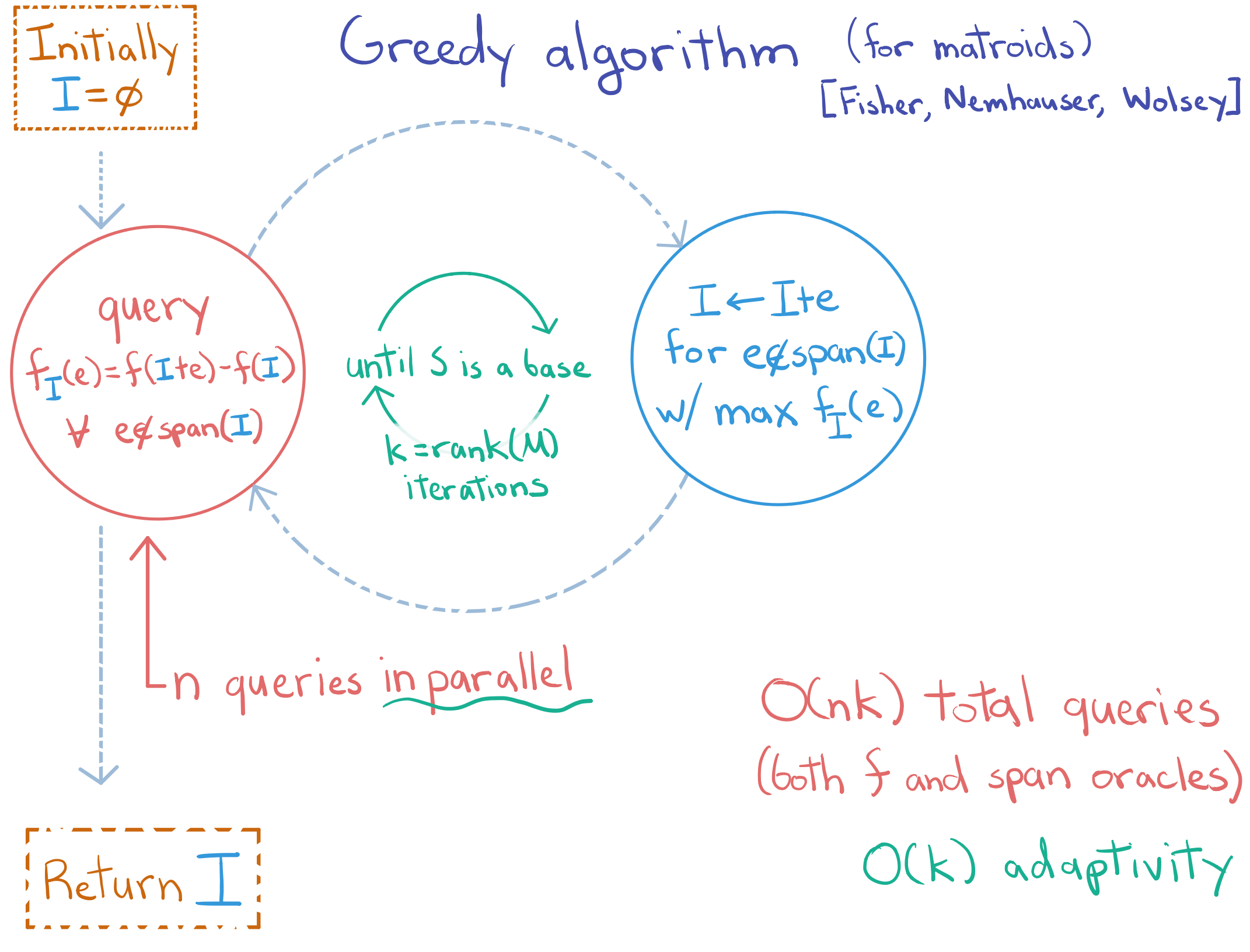
$\frac{1}{2}$-APX for monotone f

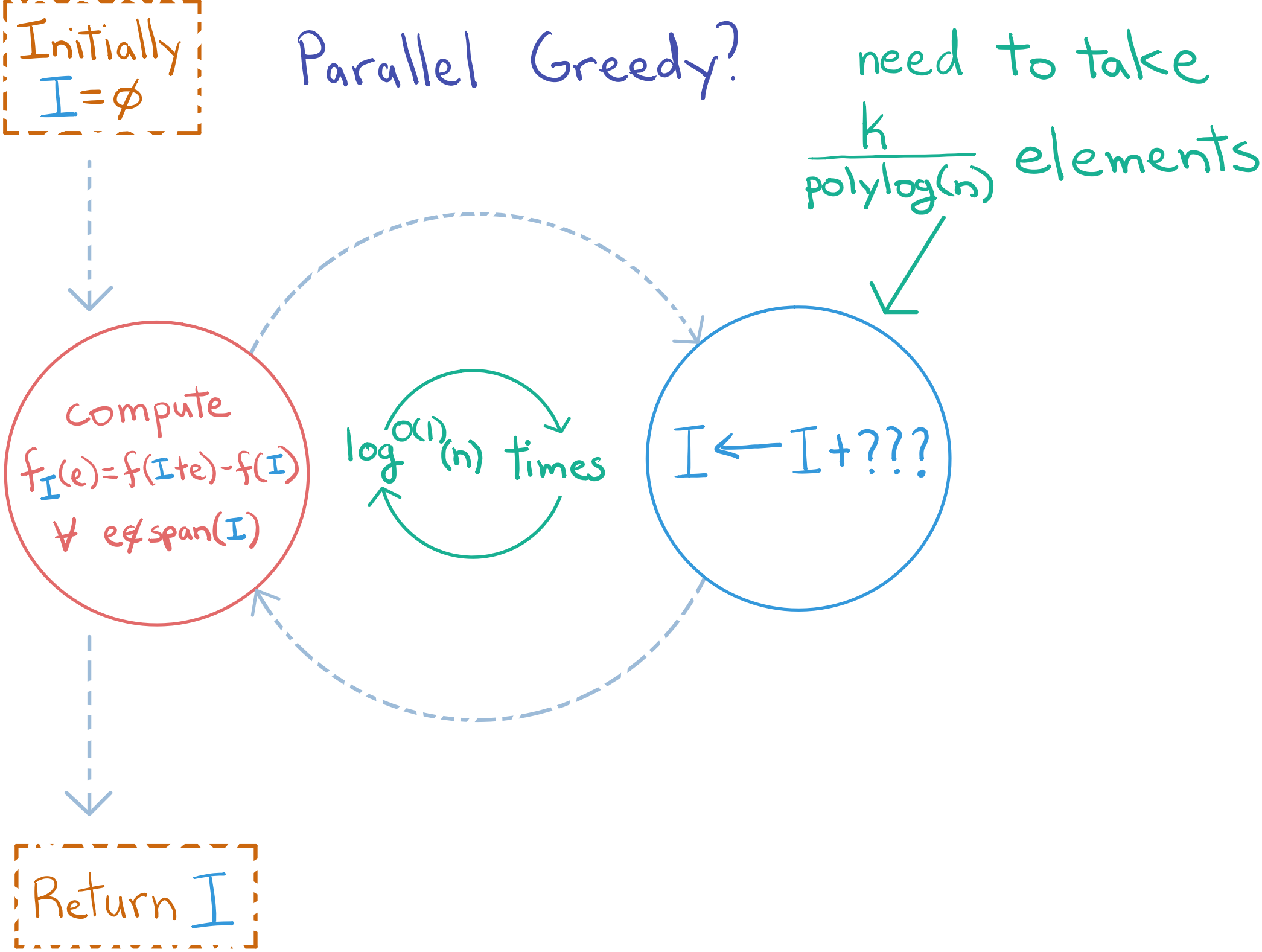


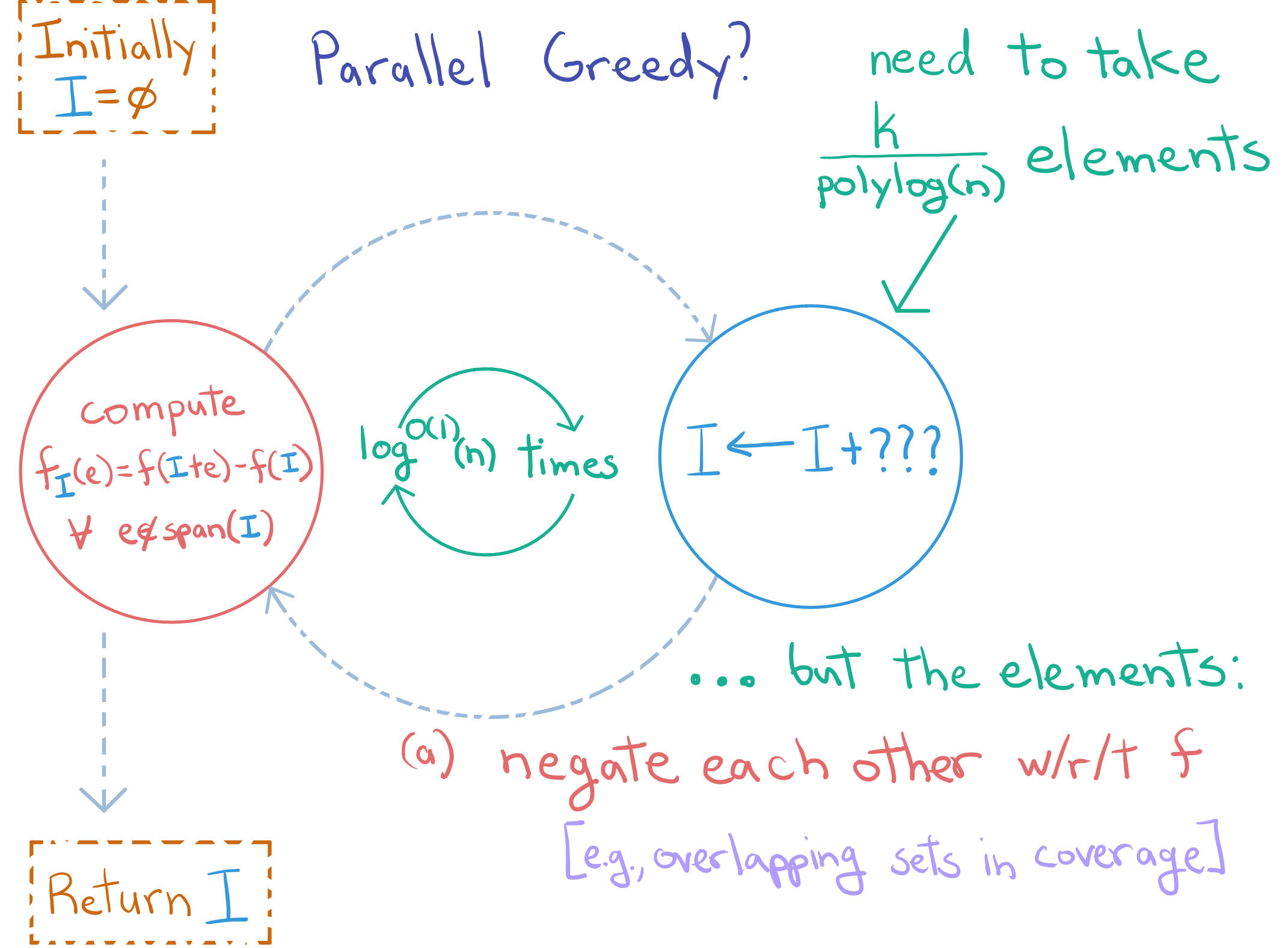
Greedy algorithm (for matroids)

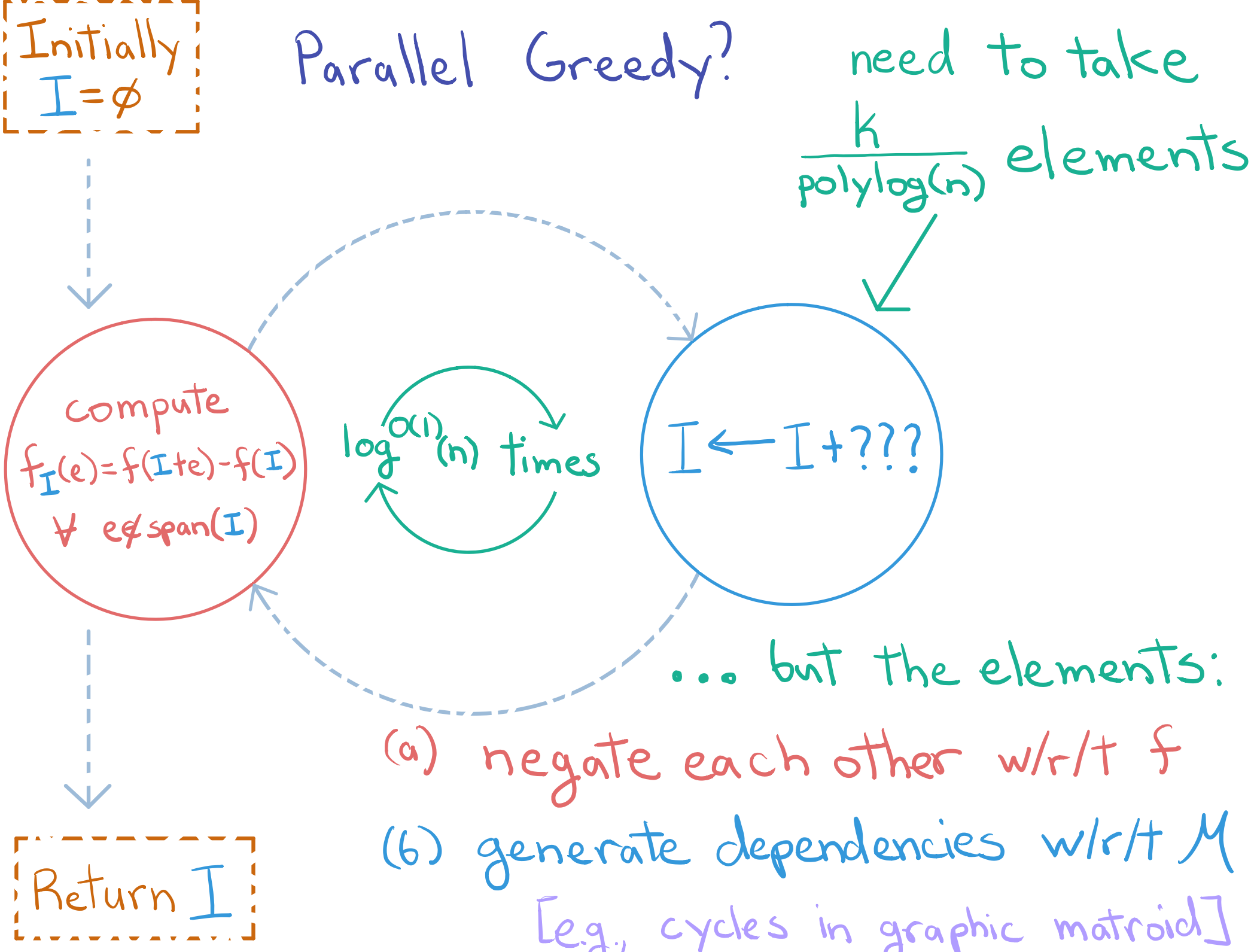
[Fisher, Nemhauser, Wolsey]

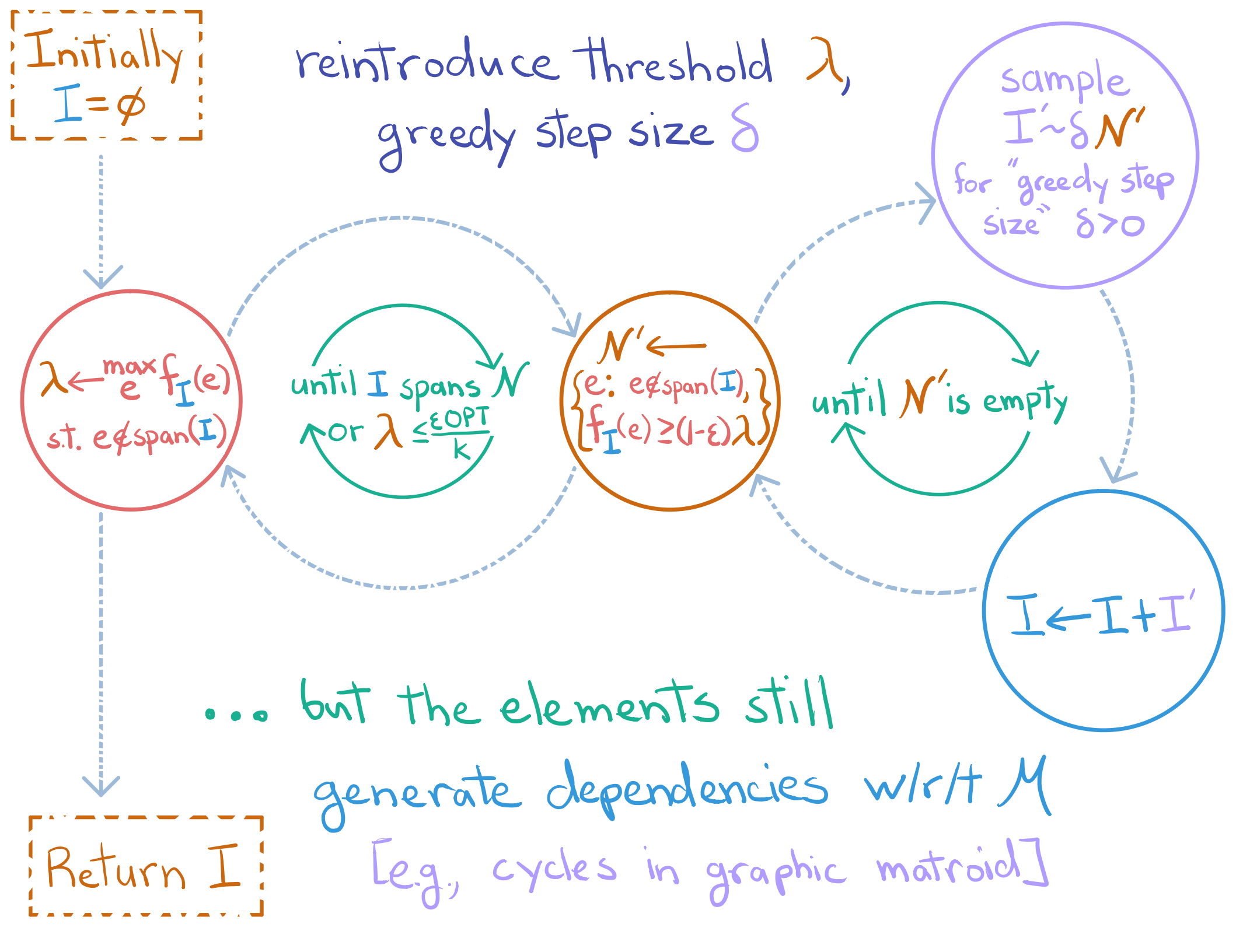


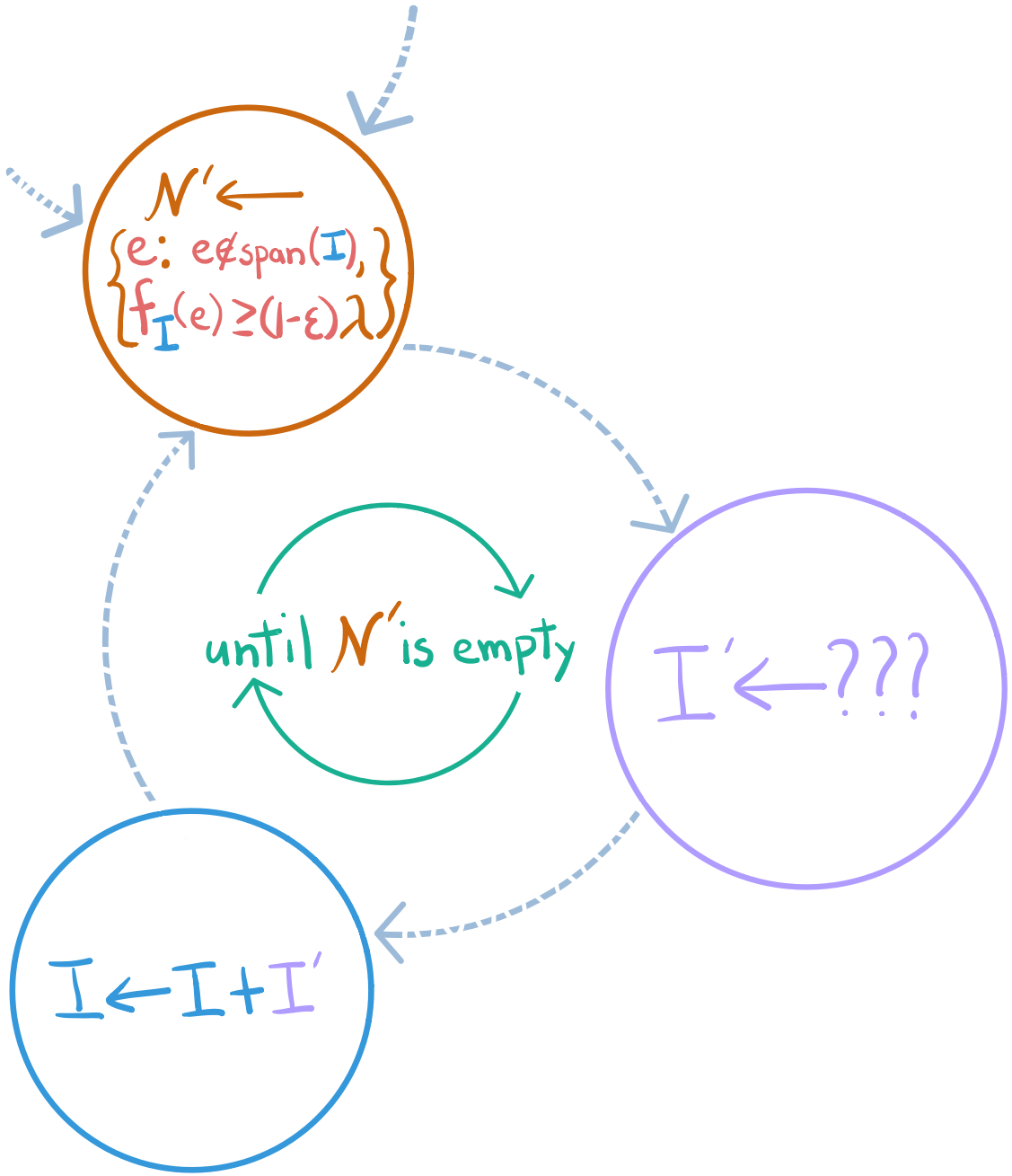

Initially
I = ∅
Greedy algorithm (for matroids)
[Fisher, Nemhauser, Wolsey]
query
$f_I(e) = f(I+e) - f(I)$
$\forall\ e \notin \text{span}(I)$
until S is a base
k = rank(M)
iterations
I ← I+e
for $e \notin \text{span}(I)$
w/ max $f_I(e)$
n queries in parallel
Return I
O(nk) total queries
(both f and span oracles)
O(k) adaptivity


Initially
$I = \emptyset$
Parallel Greedy?
need to take
$\frac{k}{\text{polylog}(n)}$ elements
compute
$f_I(e) = f(I+e) - f(I)$
$\forall\ e \notin \text{span}(I)$
$\log^{O(1)}(n)$ times
$I \leftarrow I + ???$
Return I


Initially
I = ∅
Parallel Greedy?
need to take
k / polylog(n) elements
compute
$f_I(e) = f(I+e) - f(I)$
∀ e ∉ span(I)
$\log^{O(1)}(n)$ times
I ← I + ???
... but the elements:
(a) negate each other w/r/t f
[e.g., overlapping sets in coverage]
Return I


Initially
$I = \emptyset$
Parallel Greedy?
need to take
$\frac{k}{\text{polylog}(n)}$ elements
compute
$f_I(e) = f(I+e) - f(I)$
$\forall\ e \notin \text{span}(I)$
$\log^{O(1)}(n)$ times
$I \leftarrow I + ???$
Return I
... but the elements:
(a) negate each other w/r/t f
(b) generate dependencies w/r/t M
[e.g., cycles in graphic matroid]


Initially
$I = \emptyset$
reintroduce threshold λ,
greedy step size δ
sample
$I' \sim \delta N'$
for "greedy step size" $\delta > 0$
$\lambda \leftarrow \max_e f_I(e)$
s.t. $e \notin \mathrm{span}(I)$
until I spans N
or $\lambda \leq \frac{\varepsilon OPT}{k}$
$N' \leftarrow$
$\{e: e \notin \mathrm{span}(I), f_I(e) \geq (1-\varepsilon)\lambda\}$
until N' is empty
$I \leftarrow I + I'$
... but the elements still
generate dependencies w/r/t M
[e.g., cycles in graphic matroid]
Return I



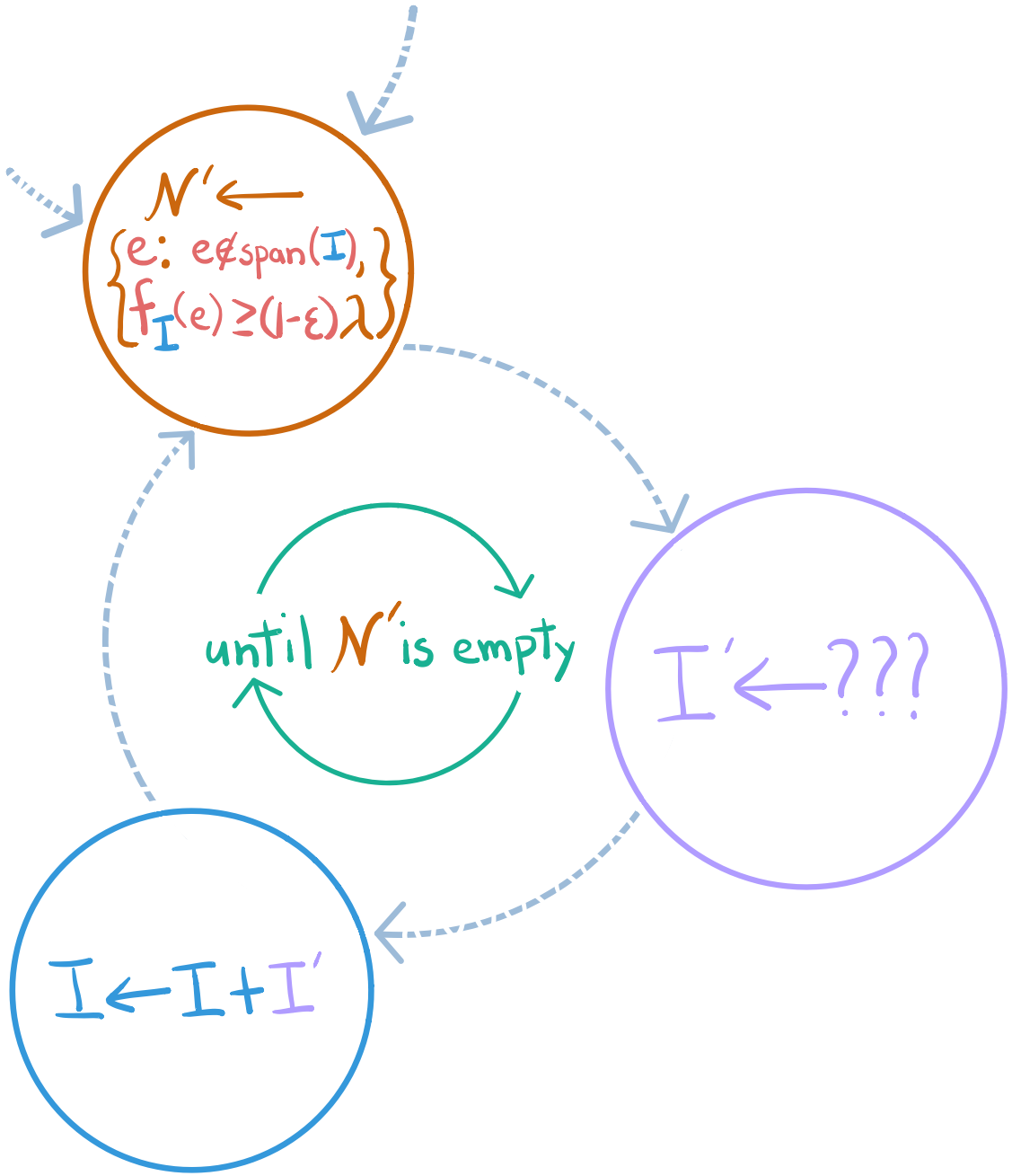
Goal

select a set $I' \subseteq N'$ to add to I that is both

"good"

and

"large"



Goal

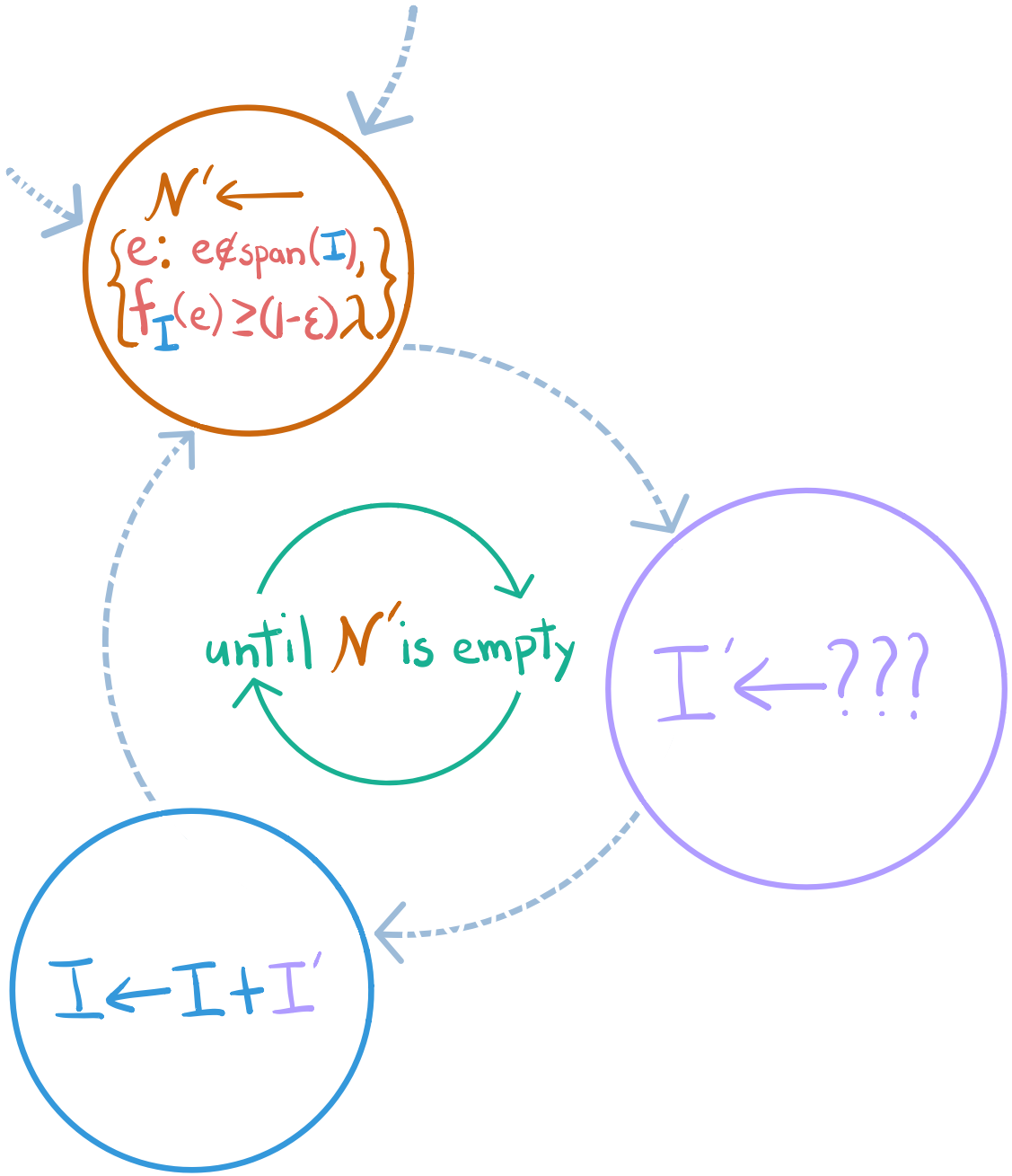
select a set $I' \subseteq N'$ s.t.

(a) $I + I'$ is independent and $f_I(I') \approx \lambda |I'|$

(b) $|N'|$ ⬇ a lot w/r/t $I + I'$ by combination of:

(i) $I + I'$ spans many $e \in N'$

(ii) $f_{I+I'}(e) \ll f_I(e)$ for many $e \in N'$



Goal

select a set $I' \subseteq N'$ s.t.

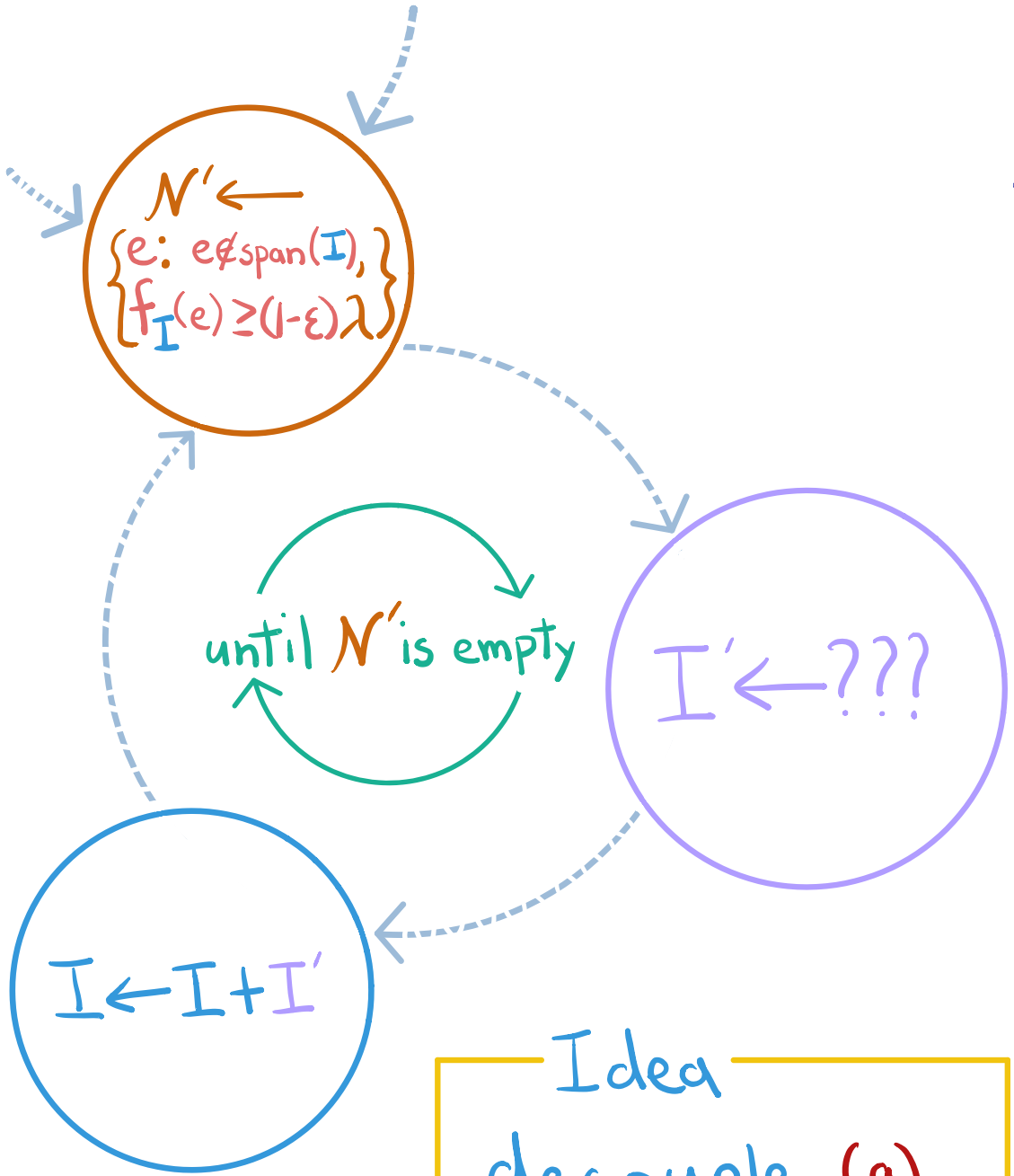
(a) $I + I'$ is independent and $f_I(I') \approx \lambda |I'|$

(b) $|N'|$ ⬇ a lot w/r/t $I + I'$ by combination of:

(i) $I + I'$ spans many $e \in N'$

(ii) $f_{I+I'}(e) \ll f_I(e)$ for many $e \in N'$

But (a) $=\not\Rightarrow=$ (b)



Goal

select a set $I' \subseteq N'$ s.t.

(a) $I + I'$ is independent and $f_I(I') \approx \lambda |I'|$

(b) $|N'|$ ⬇ a lot w/r/t $I + I'$ by combination of:

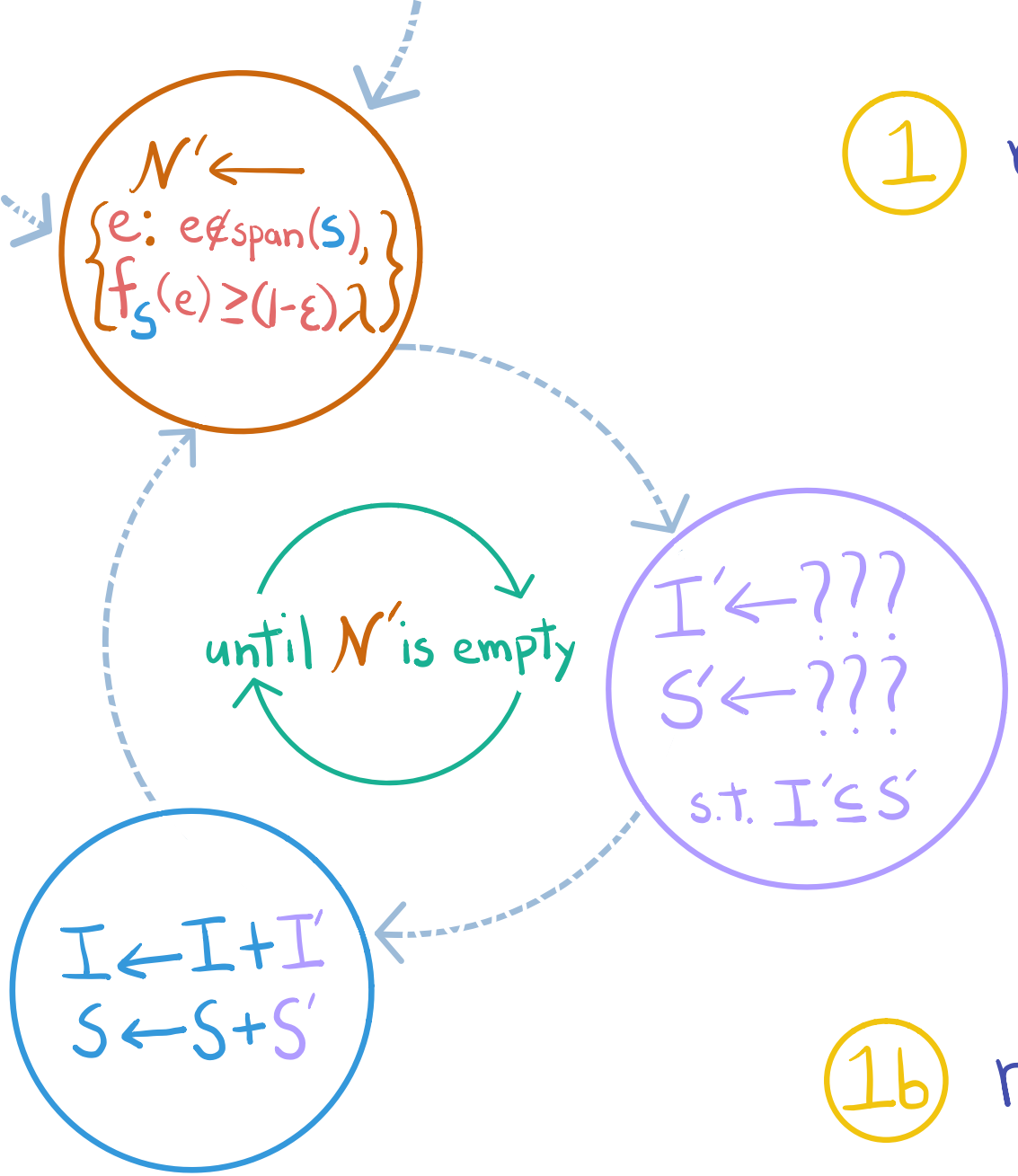
(i) $I + I'$ spans many $e \in N'$

(ii) $f_{I+I'}(e) \ll f_I(e)$ for many $e \in N'$

Idea

decouple (a) and (b) with different sets

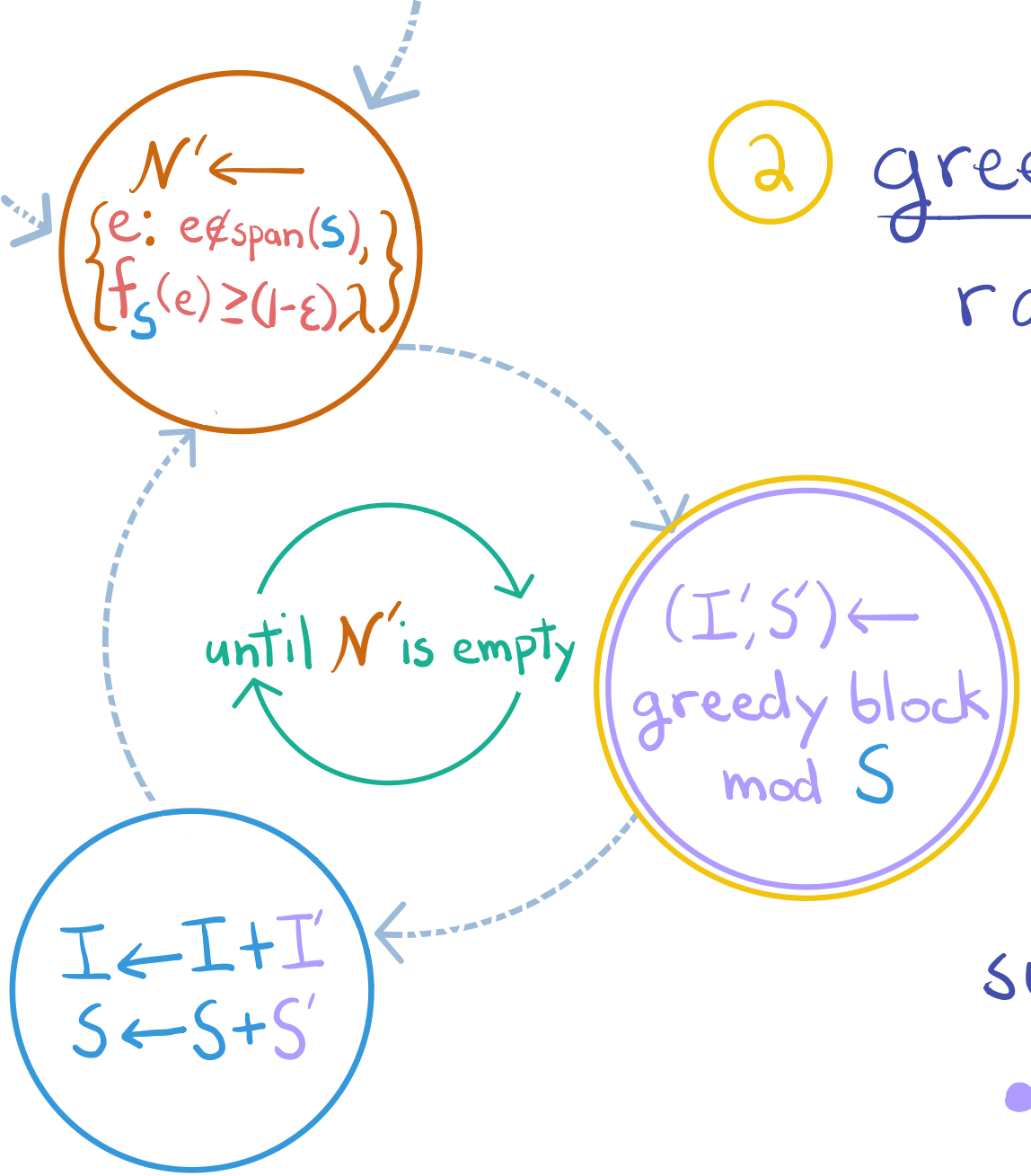
But (a) =X= (b)



(1) replace I' with 2 sets I' and S'

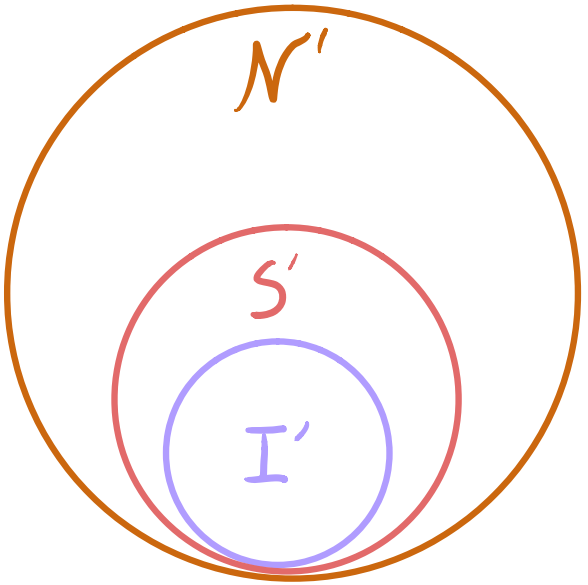
- I' = "clean" indep. set that we add
- S' = "messy" superset of I' driving $|N'|$ ⬇

(1b) replace I w/ I and S (for sake of induction)



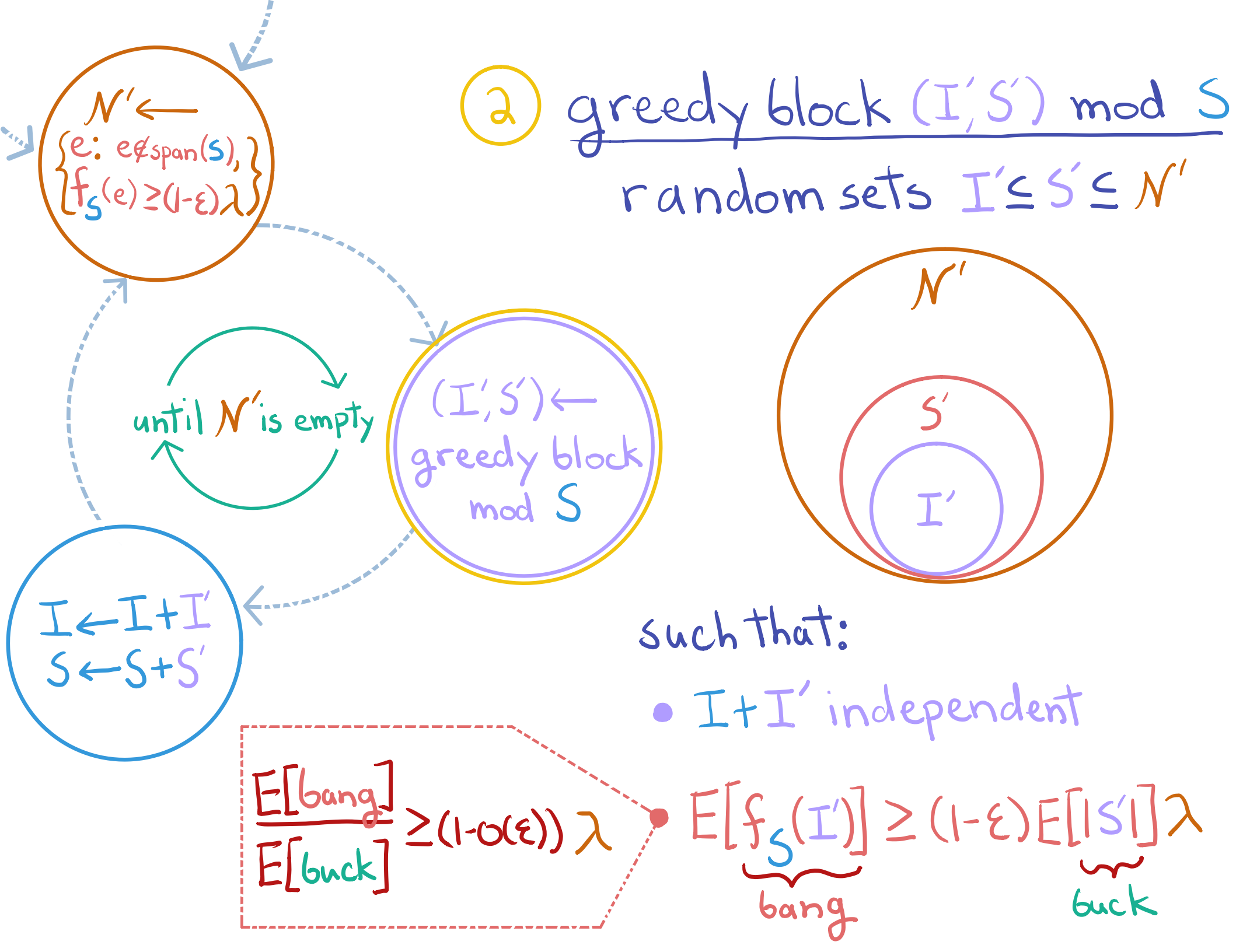
② greedy block (I', S') mod S

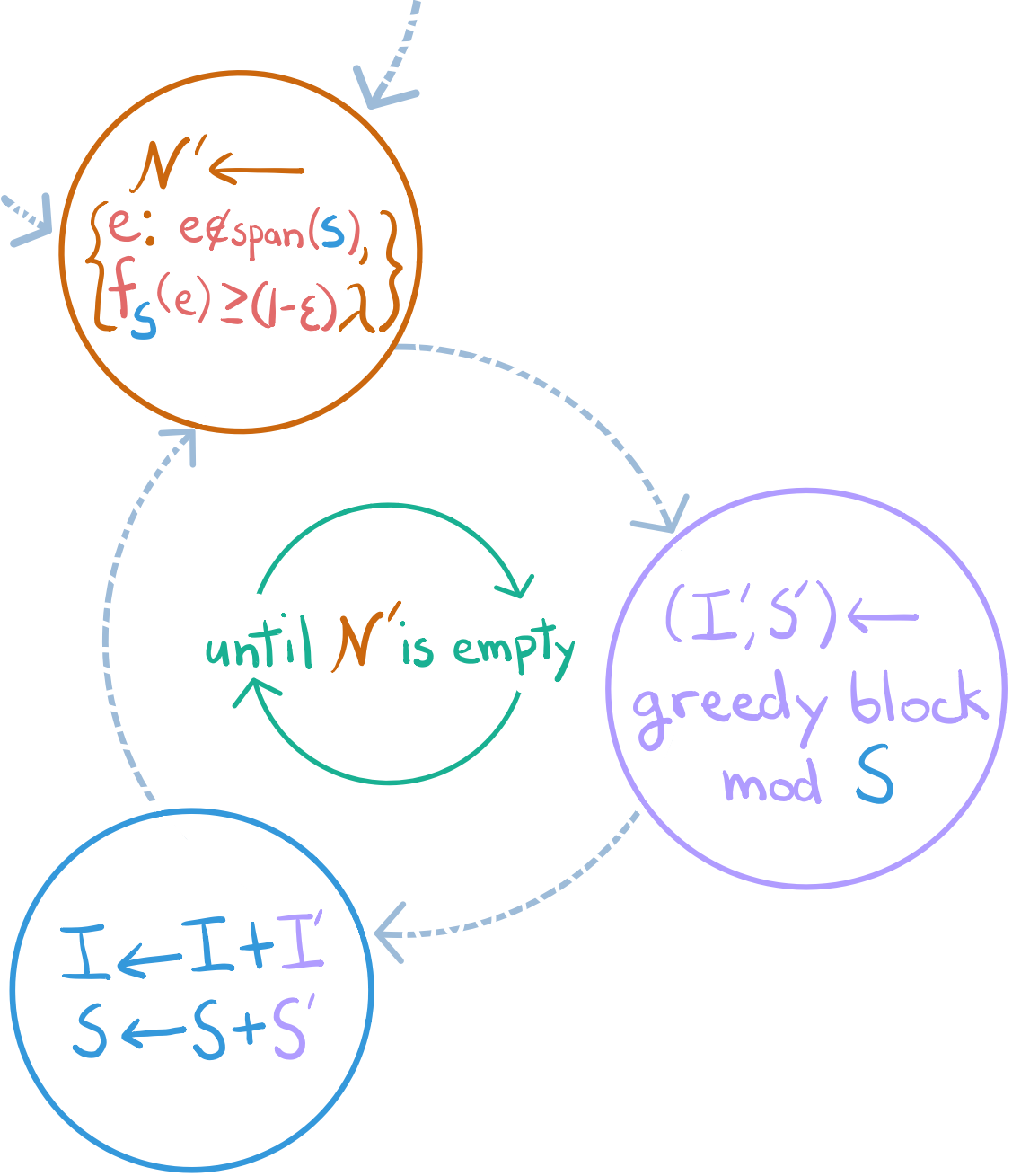
random sets $I' \subseteq S' \subseteq N'$



such that:

- $I + I'$ independent
- $E[f_S(I')] \geq (1-\varepsilon) E[|S'|] \lambda$


N' ← {e: e∉span(S), f_S(e) ≥ (1-ε)λ}
until N' is empty
(I', S') ← greedy block mod S
I ← I + I'
S ← S + S'
② greedy block (I', S') mod S
random sets I' ⊆ S' ⊆ N'
N'
S'
I'
such that:
• I + I' independent
E[bang] / E[buck] ≥ (1-O(ε)) λ
E[f_S(I')] ≥ (1-ε) E[|S'|] λ
bang
buck



Need to show

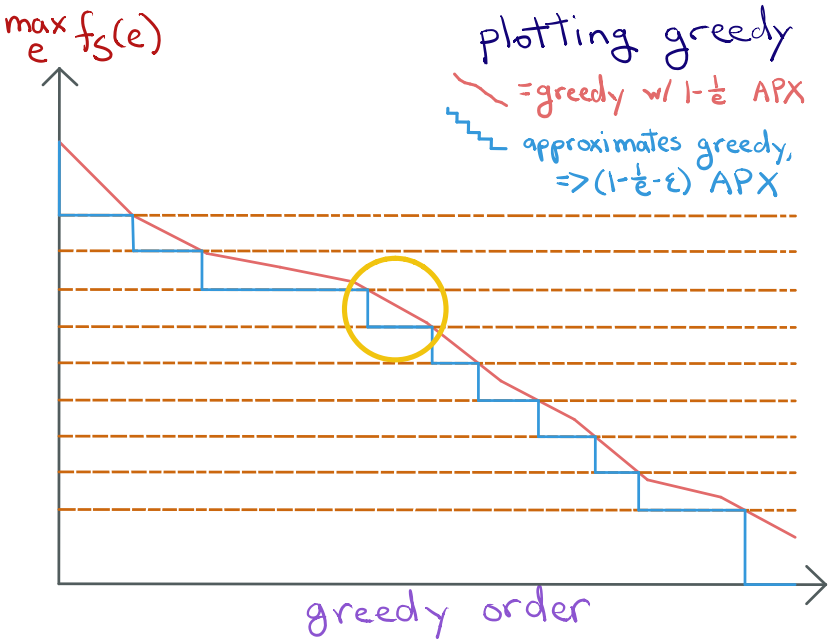
(a) APX factor

(b) how? and depth

greedy blocks guarantee:

$$\frac{E[\text{bang}]}{E[\text{buck}]} \geq (1-O(\varepsilon))\,\lambda$$

(more or less) follows greedy



(complicated by bang = f(I) decoupled from buck = S)

$\Rightarrow \approx \Rightarrow$

same APX guarantees as greedy

APX guarantees

Standard greedy

returns $I \in \mathcal{I}$ s.t.

- for all $J \in \mathcal{I}$, $f(I) \geq f_I(J)$

"block greedy" (Lemma)

returns $I \in \mathcal{I}$, $S \subseteq N$ s.t.

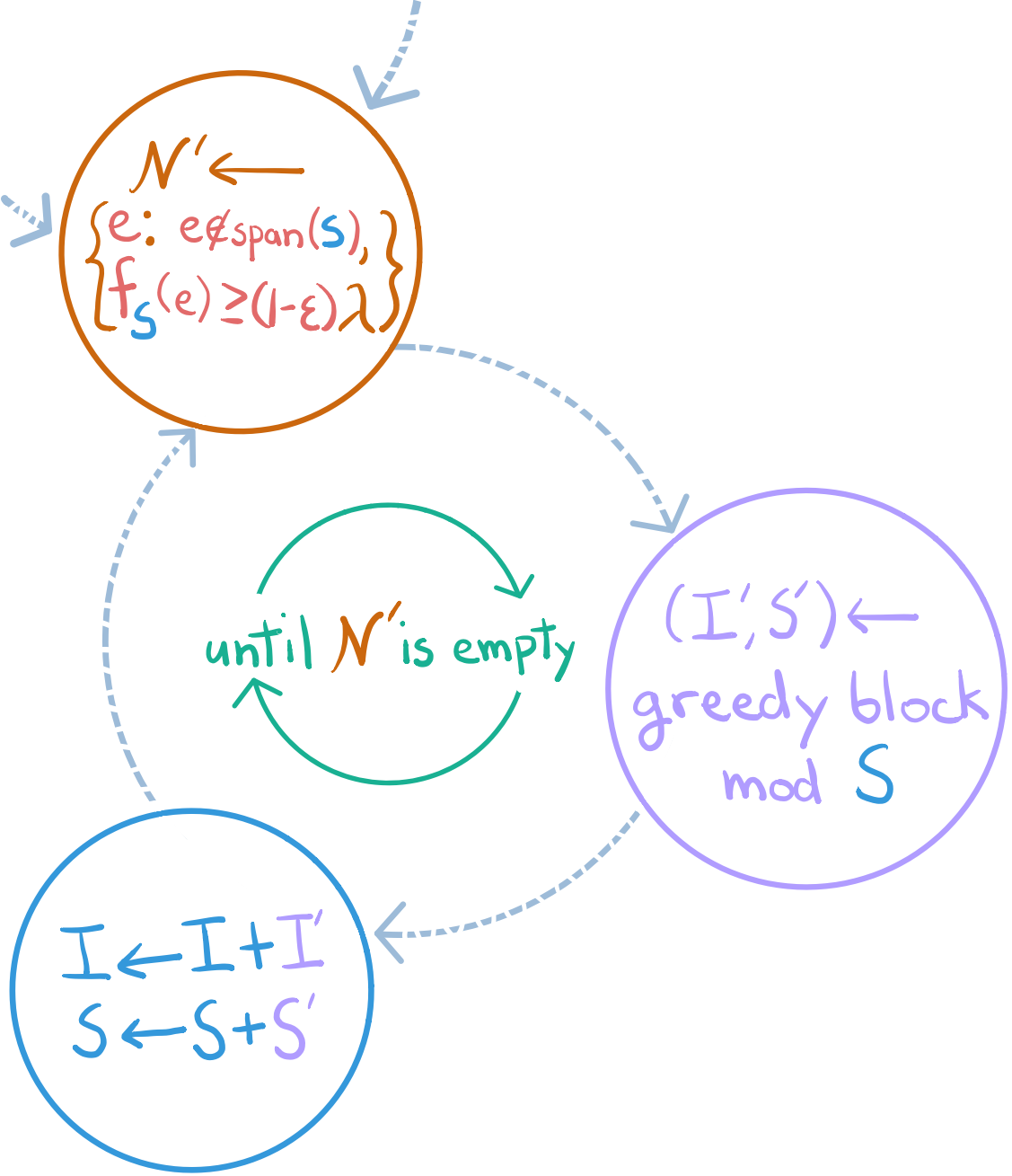
- for all $J \in \mathcal{I}$, $E[f(I)] \geq (1-\varepsilon) E[f_S(J)]$
- $E[f(I)] \geq (1-\varepsilon) E[f(S)]$

[omit proof]

Corollary

$(1-\varepsilon)\frac{1}{2}$-APX for monotone f,

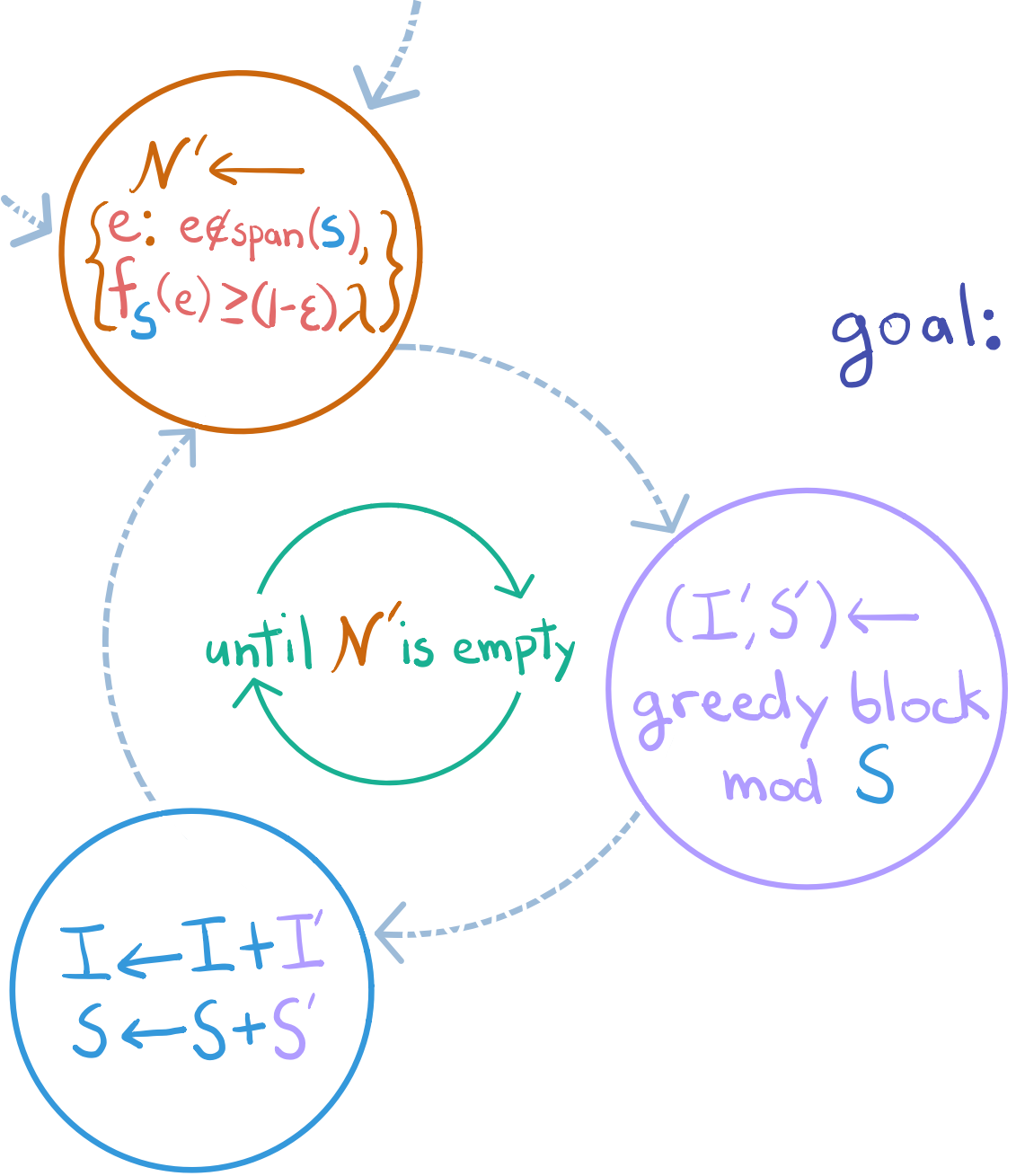
constant APX for nonnegative f



Need to show

(a) ~~APX factor~~

(b) how? and depth



"greedy sampling"

goal: compute greedy block s.t. $|N'|$ decreases a lot

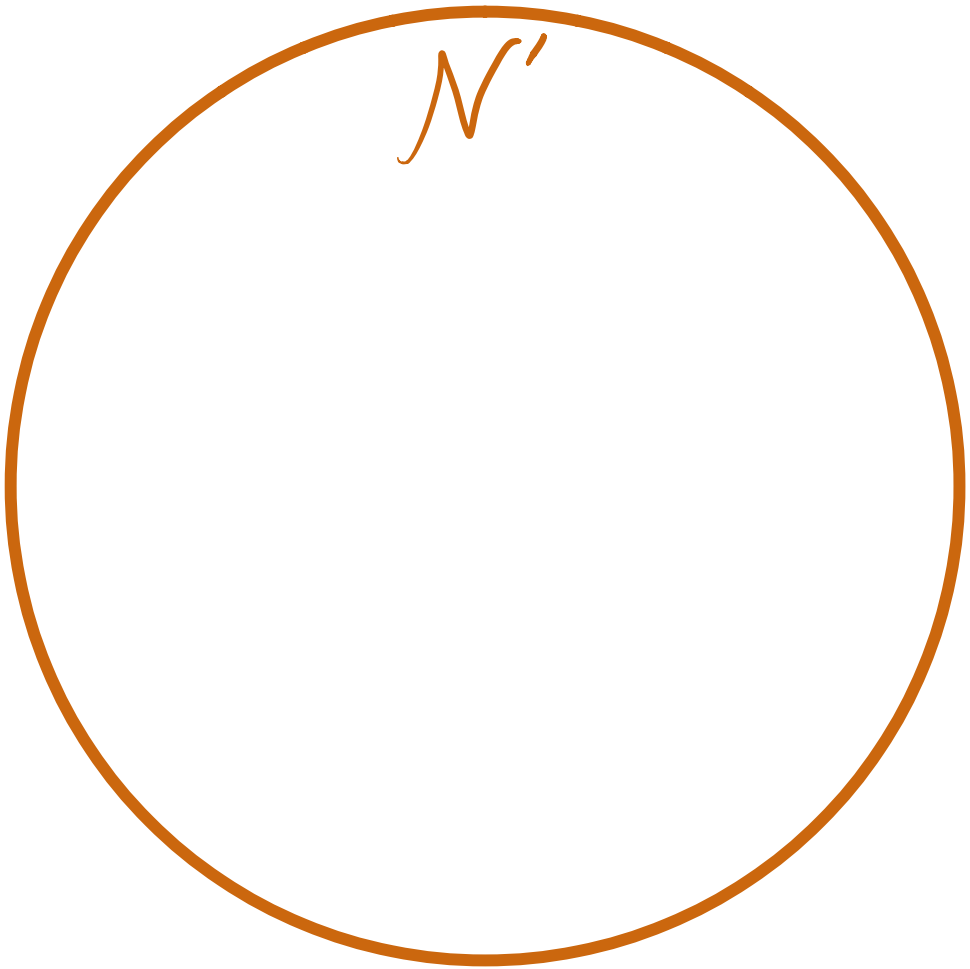
Need to show

(a) ~~APX factor~~

(b) how? and depth

"greedy sampling"

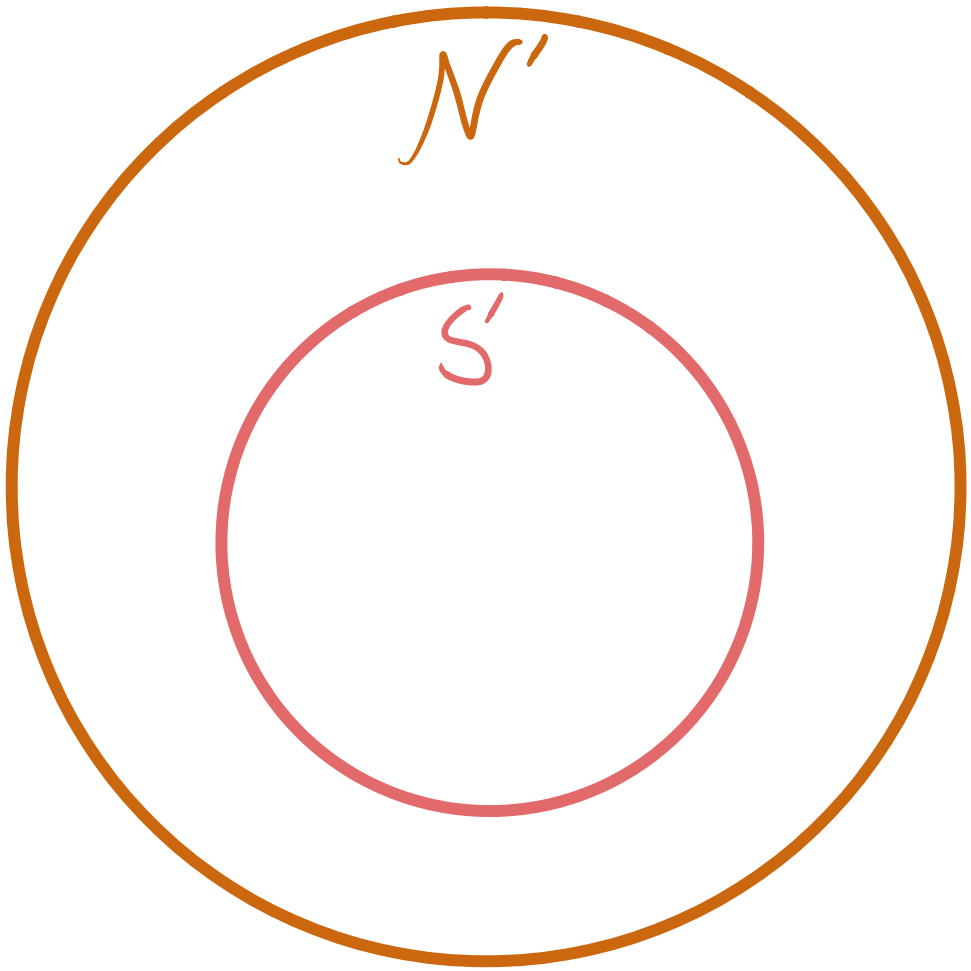
let $\delta > 0$ TBD



"greedy sampling"

let $\delta > 0$ TBD

(a) sample $S' \sim \delta N'$

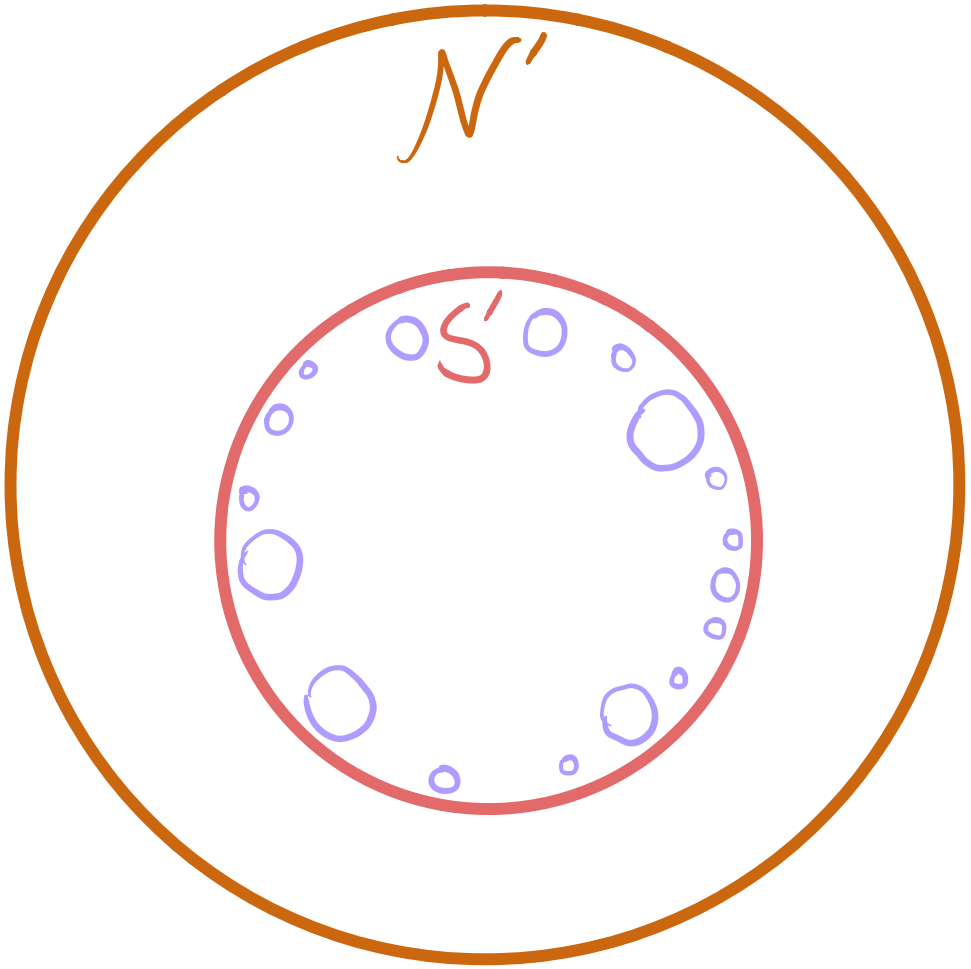


"greedy sampling"

let $\delta > 0$ TBD

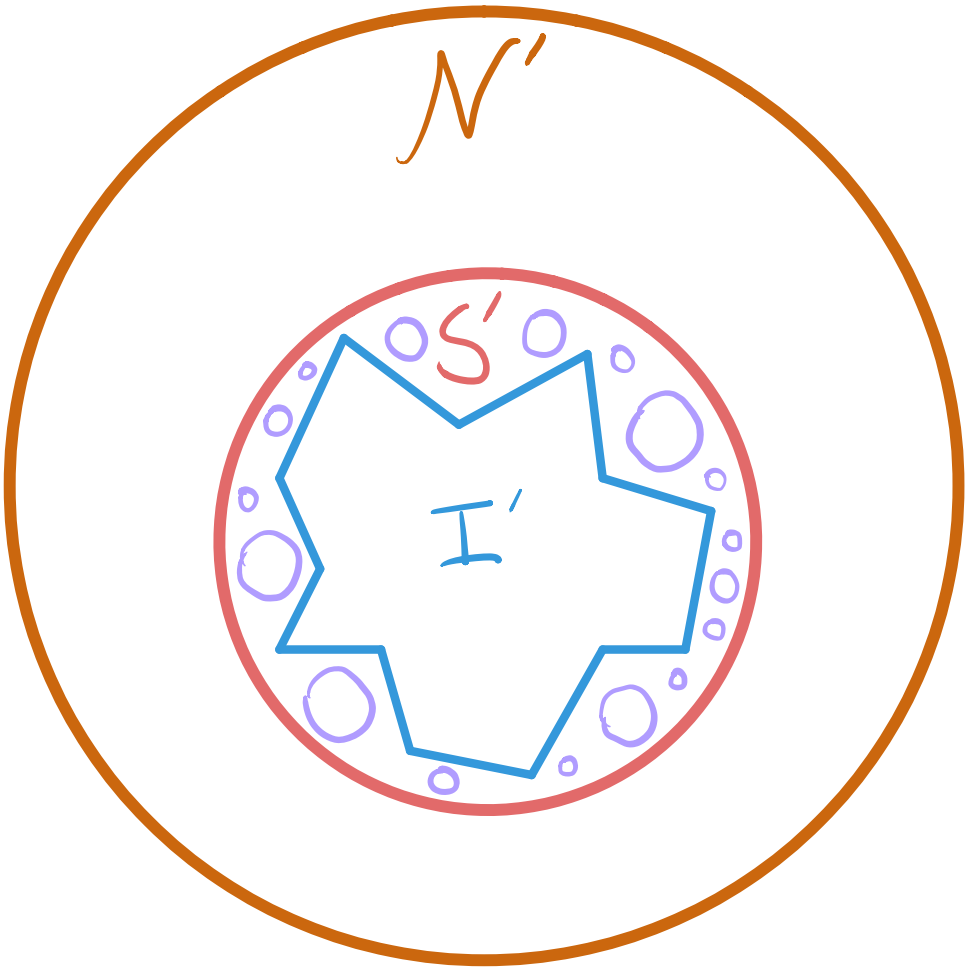
(a) sample $S' \sim_\delta N'$

[S' may have dependencies, or bad marginal values]



"greedy sampling"

let $\delta > 0$ TBD



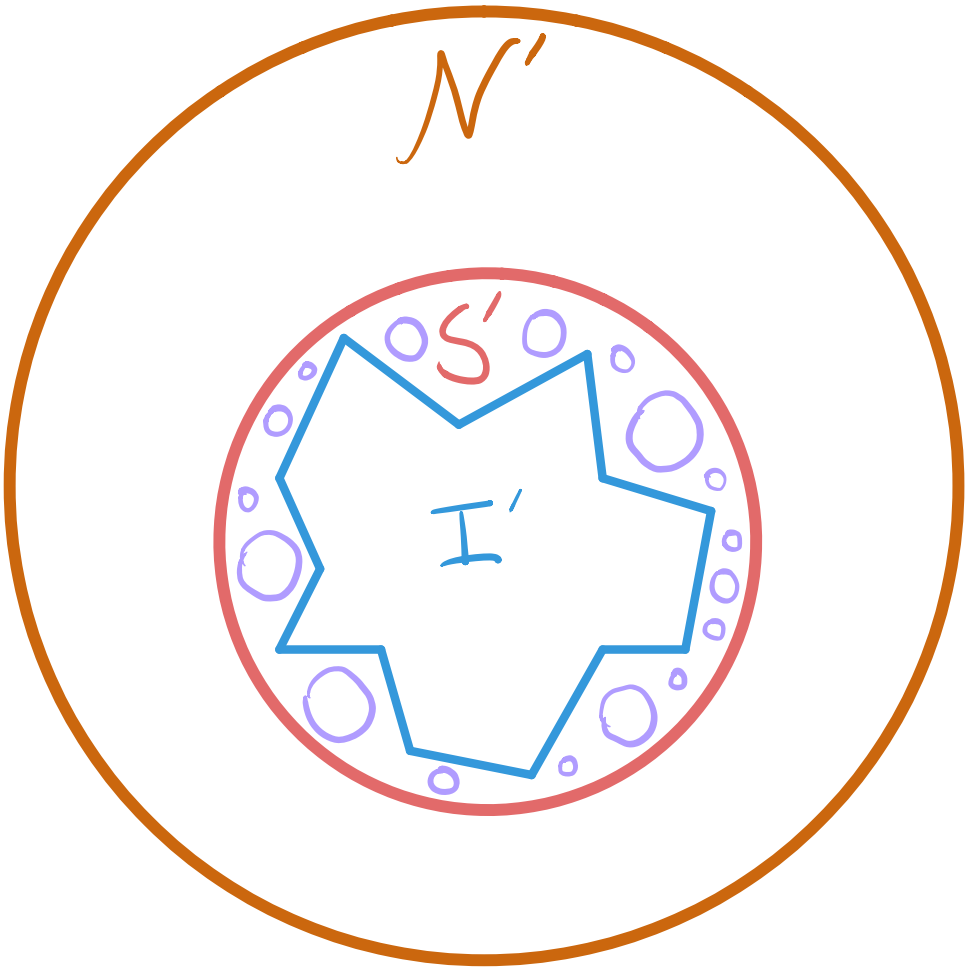
(a) sample $S' \sim \delta N'$

[S' may have dependencies, or bad marginal values]

(b) $I' \leftarrow \{ e \in S' : e \notin \text{span}(S + S' - e) \text{ and } f_{S+S'-e}(e) \geq (1-\varepsilon)\lambda \}$

"greedy sampling"

let $\delta > 0$ TBD



(a) sample $S' \sim \delta N'$

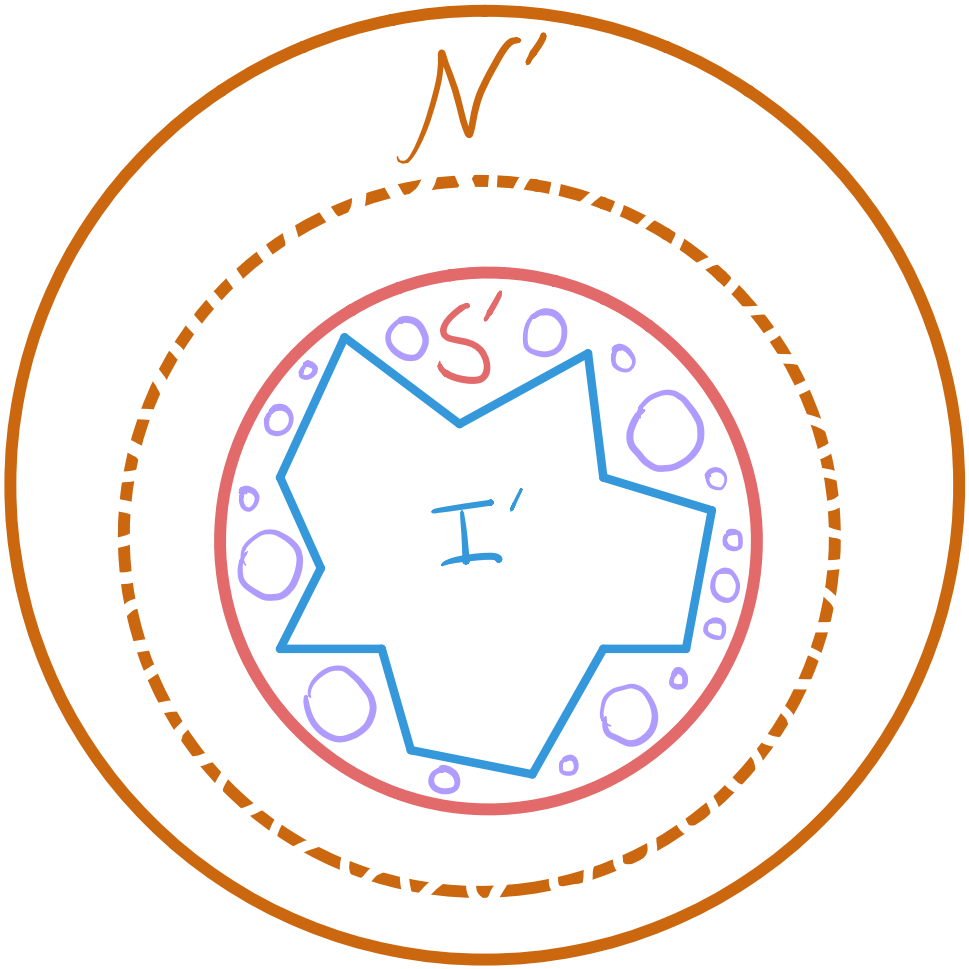
[S' may have dependencies, or bad marginal values]

(b) $I' \leftarrow \{ e \in S' : e \notin \text{span}(S + S' - e) \text{ and } f_{S+S'-e}(e) \geq (1-\varepsilon)\lambda \}$

<---------- in parallel w/ independent queries to f and span oracles

"greedy sampling"

let $\delta > 0$ TBD



(a) sample $S' \sim \delta N'$

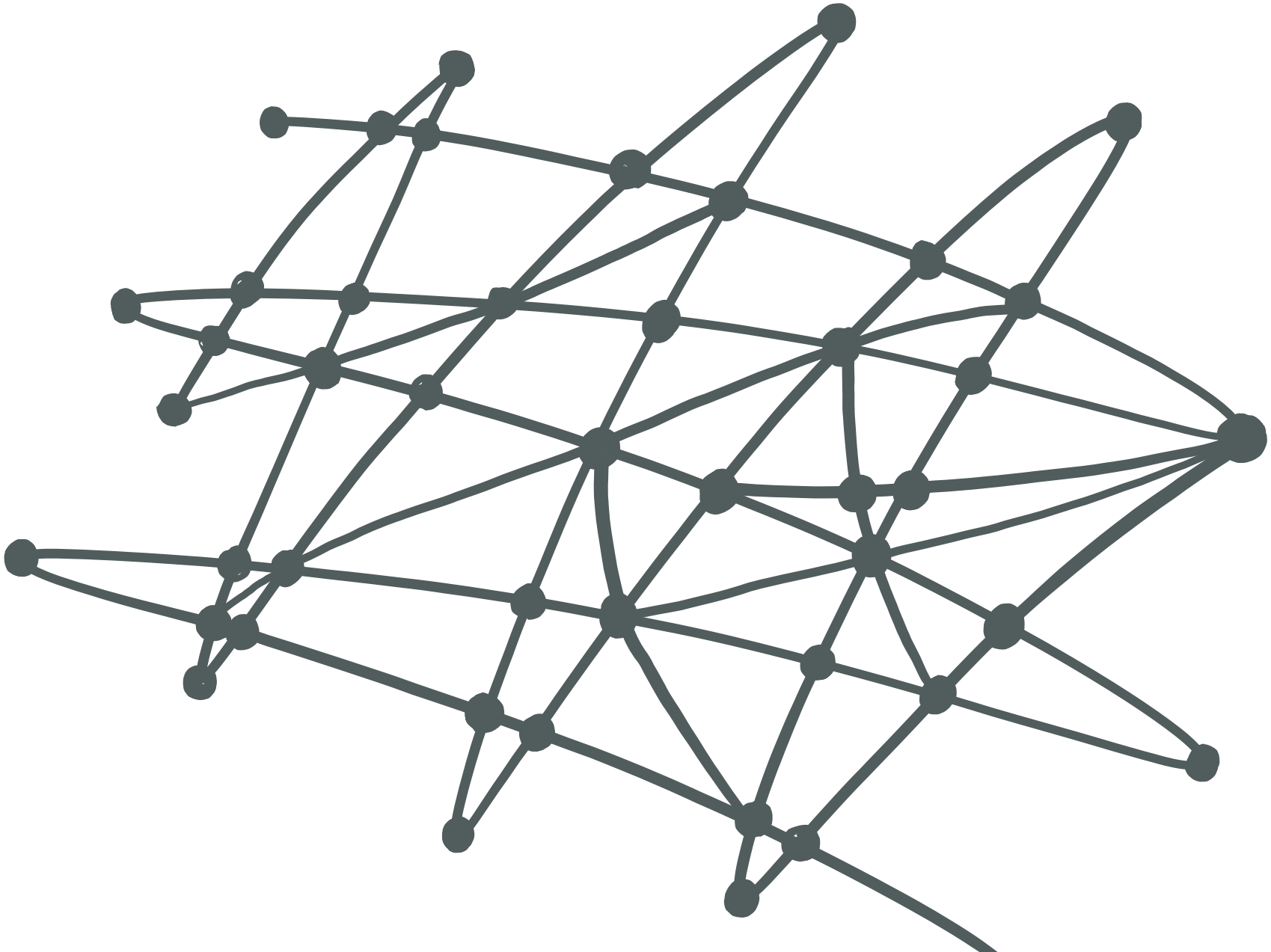
[S' may have dependencies, or bad marginal values]

(b) $I' \leftarrow \{ e \in S' : e \notin \text{span}(S + S' - e) \text{ and } f_{S+S'-e}(e) \geq (1-\varepsilon)\lambda \}$

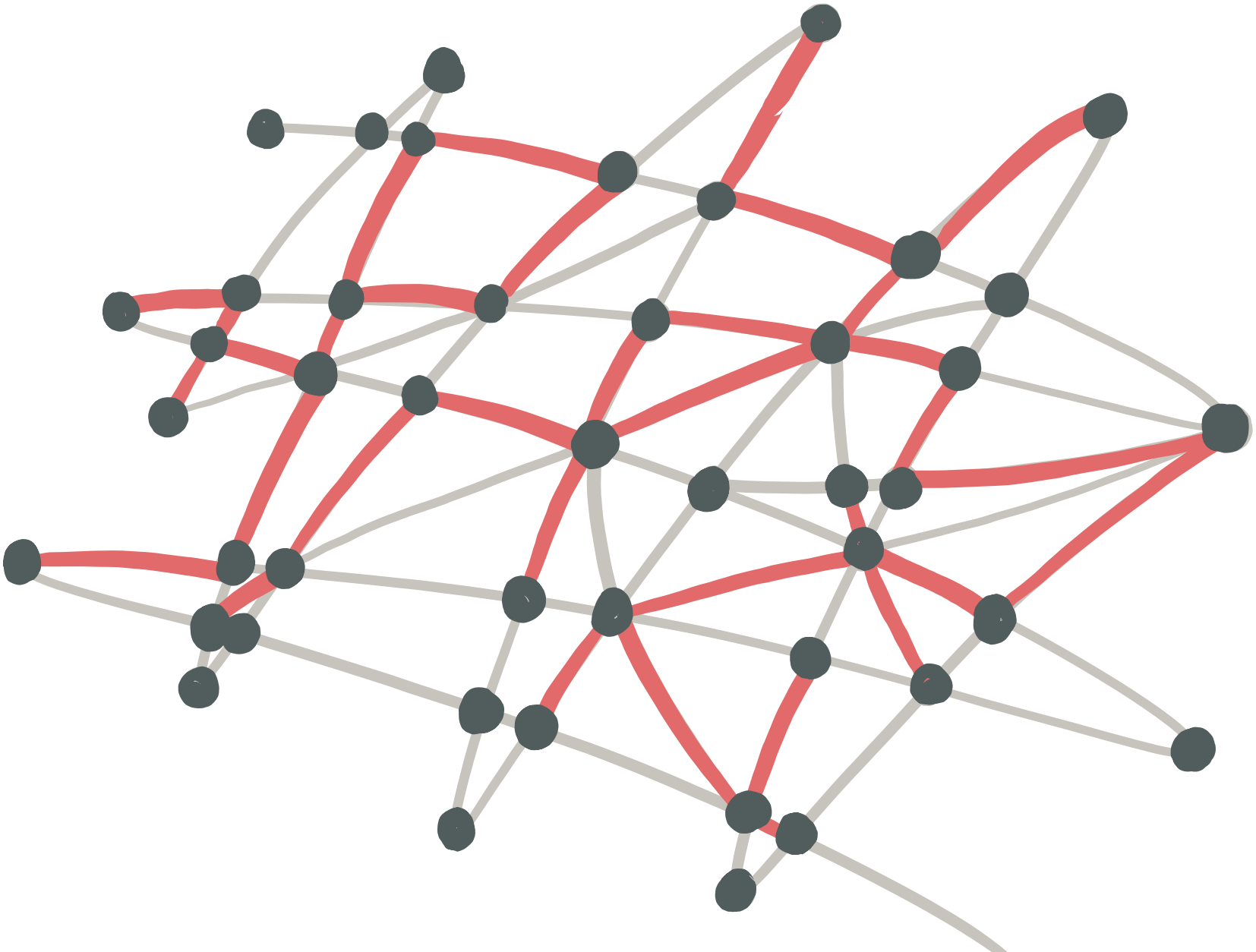
<------ in parallel w/ independent queries to f and span oracles

(and then N' shrinks)

e.g. Graphic Matroid $(N=E,\ \mathcal{I}=\{\text{forests}\})$

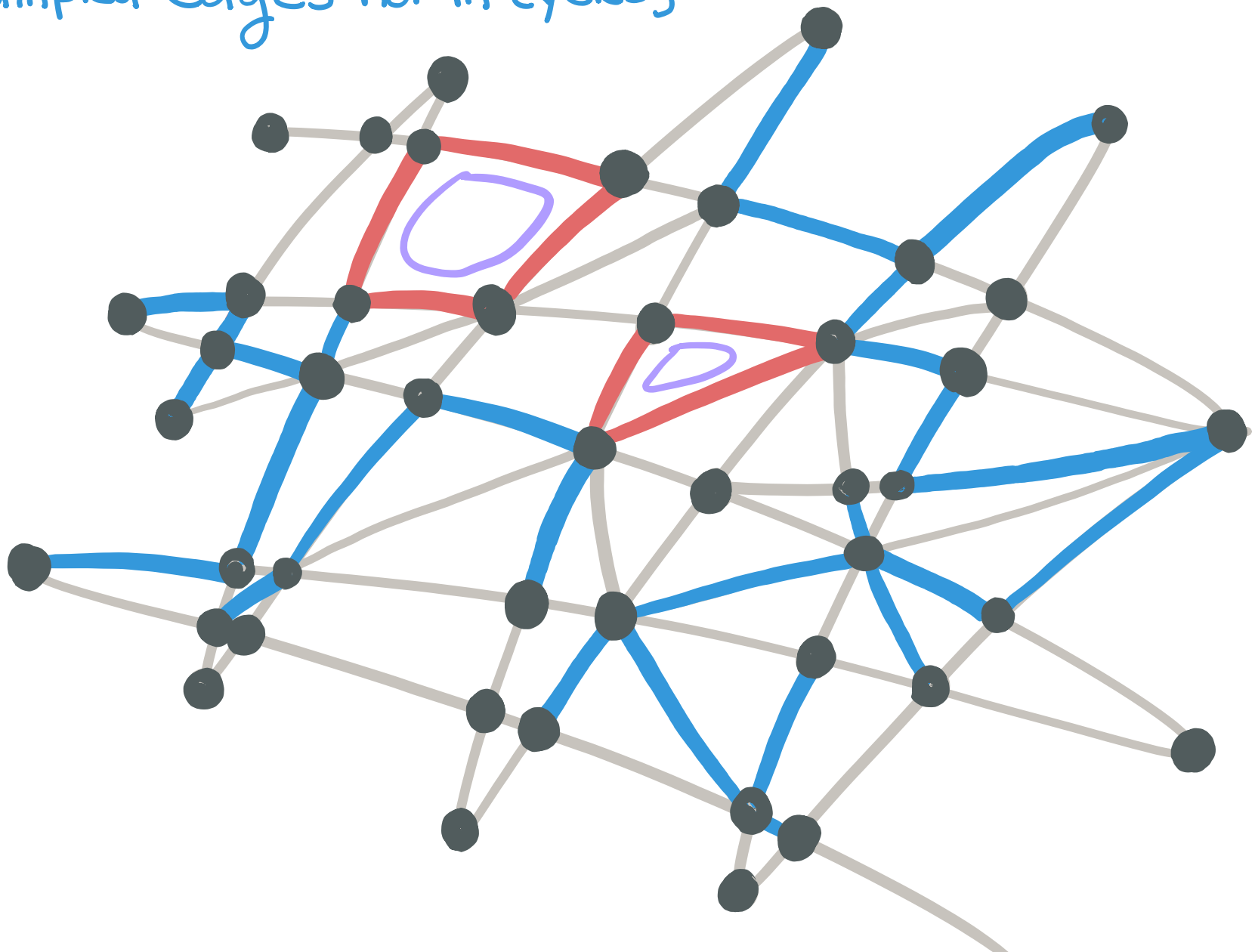


① let S' sample each edge w/ prob δ



① let S' sample each edge w/ prob δ

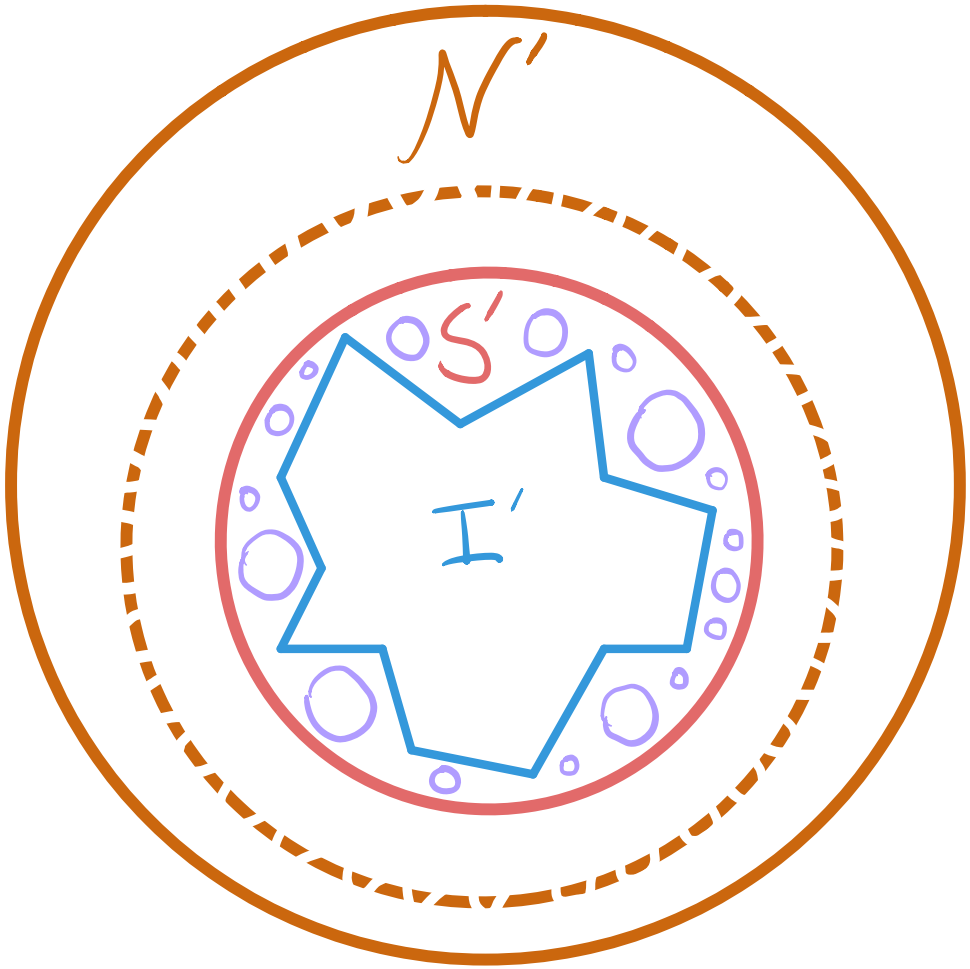
② $I' \subseteq$ {sampled edges not in cycles}



"greedy sampling"

let $\delta > 0$ TBD





(a) sample $S' \sim_\delta N'$

[S' may have dependencies, or bad marginal values]

(b) $I' \leftarrow \{ e \in S' : e \notin \text{span}(S + S' - e) \text{ and } f_{S+S'-e}(e) \geq (1-\varepsilon)\lambda \}$

(and then N' shrinks)

"greedy sampling"

let $\delta > 0$ TBD

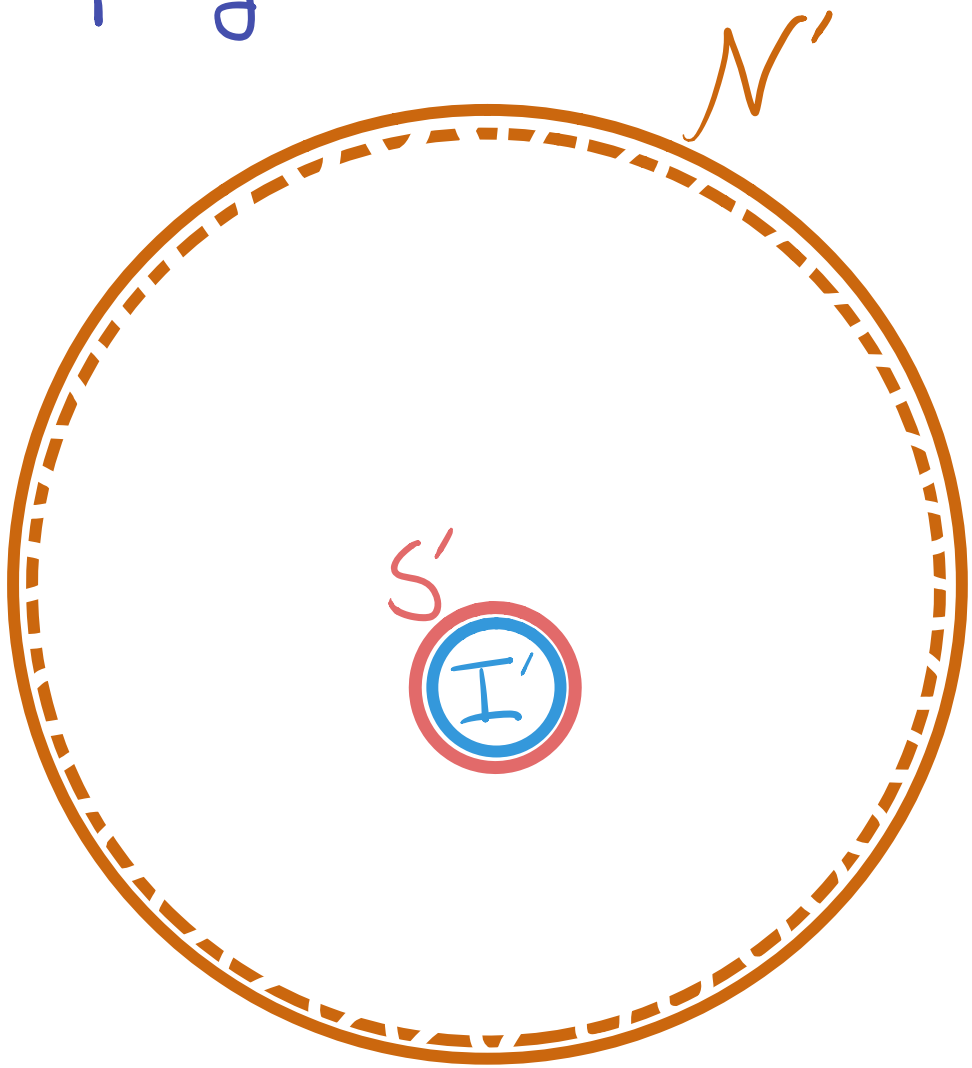
δ very very small

(a) sample $S' \sim \delta N'$

[S' may have dependencies, or bad marginal values]

(b) $I' \leftarrow \{ e \in S' : e \notin \text{span}(S + S' - e) \text{ and } f_{S+S'-e}(e) \geq (1-\varepsilon)\lambda \}$

(and then N' shrinks)



I' retains most of S'

but

S' too small to shrink N'

"greedy sampling"

let $\delta > 0$ TBD

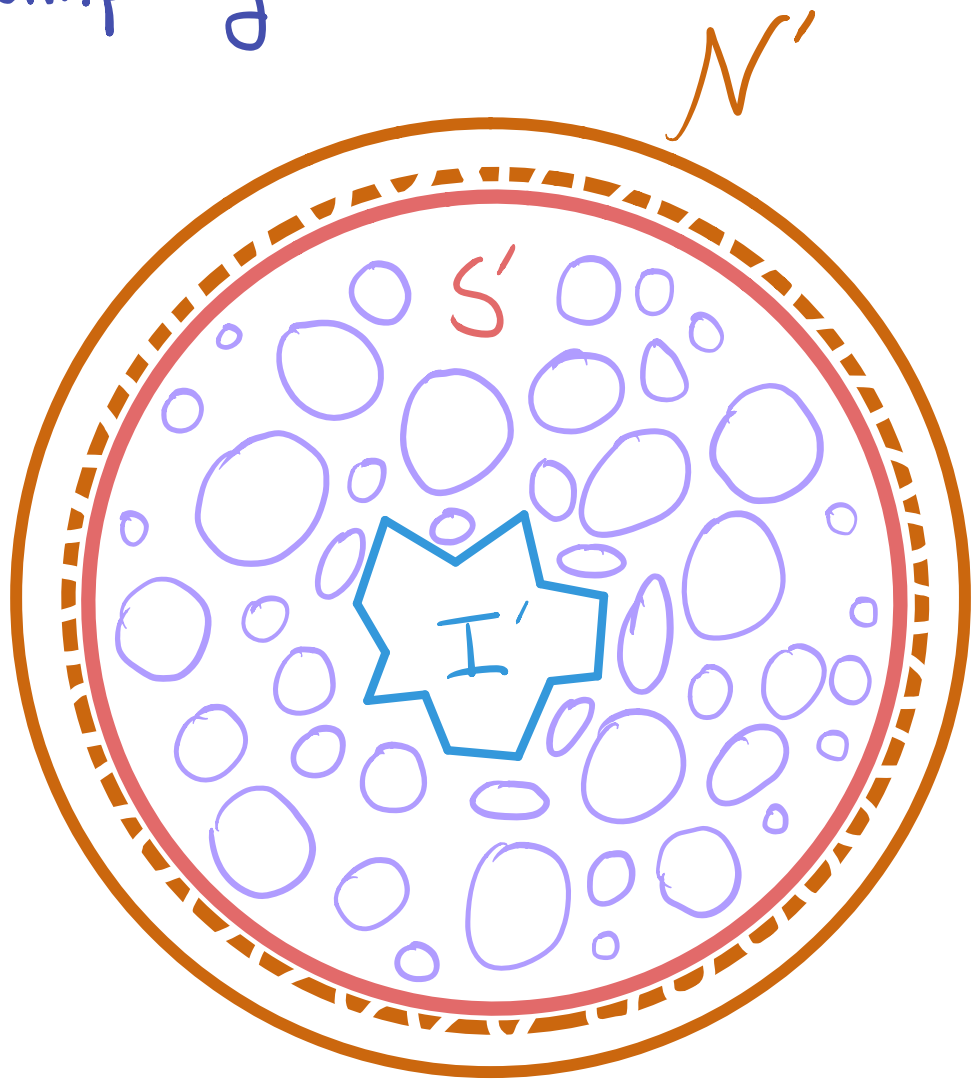
δ very very big

(a) sample $S' \sim \delta_{N'}$

[S' may have dependencies, or bad marginal values]

(b) $I' \leftarrow \{ e \in S' : e \notin \text{span}(S + S' - e) \text{ and } f_{S+S'-e}(e) \geq (1-\varepsilon)\lambda \}$

(and then N' shrinks)

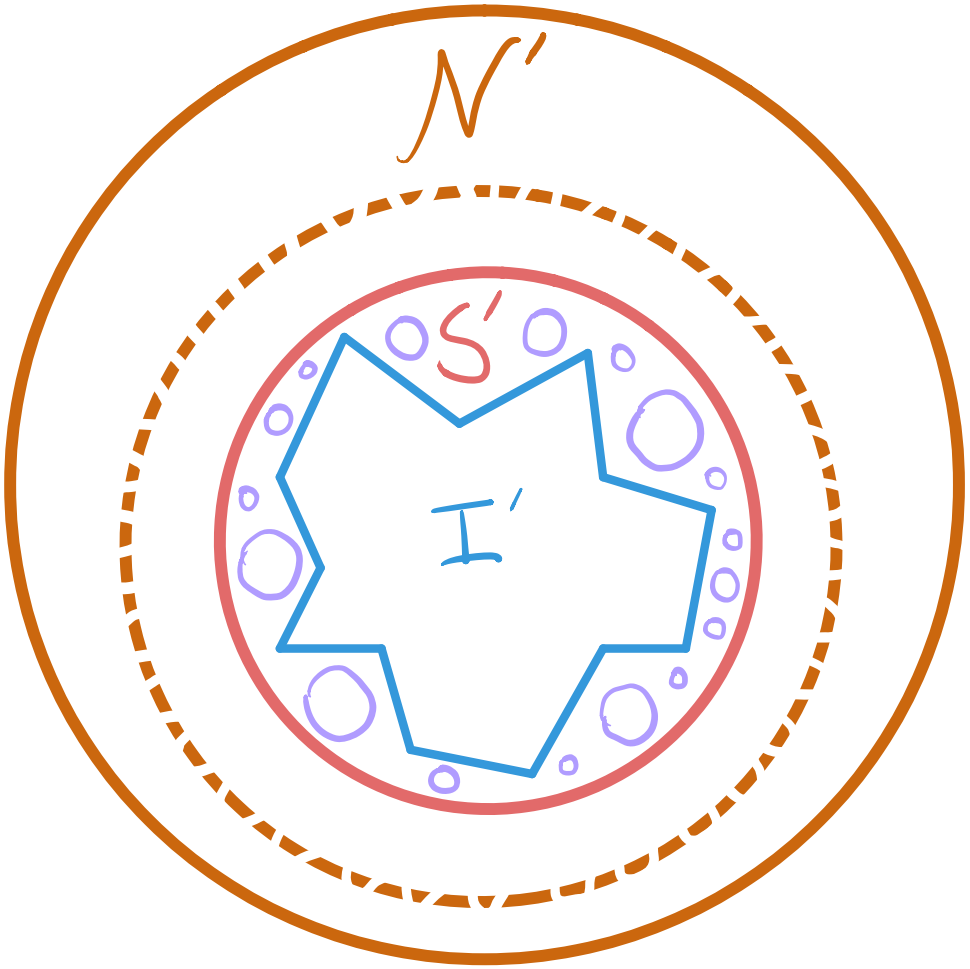


S shrinks N' but many dependencies leads to small I

"greedy sampling"

let $\delta > 0$ TBD





(a) sample $S' \sim_\delta N'$

[S' may have dependencies, or bad marginal values]

(b) $I' \leftarrow \{ e \in S' : e \notin \text{span}(S + S' - e) \text{ and } f_{S+S'-e}(e) \geq (1-\varepsilon)\lambda \}$

(and then N' shrinks)

choose max $\delta > 0$ s.t.

"greedy sampling"

for $S' \sim \delta N'$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

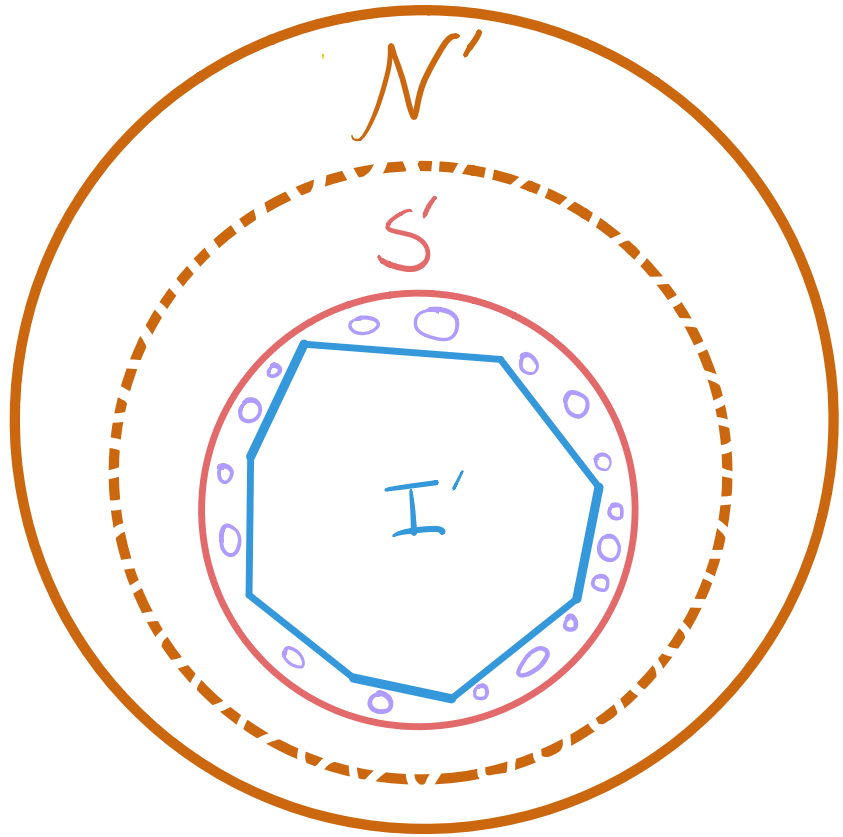
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

[S' may have dependencies, or bad marginal values]

(b) $I' \leftarrow \{e \in S' : e \notin \text{span}(S+S'-e)$ and $f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

(and then N' shrinks)



choose max $\delta > 0$ s.t. "greedy sampling"

for $S' \sim \delta N'$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

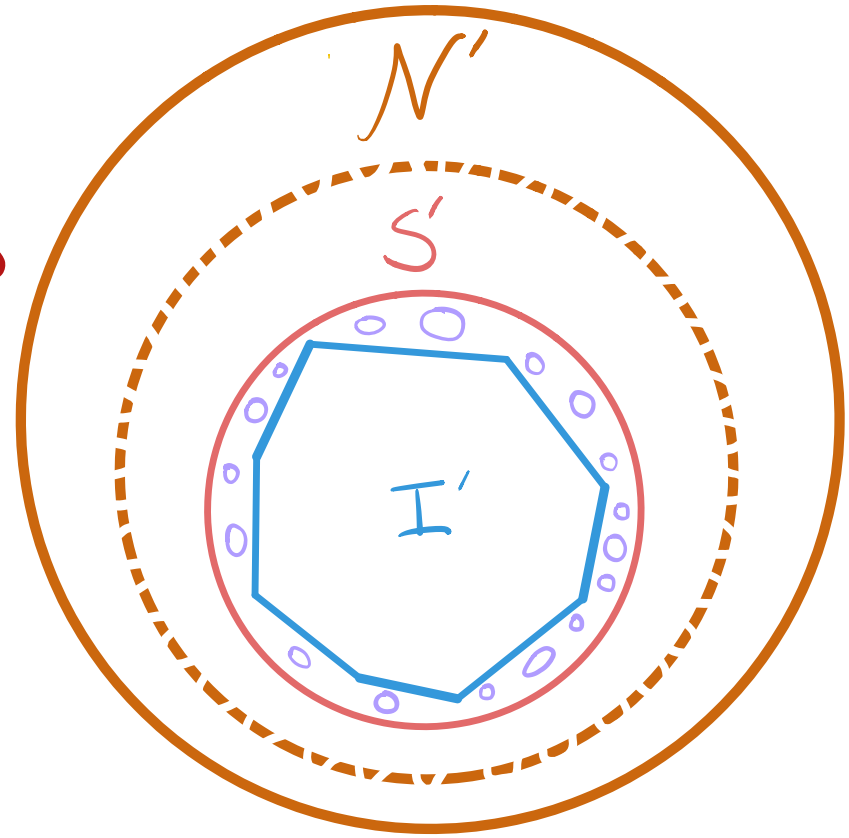
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

Key Claims

(1) (I', S') is greedy block mod S

(namely, $\mathbb{E}[f_S(I')] \ge (1-O(\varepsilon))\lambda \mathbb{E}[|S'|]$)

(2) N' shrinks by $(1-\varepsilon)$-factor



(1) (I', S') is greedy block models

(namely, $E[f_S(I')] \geq (1-O(\varepsilon))\lambda E[|S'|]$)

Proof is subtle but main ideas are:

(1) δ is chosen conservatively small

(2) submodular functions and matroids are well-behaved under sampling

proof slides in appendix

choose max $\delta > 0$ s.t. "greedy sampling"

---

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

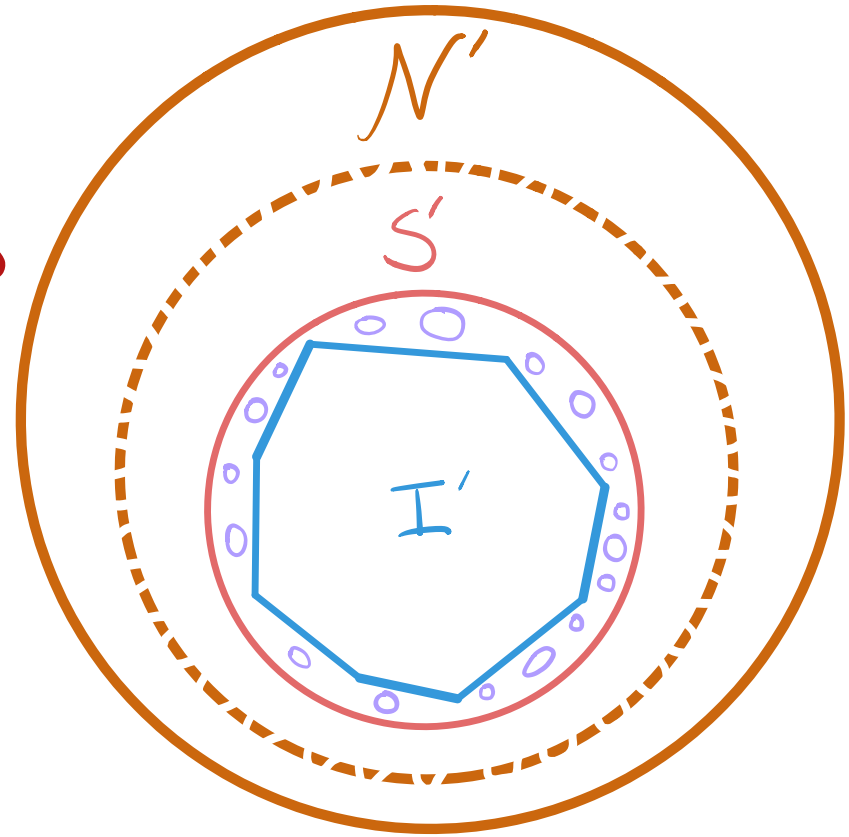
(b) $\mathbb{E}[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \le \varepsilon |N'|$

Key Claims

(1) (I', S') is greedy block mod S

(namely, $\mathbb{E}[f_S(I')] \ge (1-O(\varepsilon))\lambda \mathbb{E}[|S'|]$)

(2) N' shrinks by $(1-\varepsilon)$-factor



choose max $\delta > 0$ s.t.

---

for $S' \sim_\delta S$:

both:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \leq (1-\varepsilon)\lambda\}|] \leq \varepsilon |N'|$

(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \leq \varepsilon |N'|$

---

both $\mathbb{E}[|\{\dots\}|]$ terms ⬆ as δ ⬆

choose ~~max~~ min $\delta > 0$ s.t.

---

for $S' \sim_\delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \leq (1-\varepsilon)\lambda\}|] \geq \varepsilon |N'|$

~~both~~ either

(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \geq \varepsilon |N'|$

---

both $\mathbb{E}[|\{\dots\}|]$ terms ↑ as δ ↑

choose ~~max~~ min $\delta > 0$ s.t.

---

for $S' \sim_\delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \geq \varepsilon |N'|$

~~both~~ either

(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \geq \varepsilon |N'|$

---

both $\mathbb{E}[|\{\dots\}|]$ terms ↑ as δ ↑

i.e., min δ s.t. $|N'|$ ↓ by $(1-\varepsilon)$ factor

choose max $\delta > 0$ s.t.

"greedy sampling"

---

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

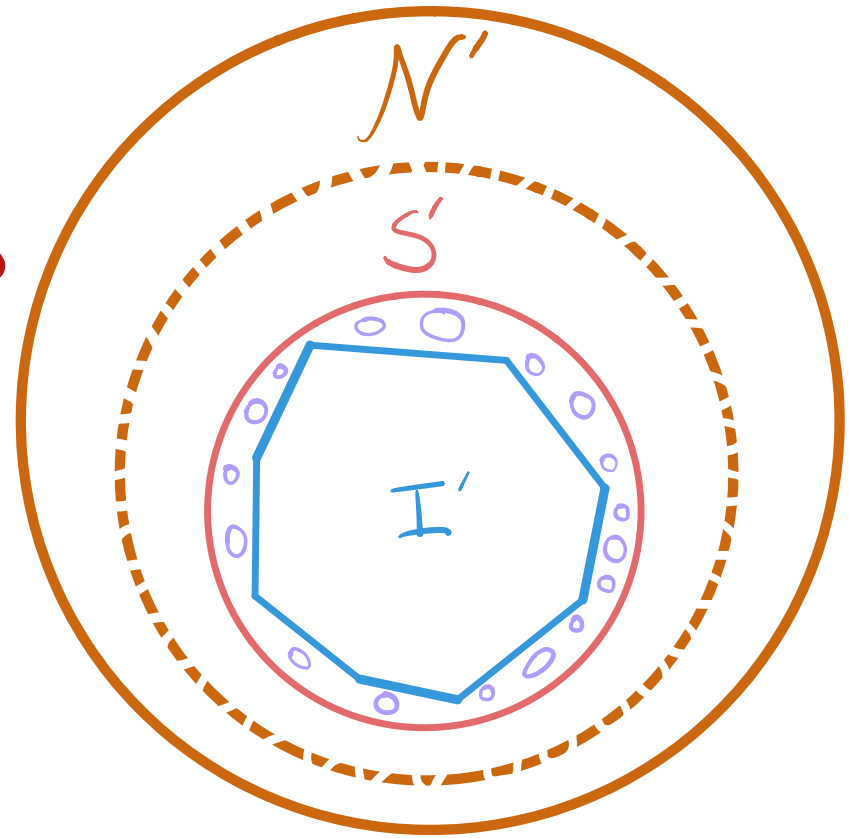
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

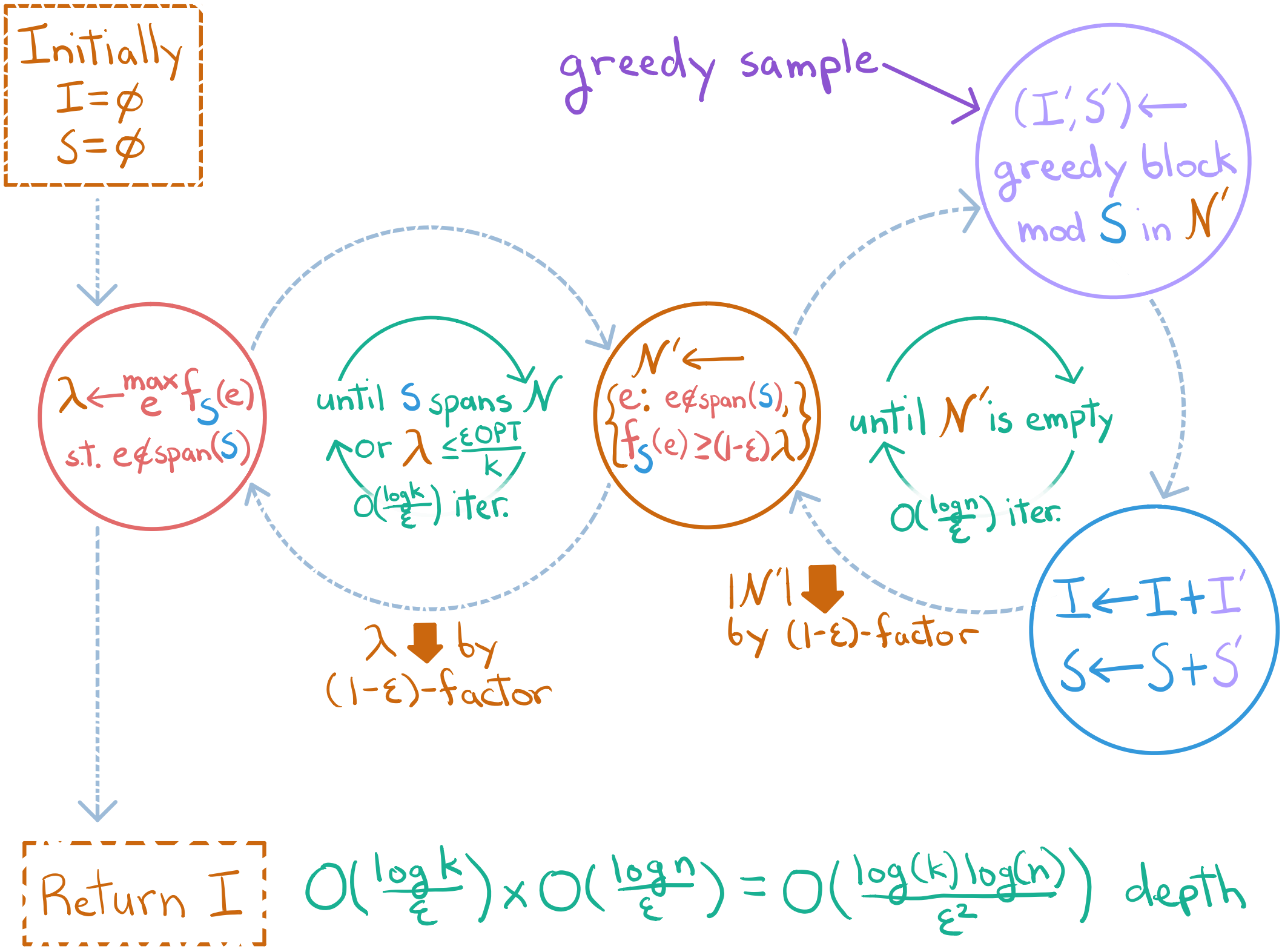
Key Claims

(1) (I', S') is greedy block mod S

(namely, $\mathbb{E}[f_S(I')] \ge (1-O(\varepsilon))\lambda \mathbb{E}[|S'|]$)

(2) N' shrinks by $(1-\varepsilon)$-factor




Initially
$I=\emptyset$
$S=\emptyset$
greedy sample
$(I', S') \leftarrow$ greedy block mod S in N'
$\lambda \leftarrow \max_e f_S(e)$ s.t. $e \notin \text{span}(S)$
until S spans N or $\lambda \leq \frac{\varepsilon OPT}{k}$
$O(\frac{\log k}{\varepsilon})$ iter.
$N' \leftarrow \{e: e \notin \text{span}(S), f_S(e) \geq (1-\varepsilon)\lambda\}$
until N' is empty
$O(\frac{\log n}{\varepsilon})$ iter.
$I \leftarrow I + I'$
$S \leftarrow S + S'$
$\lambda \downarrow$ by $(1-\varepsilon)$-factor
$|N'| \downarrow$ by $(1-\varepsilon)$-factor
Return I
$O(\frac{\log k}{\varepsilon}) \times O(\frac{\log n}{\varepsilon}) = O(\frac{\log(k)\log(n)}{\varepsilon^2})$ depth

$1/2$-APX for monotone f [Fisher, Nemhauser, Wolsey]

this talk

Iterate along greedy blocks

generate greedy blocks w/ greedy sampling

$\frac{1/2 - \varepsilon}{\text{constant}}$-APX for $\frac{\text{monotone}}{\text{nonnegative}}$ f w/ $\tilde{O}(\frac{1}{\varepsilon^2})$ depth

see paper

"multilinear amplification" for monotone f

[Vondrák, Badanidiyuru]

generalize to nonnegative f

fractional $\frac{1 - 1/e - \varepsilon}{1/e - \varepsilon}$-APX for $\frac{\text{monotone}}{\text{nonnegative}}$ f w/ $\tilde{O}(\frac{1}{\varepsilon^3})$ depth

Summary of results (related to matroids)

matroid constraint in polylog depth

- integral approximations matching greedy
- fractional* approx. matching continuous greedy

* can be rounded oblivious to f

k-matchoid constraints in polylog depth

- similar results to matroid

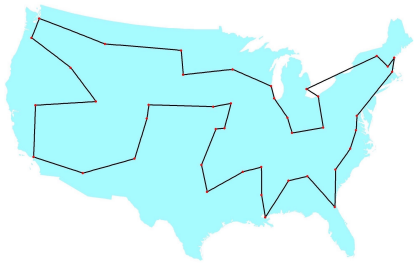
Open problems

- rounding matroids in parallel
- further (combinations) of constraints

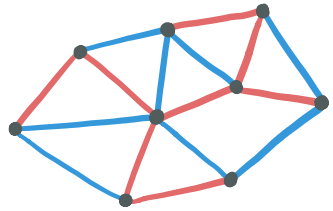
Other work

Written about comb. optimization, approximations, fast approximations, online learning, streaming, computational geometry... and trying to add more!

One line of work is on linear time approx. for implicit and explicit LP's. e.g.



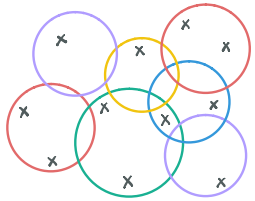
Metric TSP



Packing Spanning Trees (for network design)



Packing Intervals (for scheduling)

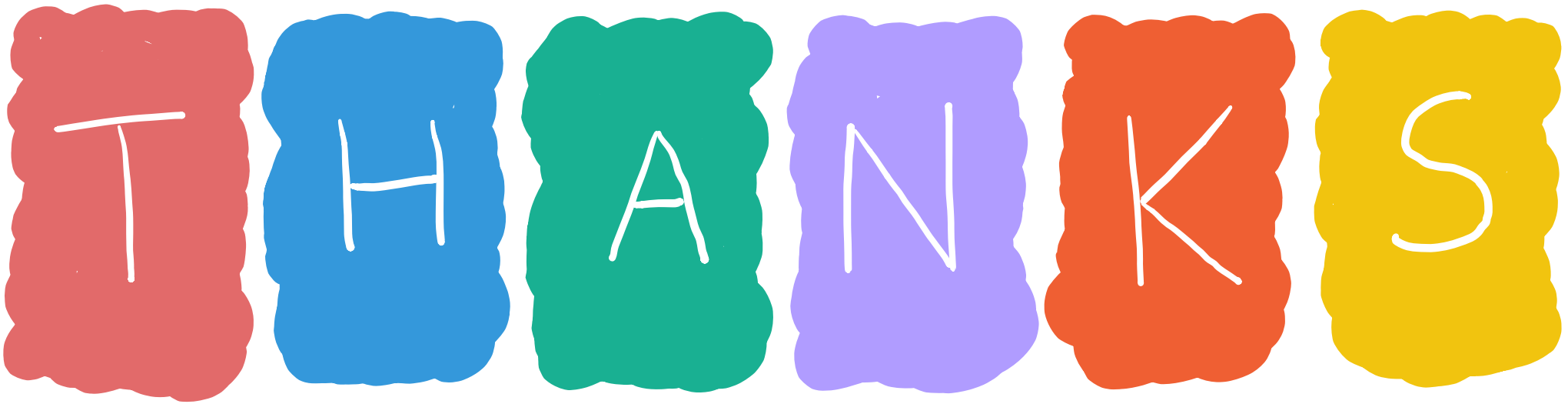


Covering points by discs (for comp. geometry)

Recurring themes:

discrete ⇅ interplay Continuous,

= FAST approx.


THANKS

(1) (I', S') is greedy block models

(namely, $E[f_S(I')] \geq (1-O(\varepsilon))\lambda E[|S'|]$)

Proof is subtle but main ideas are:

(1) δ is chosen conservatively small

(2) submodular functions and matroids are well-behaved under sampling

proof slides in appendix

"greedy sampling"

choose max $\delta > 0$ s.t.

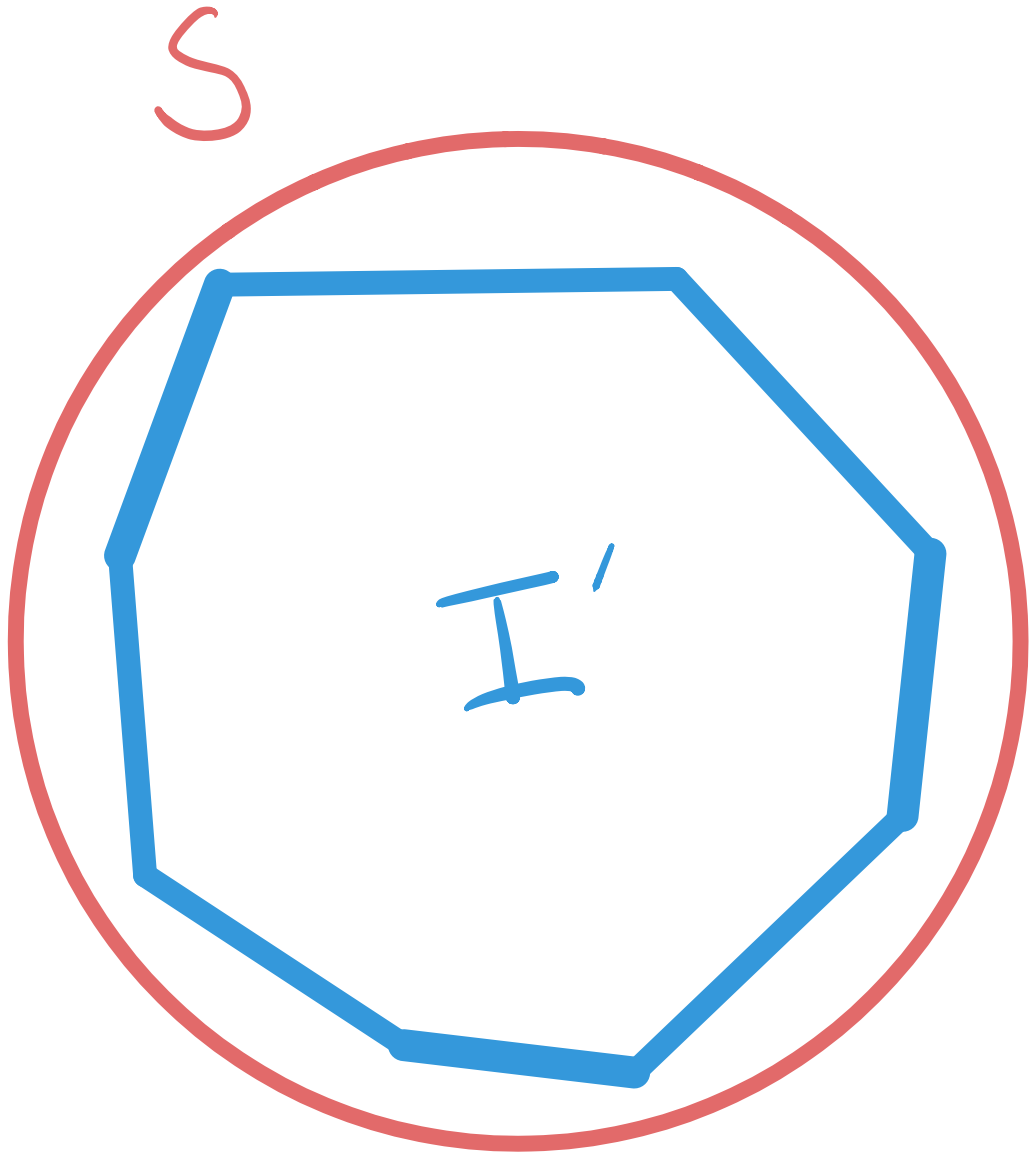
for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

(b) $\mathbb{E}[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

(b) $I' \leftarrow \{e \in S' : e \notin \mathrm{span}(S+S'-e)$ and $f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$



"greedy sampling"

choose max $\delta > 0$ s.t.

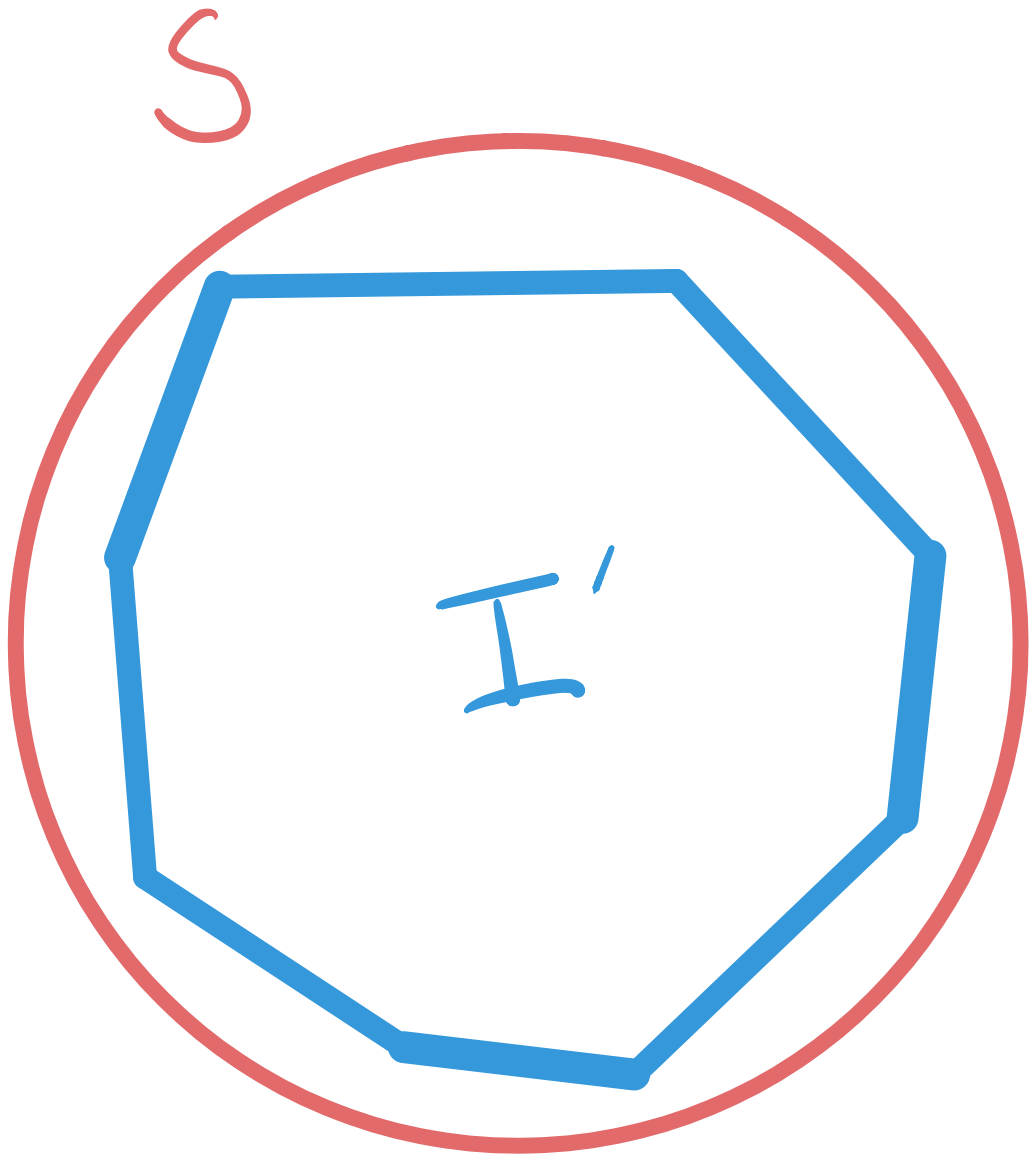
for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

(b) $I' \leftarrow \{e \in S' : e \notin \text{span}(S+S'-e)$ and $f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$



Define

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

"greedy sampling"

choose max $\delta > 0$ s.t.

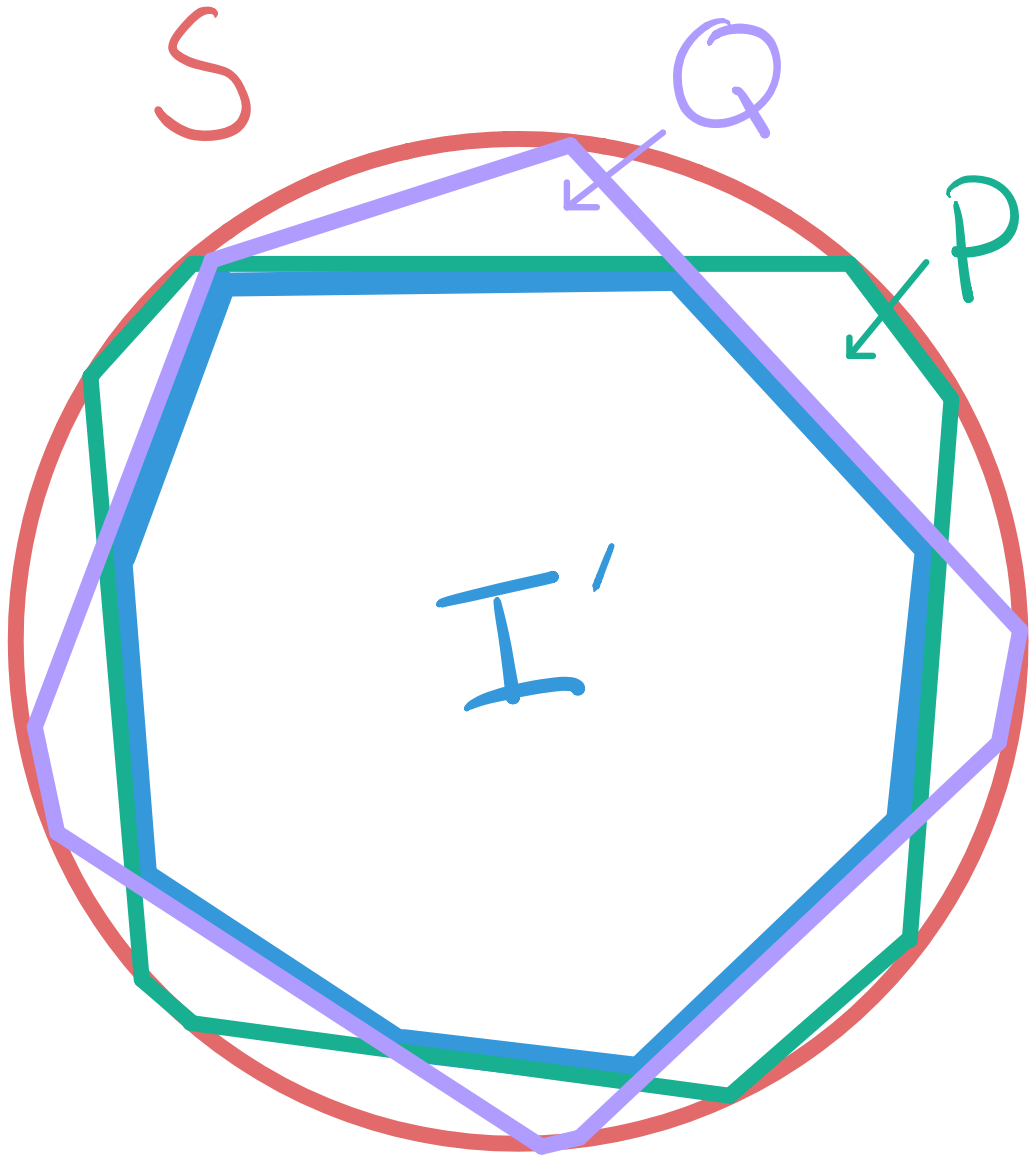
for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

(b) $I' \leftarrow \{e \in S' : e \notin \text{span}(S+S'-e)$ and $f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$



Define

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

"greedy sampling"

choose max $\delta > 0$ s.t.

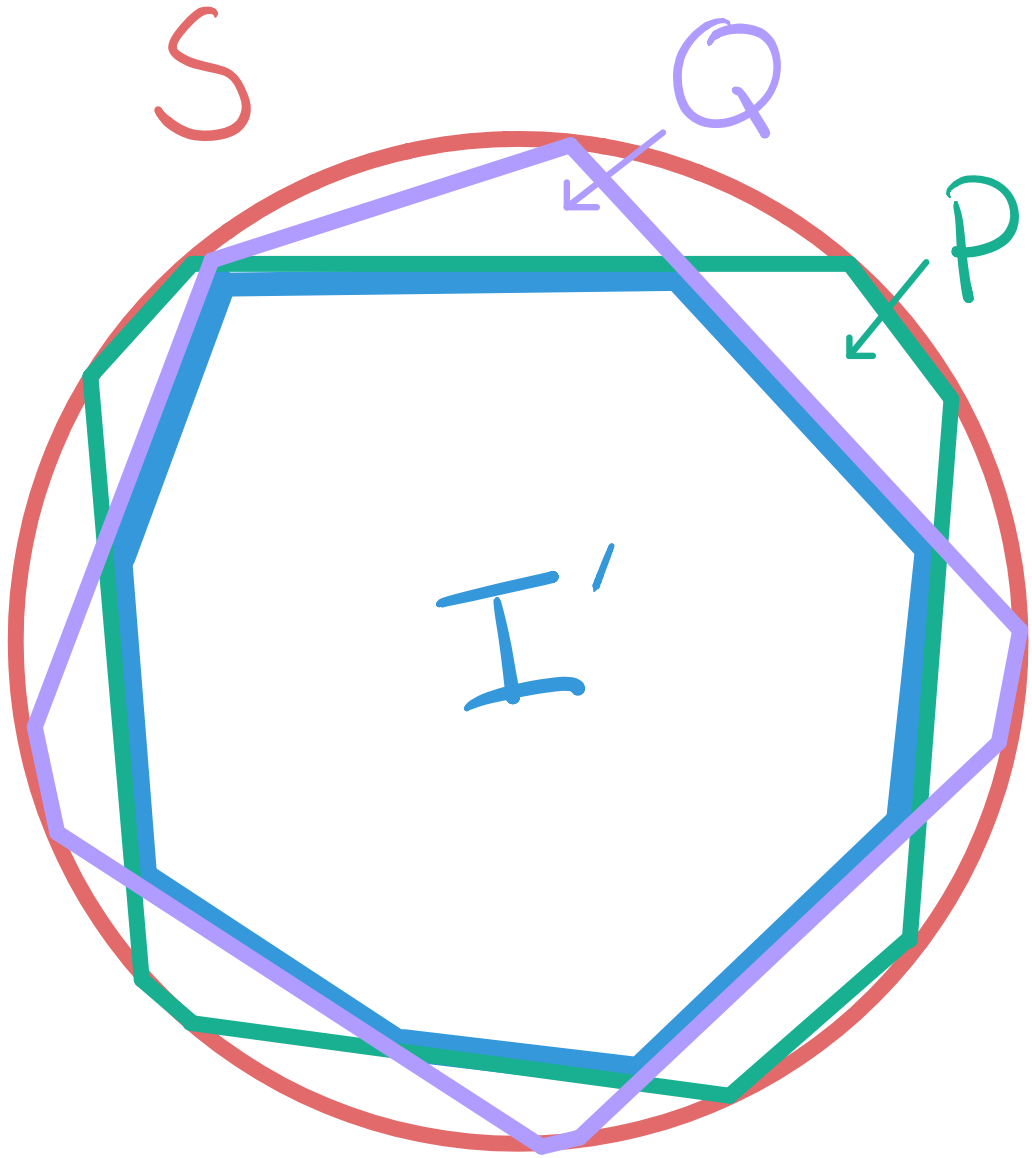
for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

(b) $I' \leftarrow \{e \in S' : e \notin \text{span}(S+S'-e)$ and $f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$



Define

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

Then $I' = Q \cap P$

3 claims

claim 1: $E[f_S(Q)] \geq (1-\varepsilon)\lambda E[|S'|]$

Q has good "bang-for-buck"

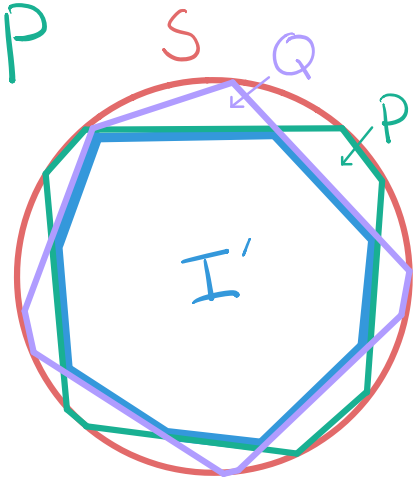
claim 2: $E[|S' \setminus P|] \leq \varepsilon[|S'|]$

most sampled elements are not in cycles, retained by P

$Q = \{e \in S' : f_{S+S'-e}(e) \geq (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(6) $I' \leftarrow Q \cap P$



claim 3: $E[f_S(I')] \geq (1-2\varepsilon)\lambda E[|S'|]$

I' has good "bang-for-buck"

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

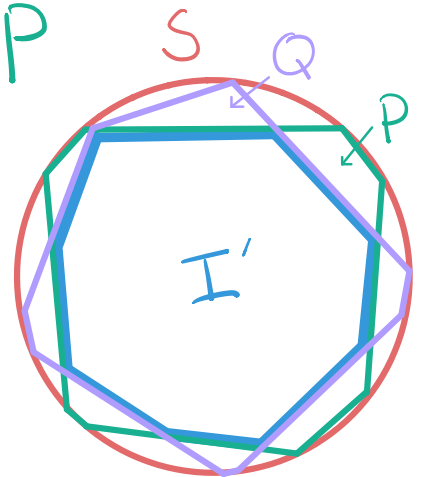
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



3 claims

claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

Q has good "bang-for-buck"

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

most sampled elements are not in cycles, retained by P

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

I' has good "bang-for-buck"

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

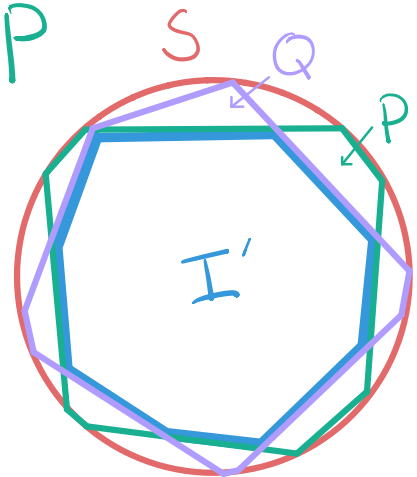
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



Proof ideas

claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

δ small enough that Q has good "bang-for-buck"

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

δ small enough that P keeps most of S'

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

submodularity allows us to combine claims 1 and 2 for $I' = Q \cap P$

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $E[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

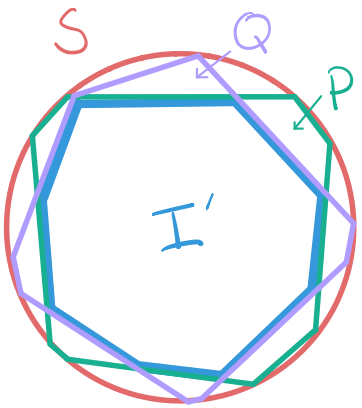
(b) $E[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



[skip]

claim 1: $E[f_S(Q)] \ge (1-\varepsilon)\lambda E[|S'|]$

Fix e.

$P[e \in Q] = P[e \in S' \ \& \ f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$

by definition of Q

claim 2: $E[|S' \setminus P|] \le \varepsilon E[|S'|]$

claim 3: $E[f_S(I')] \ge (1-2\varepsilon)\lambda E[|S'|]$

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

(b) $\mathbb{E}[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \le \varepsilon |N'|$

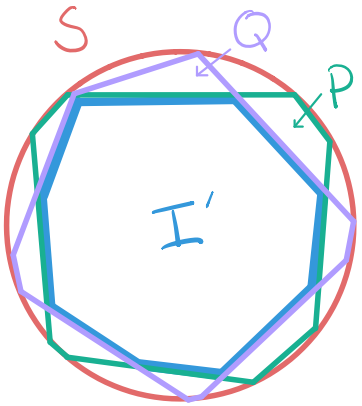
(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \mathrm{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$

[skip]



claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

Fix e.

$\mathbb{P}[e \in Q] = \mathbb{P}[e \in S' \ \&\ f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$

$= \mathbb{P}[e \in S'] \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$

by independence of &

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

(b) $\mathbb{E}[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \le \varepsilon |N'|$

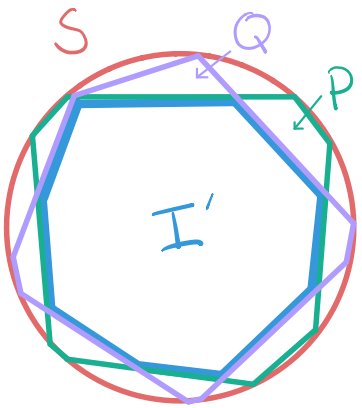
(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \mathrm{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$

[skip]



claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

Fix e.

$$\mathbb{P}[e \in Q] = \mathbb{P}[e \in S' \ \&\ f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$$

$$= \mathbb{P}[e \in S'] \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$$

$$= \delta \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$$

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

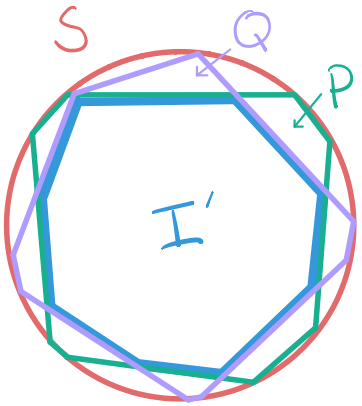
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



[skip]

claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

$\mathbb{P}[e \in Q] = \delta \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$

$\mathbb{E}[f_S(Q)] \ge \sum_{e \in N'} \mathbb{P}[e \in Q] \mathbb{E}[f_{S+Q-e}(e) \mid e \in Q]$

(by submodularity)

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

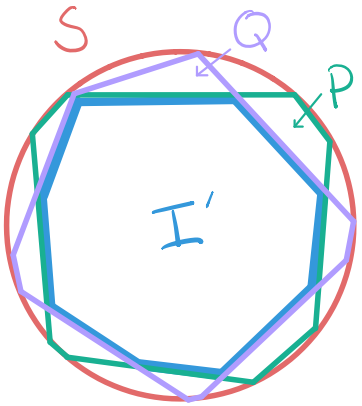
(b) $\mathbb{E}[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \mathrm{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



[skip]

claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

(*) $\mathbb{P}[e \in Q] = \delta \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$

$$\mathbb{E}[f_S(Q)] \ge \sum_{e \in N'} \mathbb{P}[e \in Q]\, \mathbb{E}[f_{S+Q-e}(e) \mid e \in Q]$$

$$= \delta \sum_{e \in N'} \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]\, \mathbb{E}[f_{S+Q-e}(e) \mid e \in Q]$$

by the above (*)

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

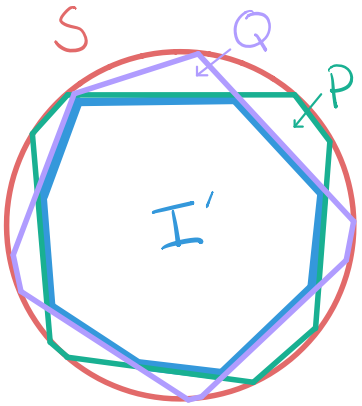
(b) $\mathbb{E}[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \mathrm{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



[skip]

claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

$$\mathbb{P}[e \in Q] = \delta \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$$

$$\mathbb{E}[f_S(Q)] \ge \sum_{e \in N'} \mathbb{P}[e \in Q]\, \mathbb{E}[f_{S+Q-e}(e) \mid e \in Q]$$

$$= \delta \sum_{e \in N'} \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]\, \mathbb{E}[f_{S+Q-e}(e) \mid e \in Q]$$

$$\ge \delta \sum_{e \in N'} \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]\, \mathbb{E}[f_{S+S'-e}(e) \mid e \in Q]$$

by submodularity

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

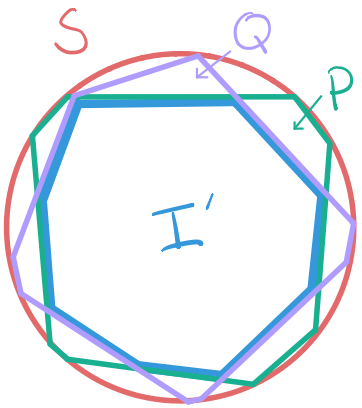
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



[skip]

claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

$$\mathbb{P}[e \in Q] = \delta \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$$

$$\mathbb{E}[f_S(Q)] \ge \sum_{e \in N'} \mathbb{P}[e \in Q]\, \mathbb{E}[f_{S+Q-e}(e) \mid e \in Q]$$

$$= \delta \sum_{e \in N'} \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]\, \mathbb{E}[f_{S+Q-e}(e) \mid e \in Q]$$

$$\ge \delta \sum_{e \in N'} \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]\, \mathbb{E}[f_{S+S'-e}(e) \mid e \in Q]$$

$$\ge \delta(1-\varepsilon)\lambda \sum_{e \in N'} \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$$

since $e \in Q \Rightarrow f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda$

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon |S'|$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

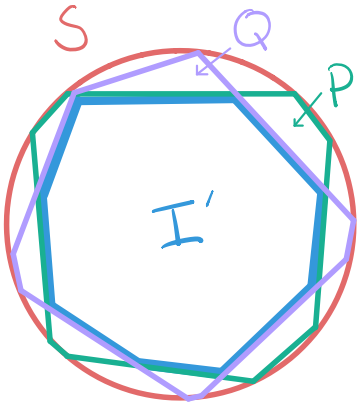
(b) $\mathbb{E}[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \mathrm{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



[skip]

claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

$$\mathbb{P}[e \in Q] = \delta \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$$

$$\mathbb{E}[f_S(Q)] \ge \sum_{e \in N'} \mathbb{P}[e \in Q]\, \mathbb{E}[f_{S+Q-e}(e) \mid e \in Q]$$

$$= \delta \sum_{e \in N'} \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]\, \mathbb{E}[f_{S+Q-e}(e) \mid e \in Q]$$

$$\ge \delta \sum_{e \in N'} \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]\, \mathbb{E}[f_{S+S'-e}(e) \mid e \in Q]$$

$$\ge \delta(1-\varepsilon)\lambda \sum_{e \in N'} \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$$

$$\ge \delta(1-\varepsilon)^2 |N'|$$

by choice of δ per (a)

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

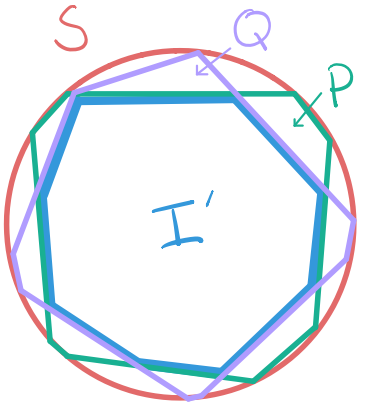
(b) $\mathbb{E}[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \mathrm{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



[skip]

claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

$$\mathbb{P}[e \in Q] = \delta \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$$

$$\mathbb{E}[f_S(Q)] \ge \sum_{e \in N'} \mathbb{P}[e \in Q]\, \mathbb{E}[f_{S+Q-e}(e) \mid e \in Q]$$

$$= \delta \sum_{e \in N'} \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]\, \mathbb{E}[f_{S+Q-e}(e) \mid e \in Q]$$

$$\ge \delta \sum_{e \in N'} \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]\, \mathbb{E}[f_{S+S'-e}(e) \mid e \in Q]$$

$$\ge \delta(1-\varepsilon)\lambda \sum_{e \in N'} \mathbb{P}[f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda]$$

$$\ge \delta(1-\varepsilon)^2 |N'| \ge (1-\varepsilon)^2 \lambda \mathbb{E}[|S|]$$

since $S \sim \delta N'$

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

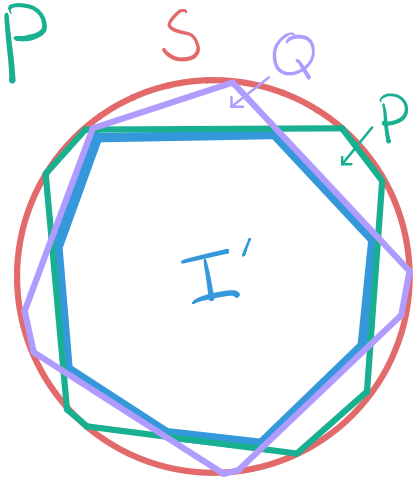
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



Proof ideas

claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

δ small enough that Q has good "bang-for-buck"

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

δ small enough that P keeps most of S'

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

submodularity allows us to combine claims 1 and 2 for $I' = Q \cap P$

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

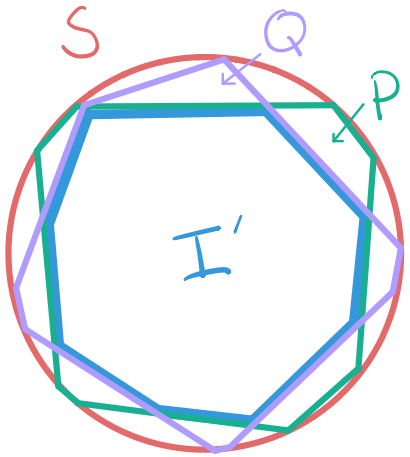
(b) $\mathbb{E}[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \mathrm{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

fix e.

events $[e \in S']$, $[e \in \mathrm{span}(S+S'-e)]$ are independent. so

$$\mathbb{P}[e \in S' \setminus P] = \underbrace{\mathbb{P}[e \in S']}_{=\delta} \mathbb{P}[e \in \mathrm{span}(S+S'-e)]$$

$$\mathbb{E}[|S' \setminus P|] = \delta \sum_{e \in N} \mathbb{P}[e \in \mathrm{span}(S-e)]$$

$$\le \delta \mathbb{E}[|\mathrm{span}(S)|]$$

$$\le \delta \varepsilon |N'| = \varepsilon \mathbb{E}[|S'|]$$

[skip]

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $E[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

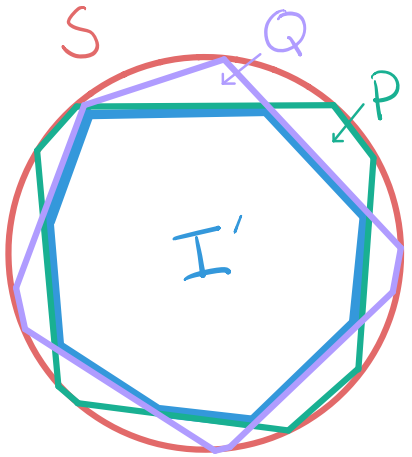
(b) $E[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



✓ claim 1: $E[f_S(Q)] \ge (1-\varepsilon)\lambda E[|S'|]$

δ small enough that Q has good "bang-for-buck"

✓ claim 2: $E[|S' \setminus P|] \le \varepsilon E[|S'|]$

δ small enough that P keeps most of S'

claim 3: $E[f_S(I')] \ge (1-2\varepsilon)\lambda E[|S'|]$

submodularity allows us to combine claims 1 and 2 for $I' = Q \cap P$

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

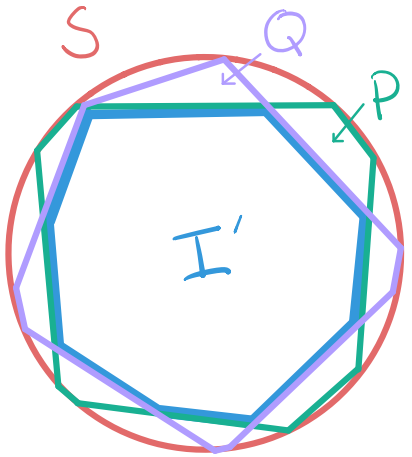
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

$\mathbb{E}[f_S(I')] = \mathbb{E}[f_S(Q)] - \mathbb{E}[f_{S+I'}(Q)]$

[since $I' \le Q$]

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

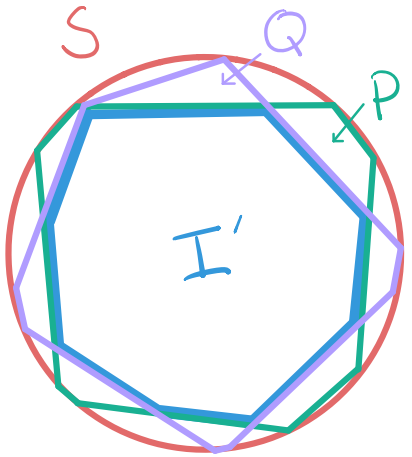
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

$$\mathbb{E}[f_S(I')] = \mathbb{E}[f_S(Q)] - \mathbb{E}[f_{S+I'}(Q)]$$
$$\ge \mathbb{E}[f_S(Q)] - \sum_{e \in Q} [f_{S+I'}(e)]$$

[by submodularity]

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

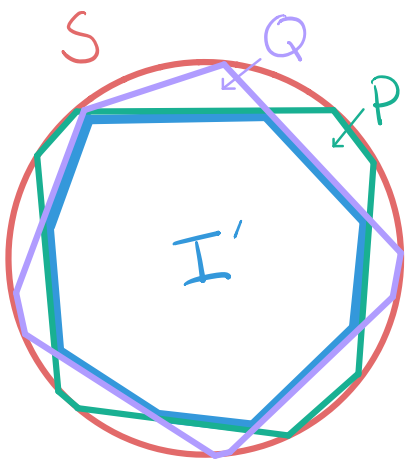
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

$$\mathbb{E}[f_S(I')] = \mathbb{E}[f_S(Q)] - \mathbb{E}[f_{S+I'}(Q)]$$
$$\ge \mathbb{E}[f_S(Q)] - \sum_{e \in Q} \mathbb{E}[f_{S+I'}(e)]$$
$$\ge \mathbb{E}[f_S(Q)] - \sum_{e \in Q} P[e \in Q \setminus I'] f_S(e)$$

[by submodularity]

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

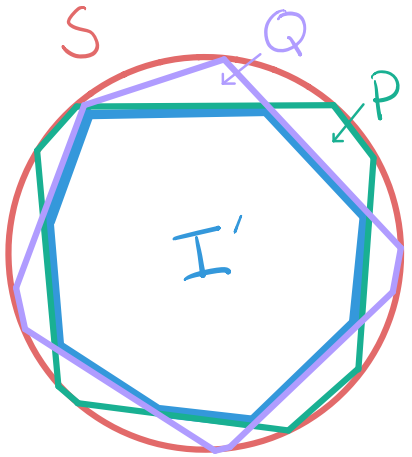
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

$$\mathbb{E}[f_S(I')] = \mathbb{E}[f_S(Q)] - \mathbb{E}[f_{S+I'}(Q)]$$
$$\ge \mathbb{E}[f_S(Q)] - \sum_{e \in Q} \mathbb{E}[f_{S+I'}(e)]$$
$$\ge \mathbb{E}[f_S(Q)] - \sum_{e \in Q} P[e \in Q \setminus I'] f_S(e)$$
$$\ge \mathbb{E}[f_S(Q)] - \lambda \mathbb{E}[|Q \setminus I'|]$$

[by $\lambda \ge \max_{e \in \text{span}(S)} f_S(e)$]

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

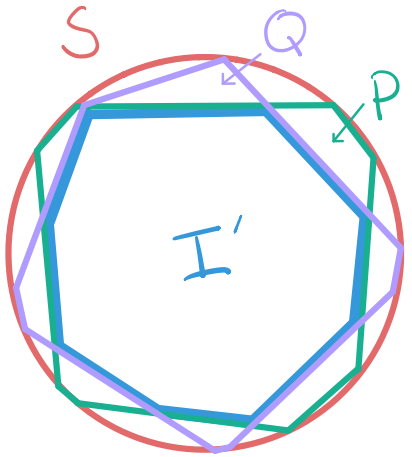
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

$$\mathbb{E}[f_S(I')] = \mathbb{E}[f_S(Q)] - \mathbb{E}[f_{S+I'}(Q)]$$

$$\ge \mathbb{E}[f_S(Q)] - \sum_{e \in Q} [f_{S+I'}(e)]$$

$$\ge \mathbb{E}[f_S(Q)] - \sum_{e \in Q} P[e \in Q \setminus I'] f_S(e)$$

$$\ge \mathbb{E}[f_S(Q)] - \lambda \mathbb{E}[|Q \setminus I'|]$$

$$\ge \mathbb{E}[f(Q)] - \lambda \mathbb{E}[|S' \setminus P|]$$

[by $Q \setminus I \subseteq S \setminus P$]

"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

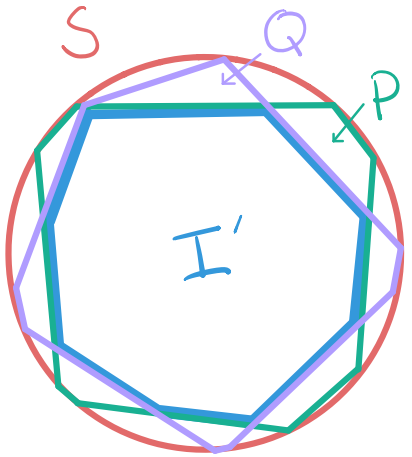
(b) $\mathbb{E}[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \mathrm{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

$$\mathbb{E}[f_S(I')] = \mathbb{E}[f_S(Q)] - \mathbb{E}[f_{S+I'}(Q)]$$
$$\ge \mathbb{E}[f_S(Q)] - \sum_{e \in Q} \mathbb{E}[f_{S+I'}(e)]$$
$$\ge \mathbb{E}[f_S(Q)] - \sum_{e \in Q} \mathbb{P}[e \in Q \setminus I'] f_S(e)$$
$$\ge \mathbb{E}[f_S(Q)] - \lambda \mathbb{E}[|Q \setminus I'|]$$
$$\ge \mathbb{E}[f(Q)] - \lambda \mathbb{E}[|S' \setminus P|]$$
$$\ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$$

by claim 1 and claim 2



"greedy sampling"

choose max $\delta > 0$ s.t.

for $S' \sim \delta S$:

(a) $\mathbb{E}[|\{e \in N' : f_{S+S'}(e) \le (1-\varepsilon)\lambda\}|] \le \varepsilon |N'|$

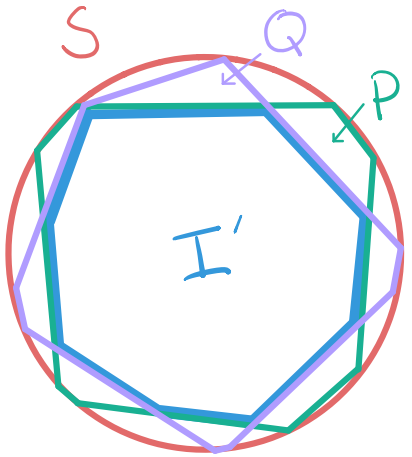
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

(a) sample $S' \sim \delta N'$

$Q = \{e \in S' : f_{S+S'-e}(e) \ge (1-\varepsilon)\lambda\}$

$P = \{e \in S' : e \notin \text{span}(S+S'-e)\}$

(b) $I' \leftarrow Q \cap P$



✓ claim 1: $\mathbb{E}[f_S(Q)] \ge (1-\varepsilon)\lambda \mathbb{E}[|S'|]$

δ small enough that Q has good "bang-for-buck"

✓ claim 2: $\mathbb{E}[|S' \setminus P|] \le \varepsilon \mathbb{E}[|S'|]$

δ small enough that P keeps most of S'

✓ claim 3: $\mathbb{E}[f_S(I')] \ge (1-2\varepsilon)\lambda \mathbb{E}[|S'|]$

submodularity allows us to combine claims 1 and 2 for $I' = Q \cap P$