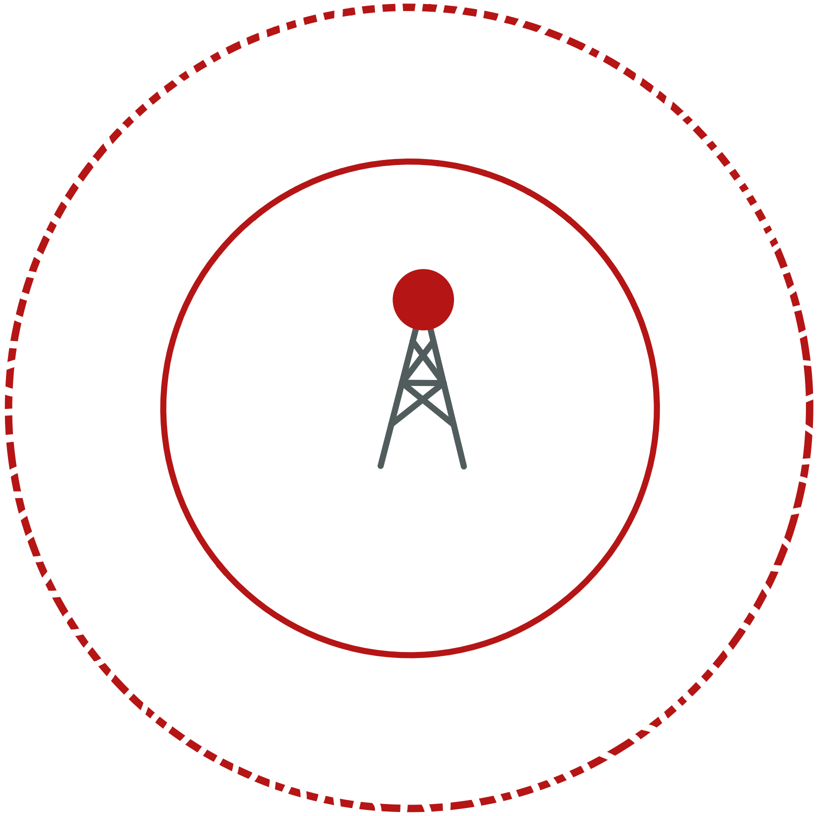
Parallelizing Greedy for Submodular Function Maximization

Kent Quanrud, UIUC

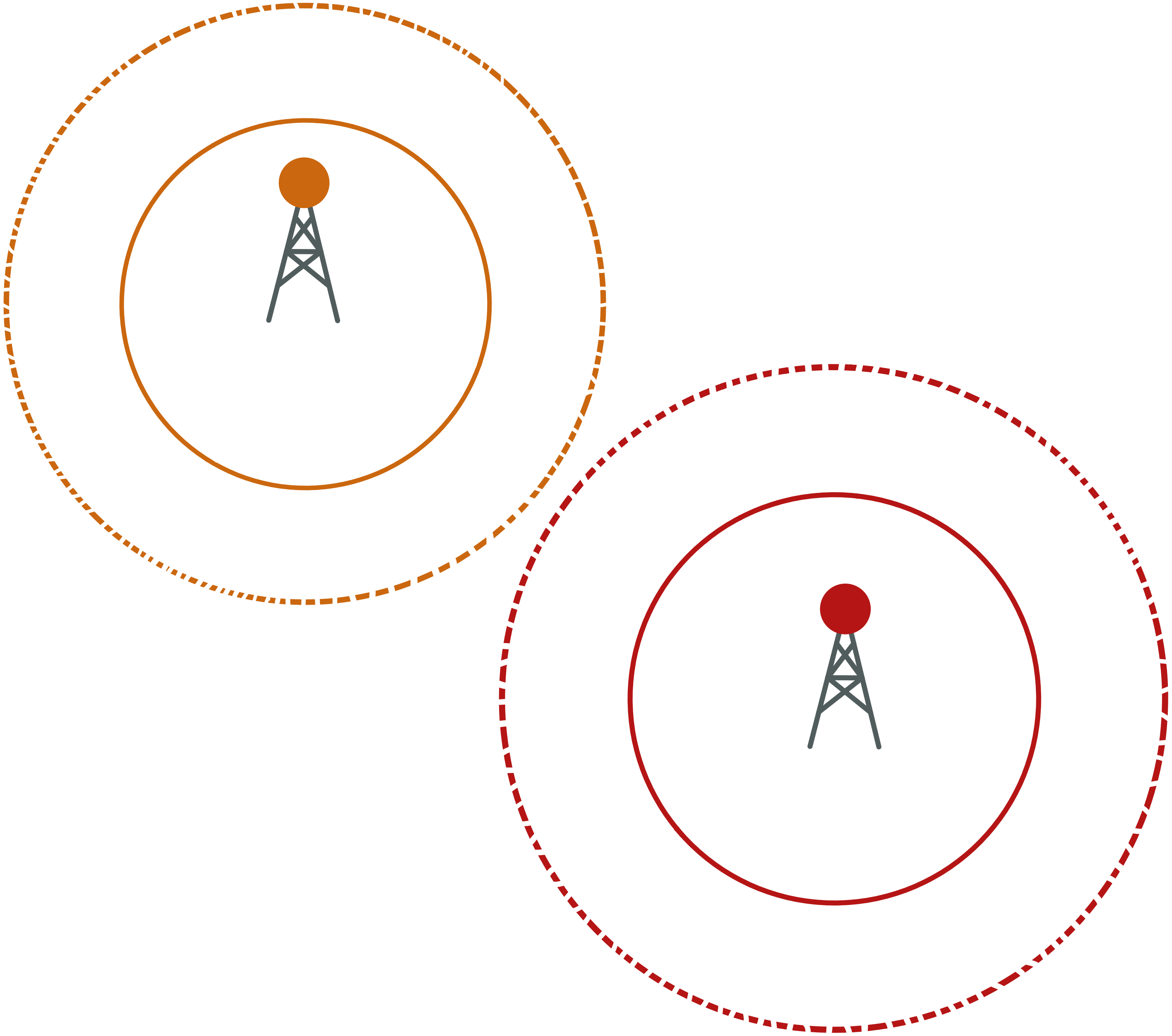
joint work w/ Chandra Chekuri, UIUC

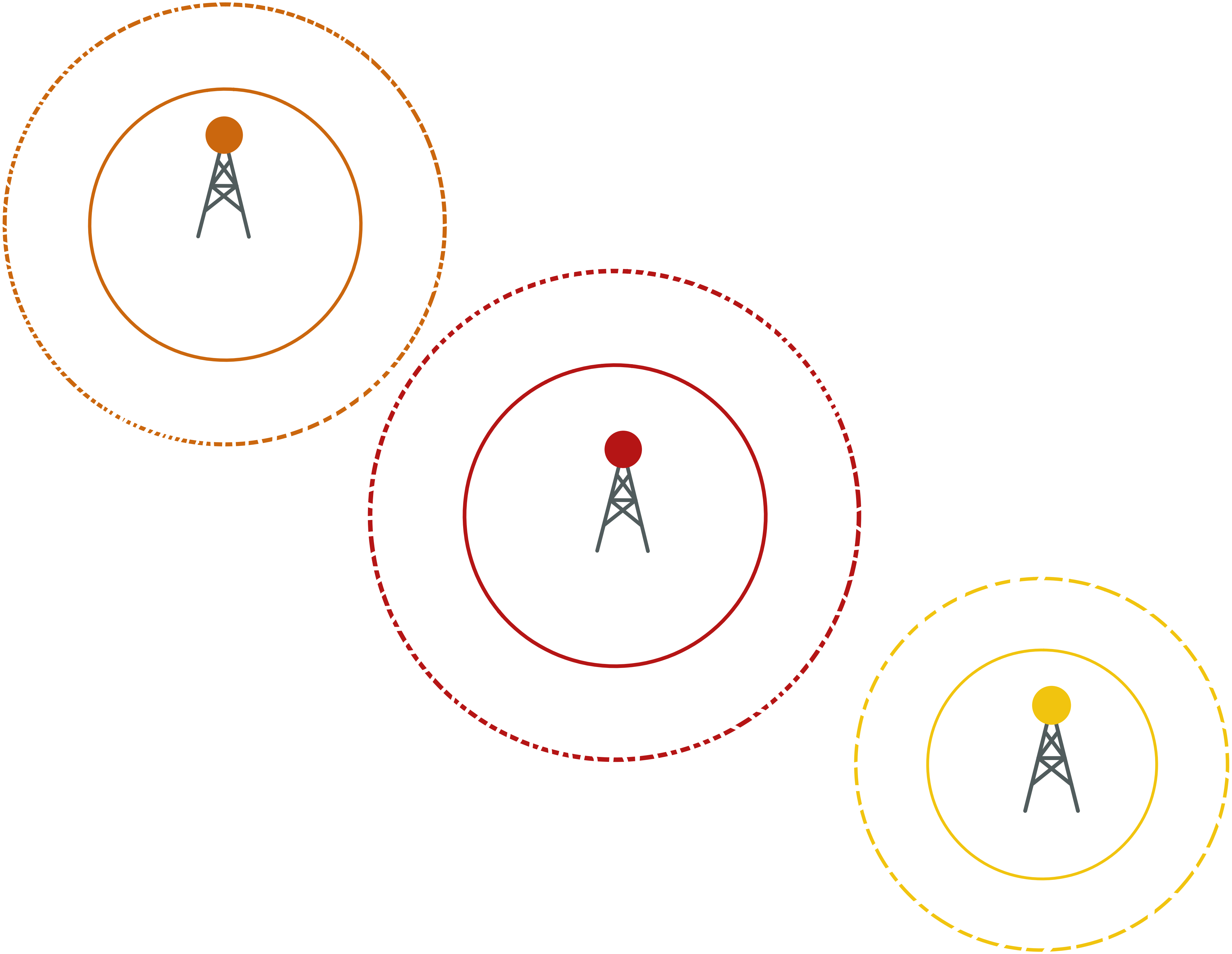
Bellairs Workshop on Discrete Optimization

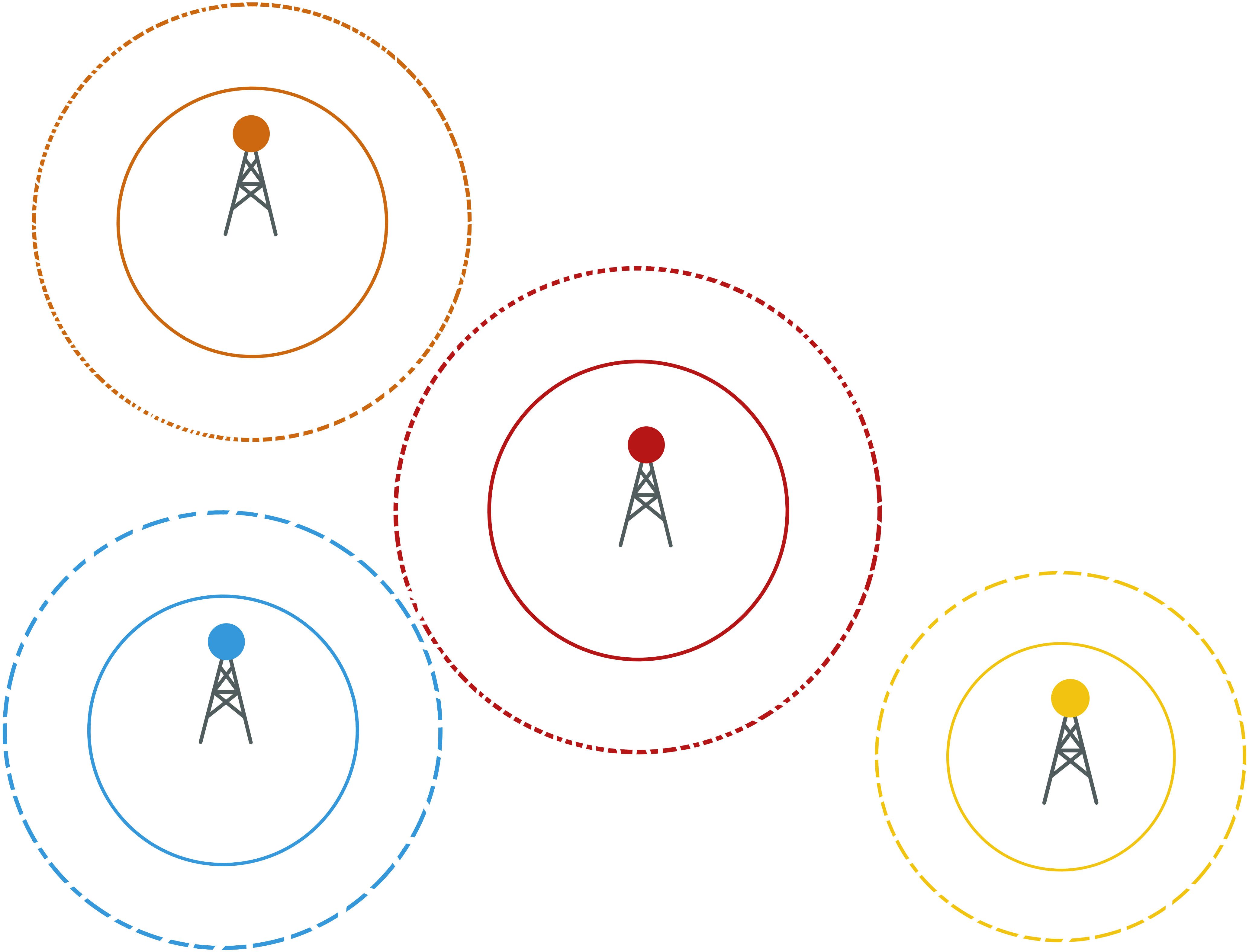
April 17, 2019

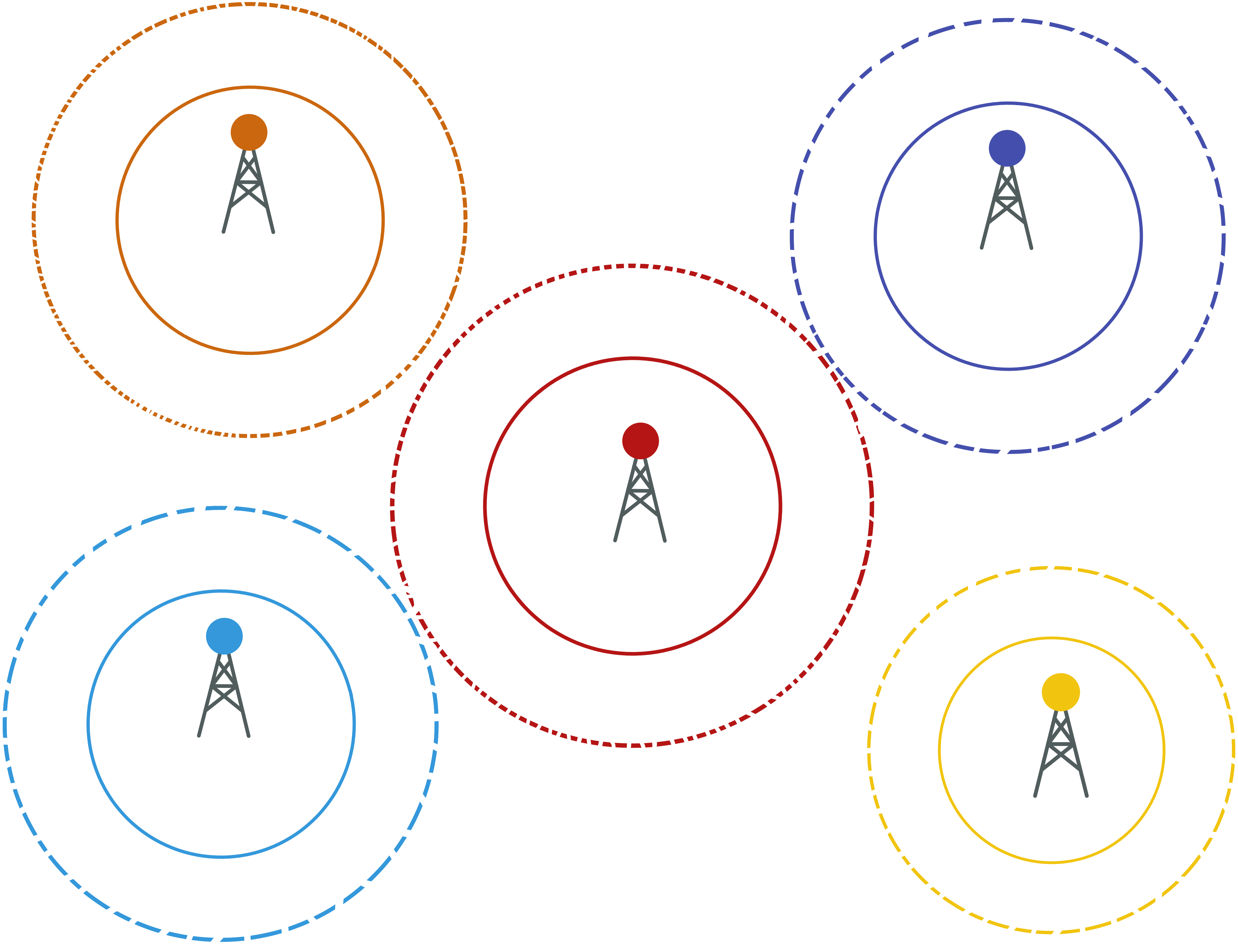


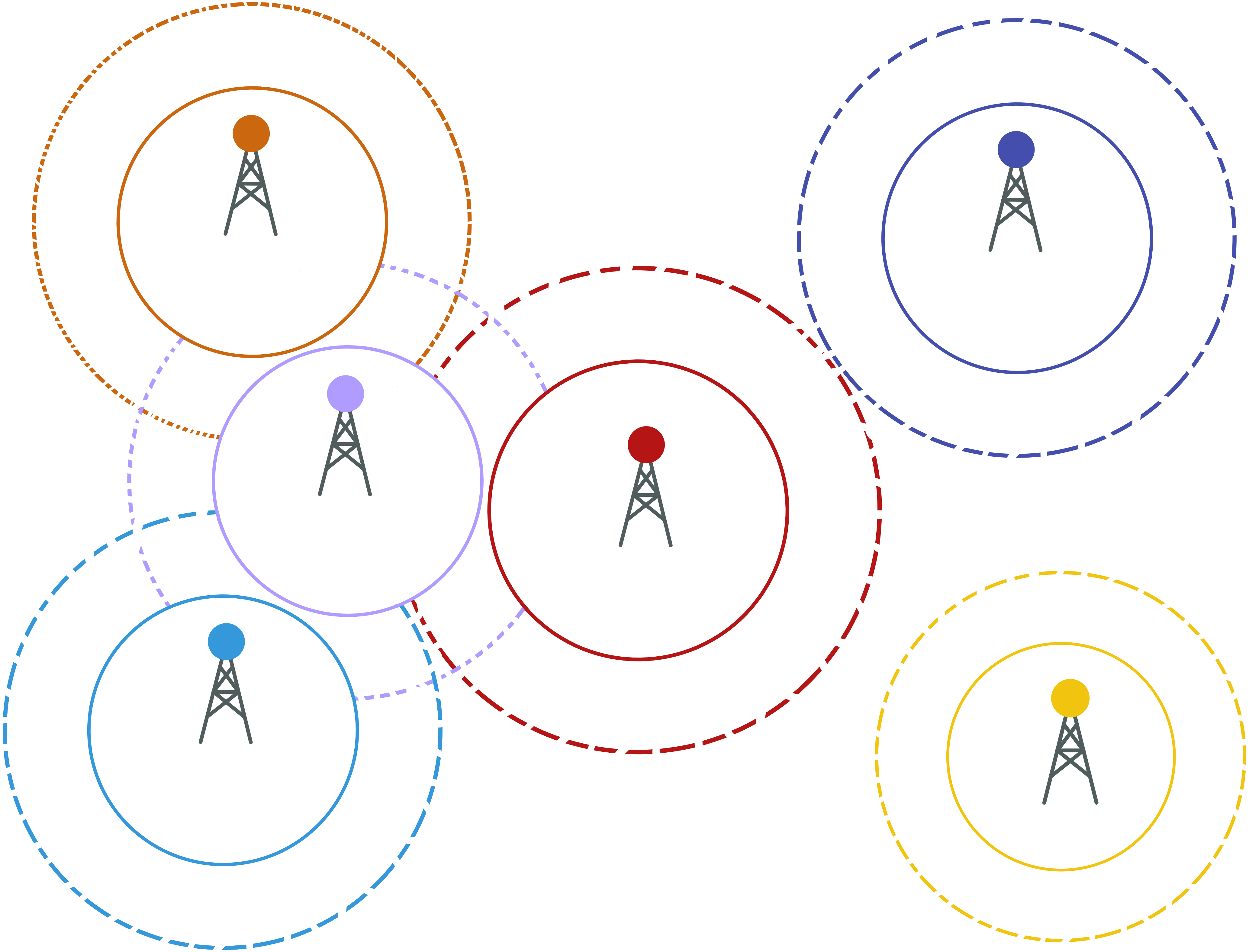
Sensor placement

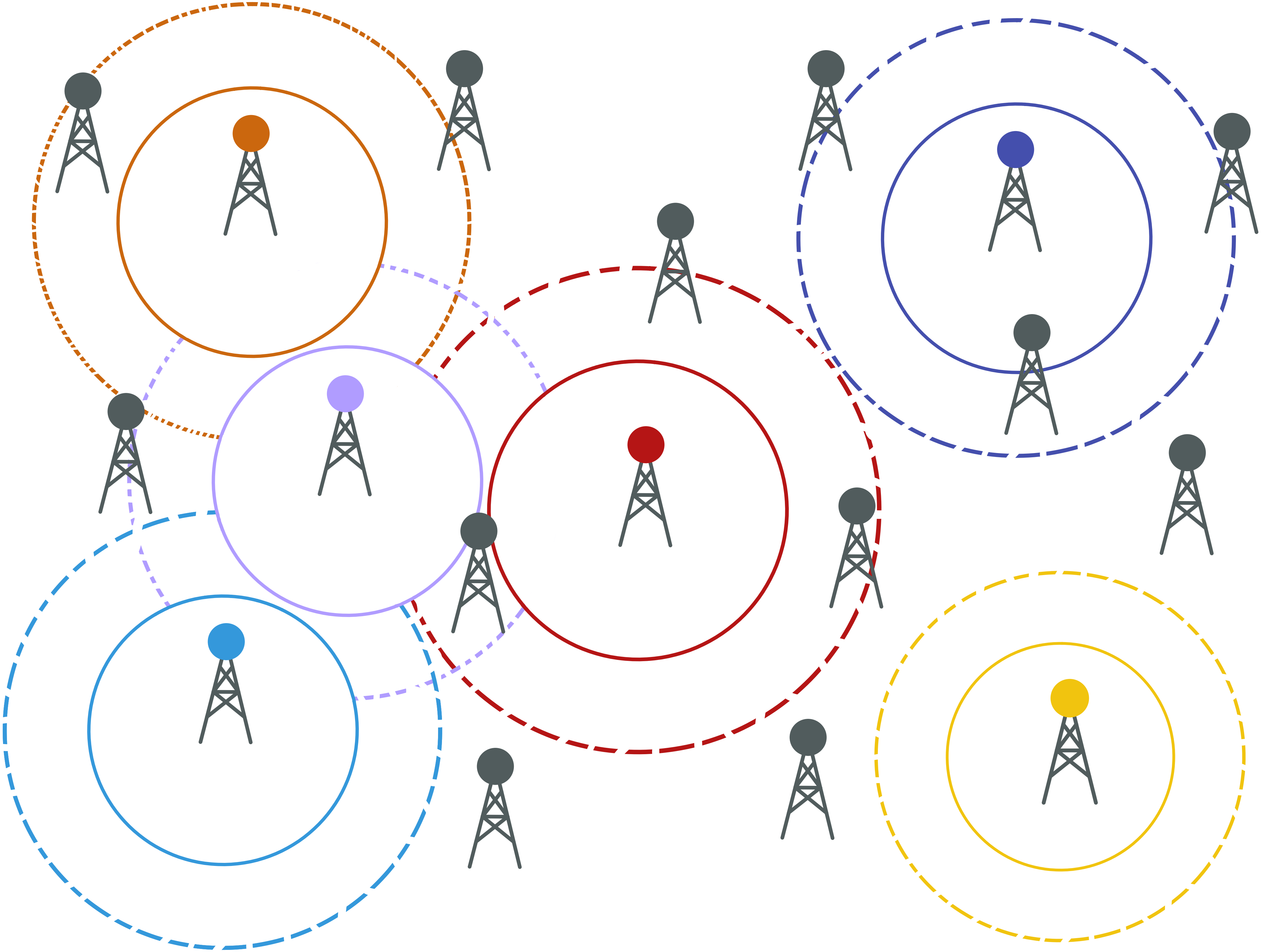








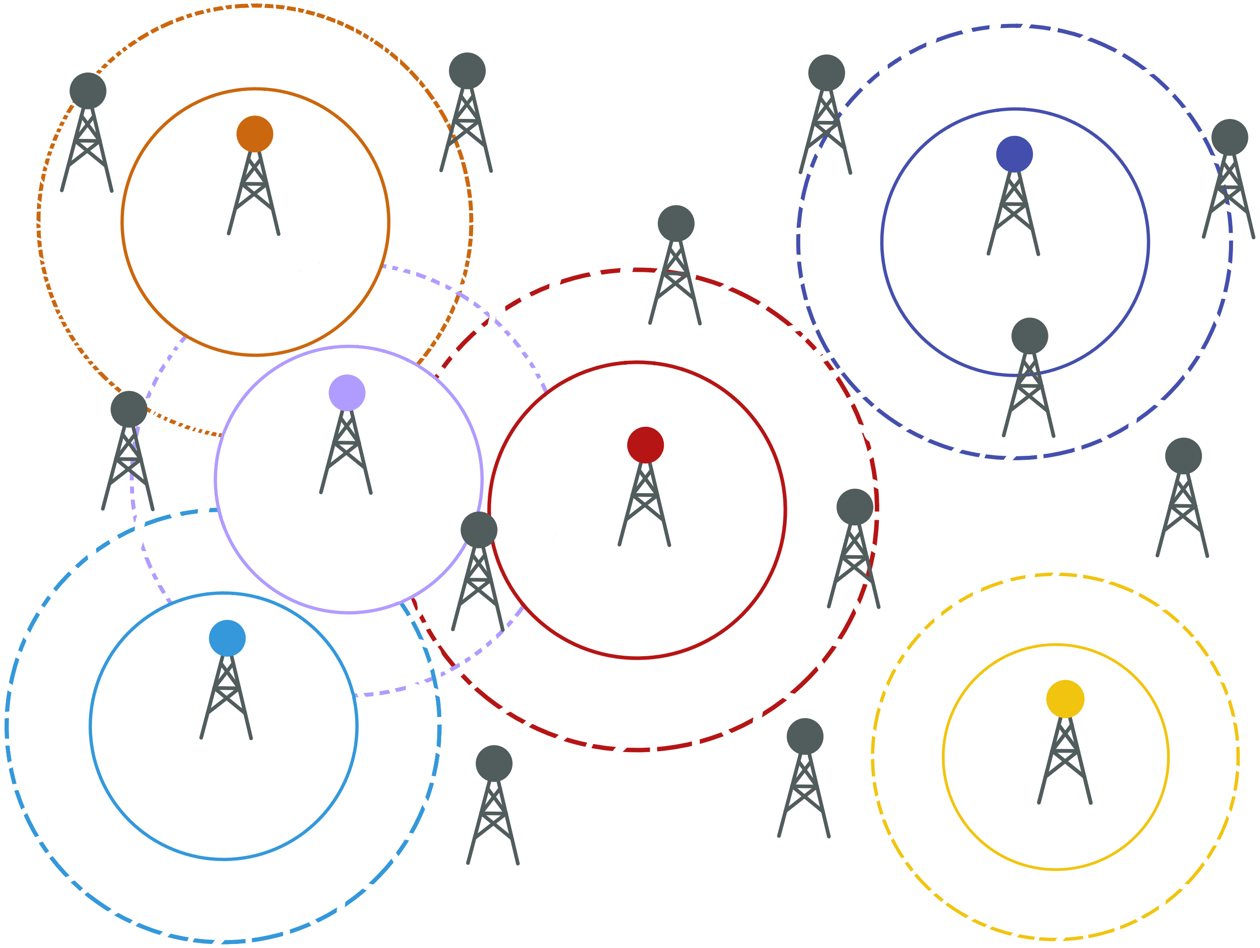




Maximum coverage

Input: sets $S_1, \ldots, S_n$; cardinality $k \in \mathbb{N}$

Goal: select k sets $S_{i_1}, \ldots, S_{i_k}$ maximizing their union $\left|\bigcup_{j=1}^{k} S_{i_j}\right|$



NP-Hard

$1 - 1/e$ APX-hardness

$1 - 1/e$ 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" $f: 2^N \to \mathbb{R}_{\geq 0}$

denote $\underbrace{f_S(Z) = f(S \cup Z) - f(S)}_{\text{"marginal value of } Z \text{ to } S\text{"}}$

[decreasing marginal returns] $S \subseteq T \Rightarrow f_S(Z) \geq f_T(Z)$

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

We say f is "monotone" if

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

[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

random arrival streams

parallel (PRAM) models

MapReduce algorithms

oracle efficient algorithms

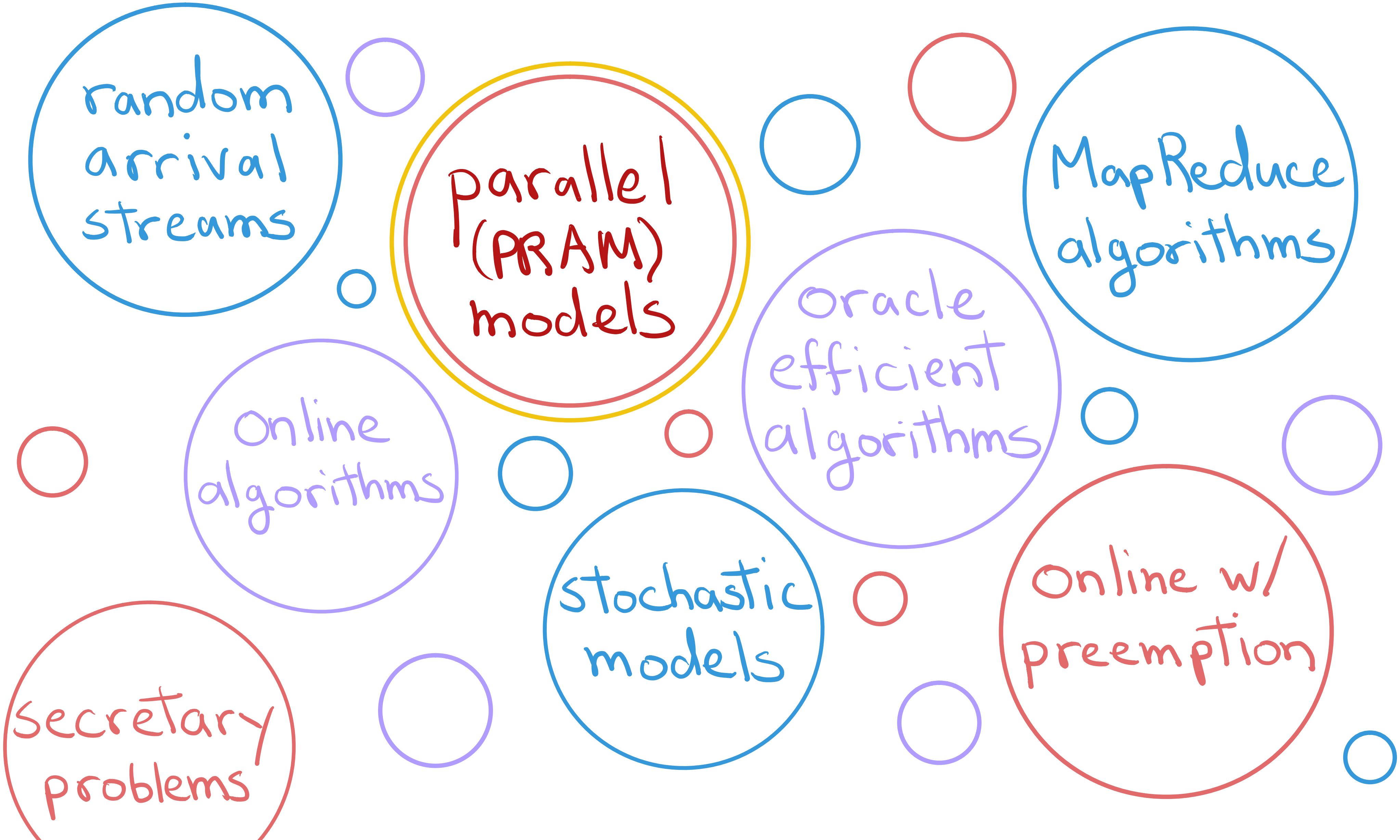
online algorithms

stochastic models

online w/ preemption

secretary problems

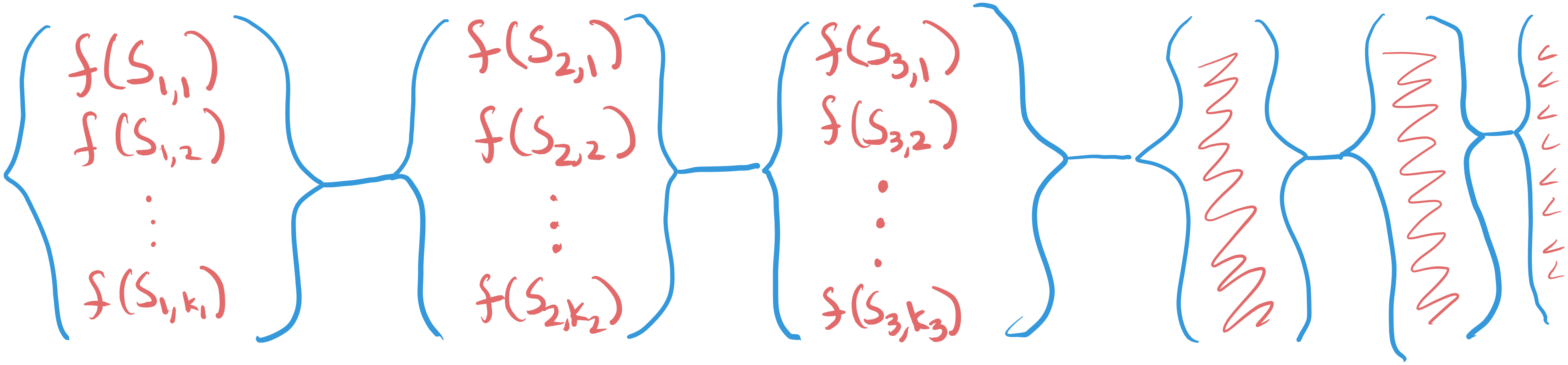
Applications ⇒ need for efficient algorithms in new models of computation



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



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] (SODA '19) [Ene, Nguyen] $(1-\frac{1}{e}-\varepsilon)$-APX in $O\left(\frac{\log n}{\varepsilon^2}\right)$ rounds

Our results: beyond cardinality

Packing constraints

[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)

Parallelizing greedy for cardinality and beyond

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

[Fisher, Nemhauser, Wolsey]

[this talk]

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

2nd half of talk

general packing constraints in parallel

matroid constraints in parallel

Initially $I=\emptyset$

query

$f_I(e)=f(I+e)-f(I)$

for all $e\in N$

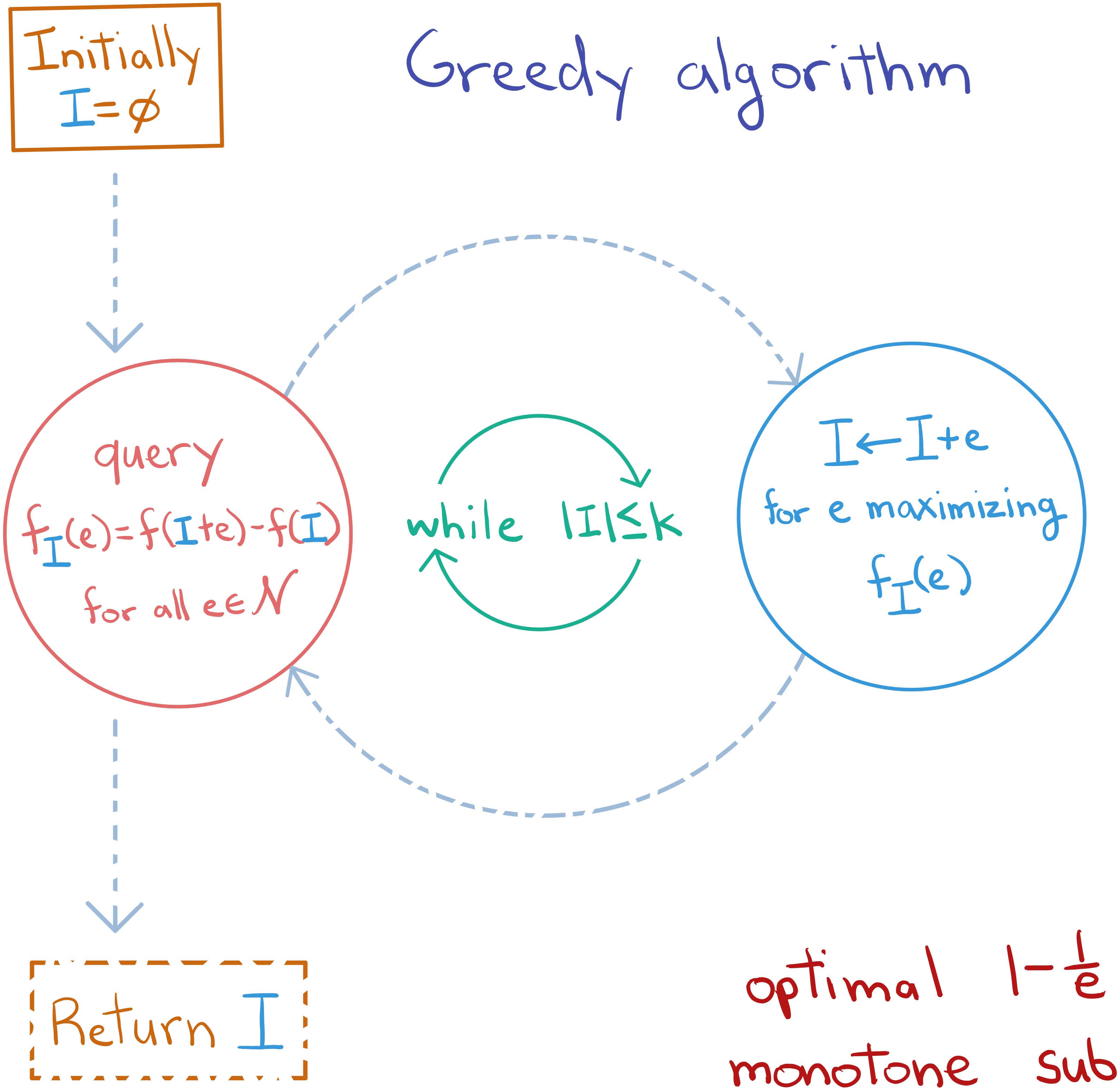
Return I

Greedy algorithm

[Nemhauser, Wolsey, Fischer '78]

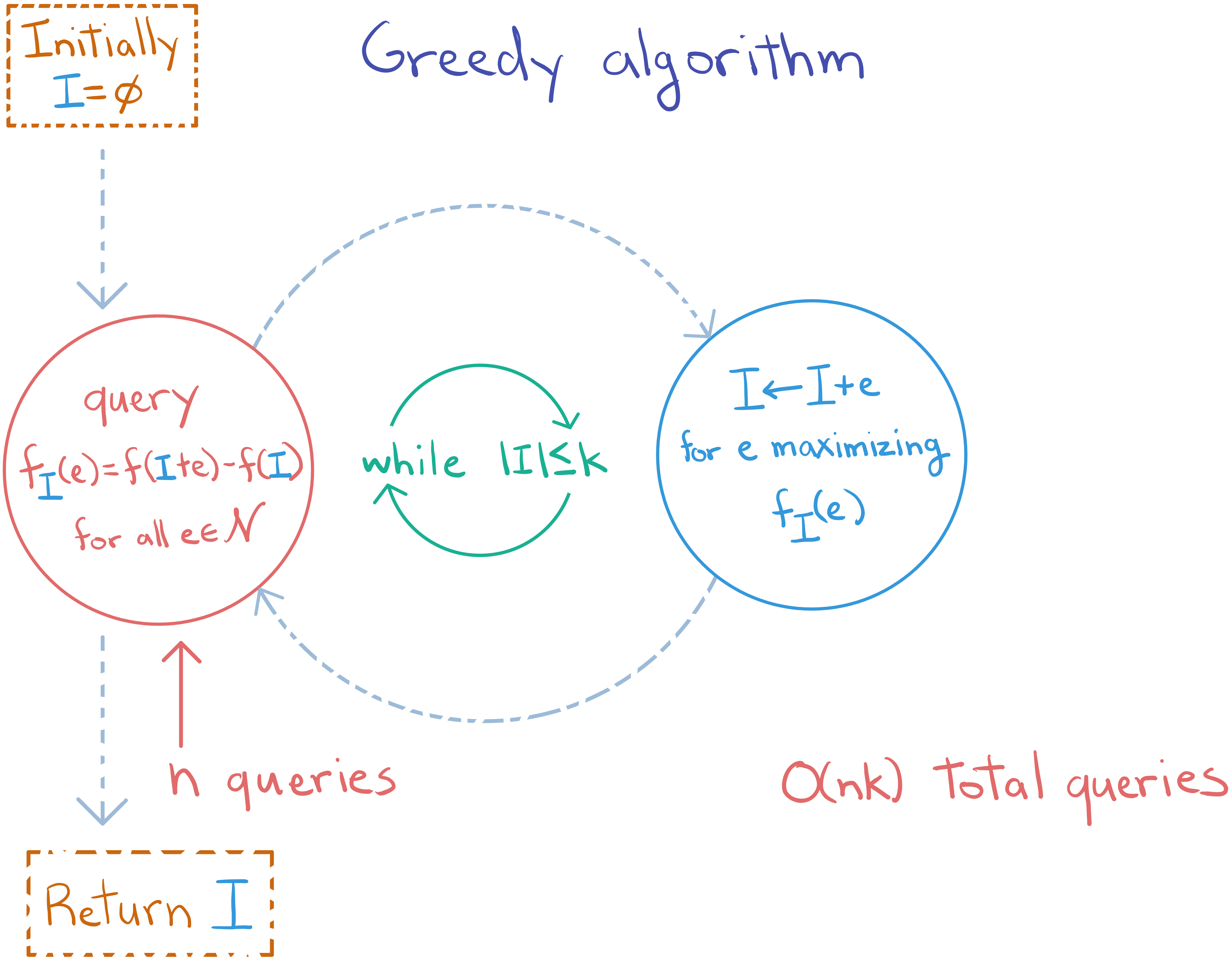
optimal $1-\frac{1}{e}$ APX for monotone submodular f

Greedy algorithm

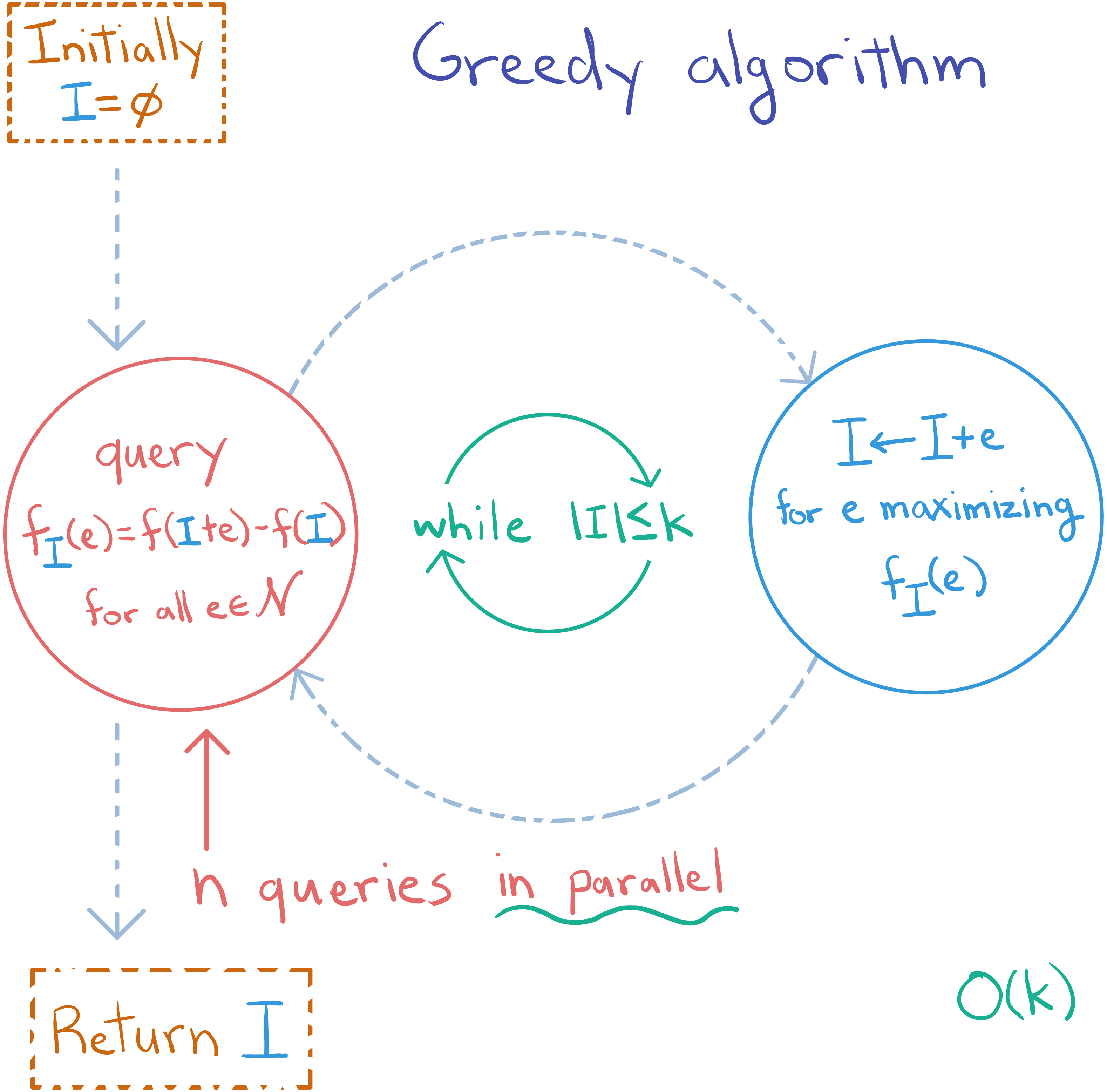


optimal $1-\frac{1}{e}$ APX for monotone submodular f

Greedy algorithm

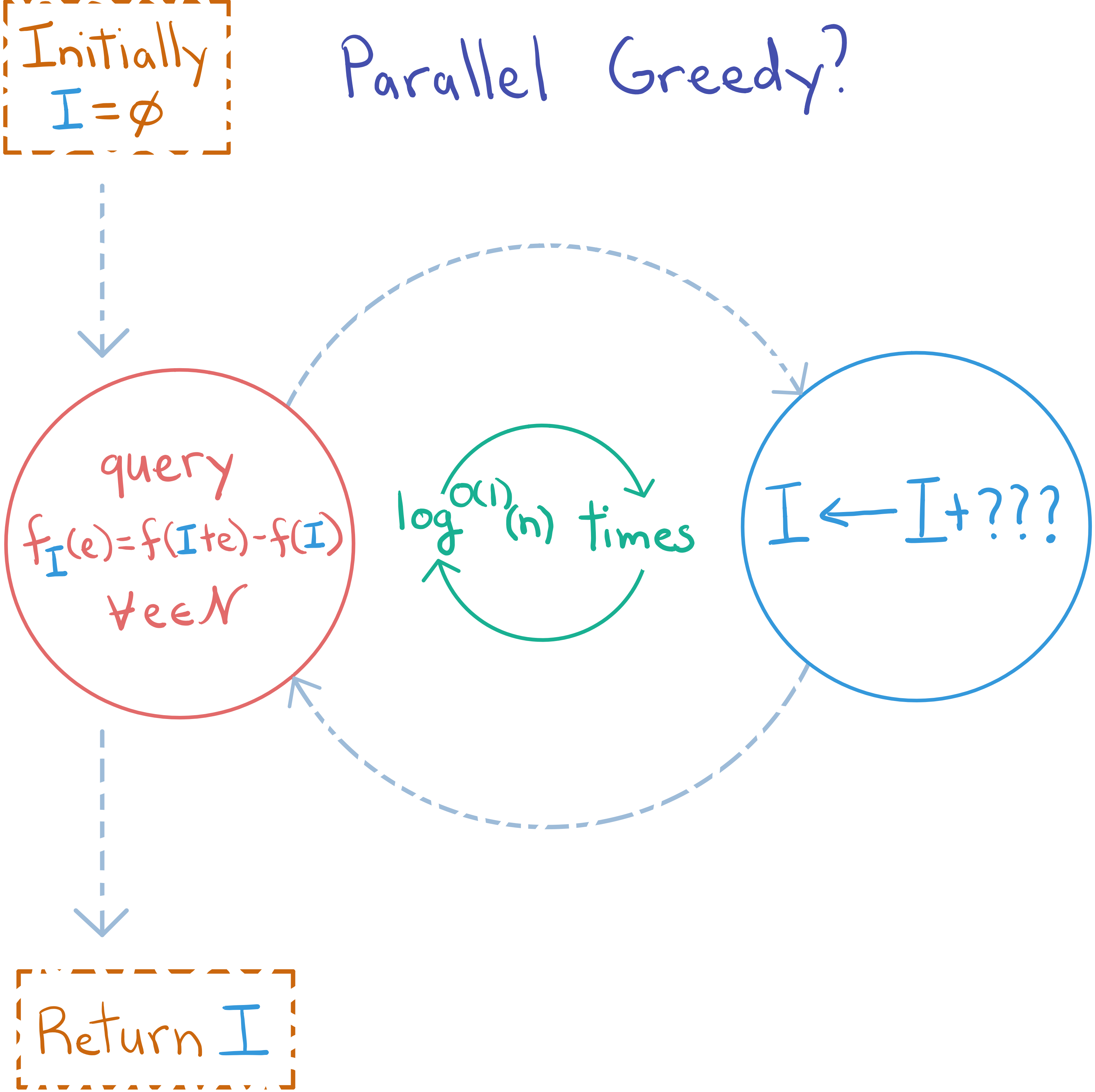


Greedy algorithm

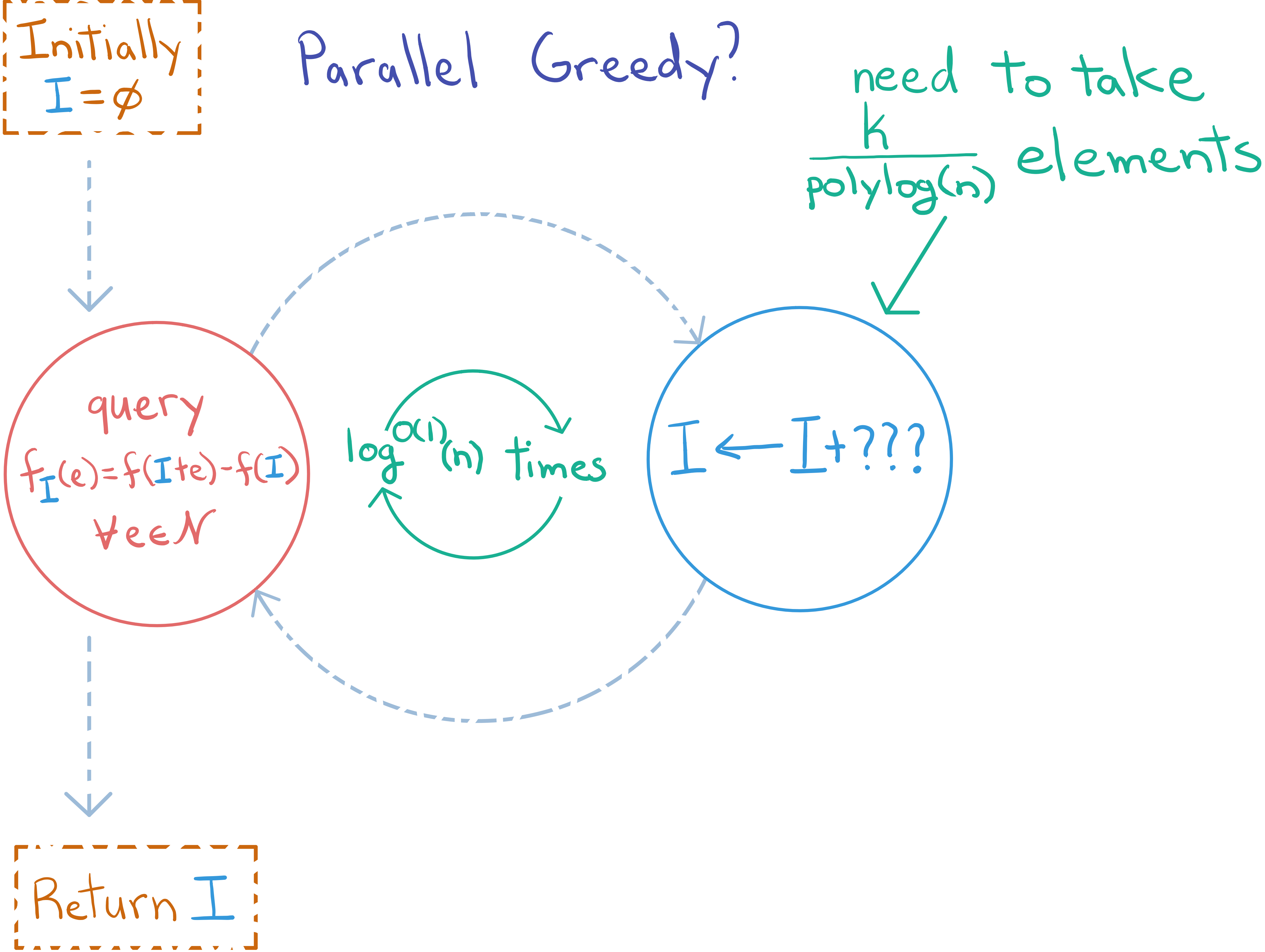


$O(k)$ adaptivity

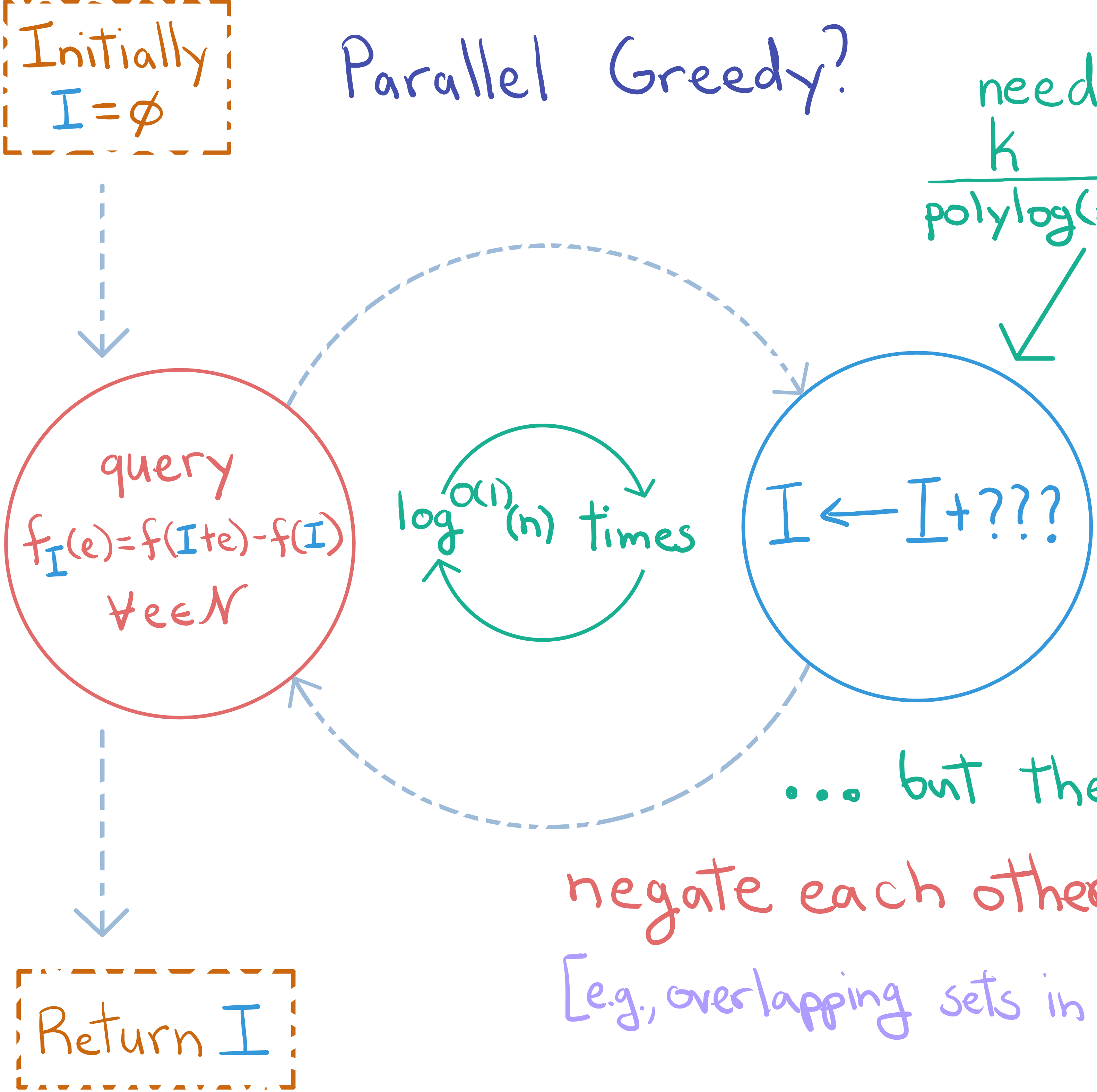
Parallel Greedy?



Parallel Greedy?



Parallel Greedy?



need to take $\frac{k}{\text{polylog}(n)}$ elements

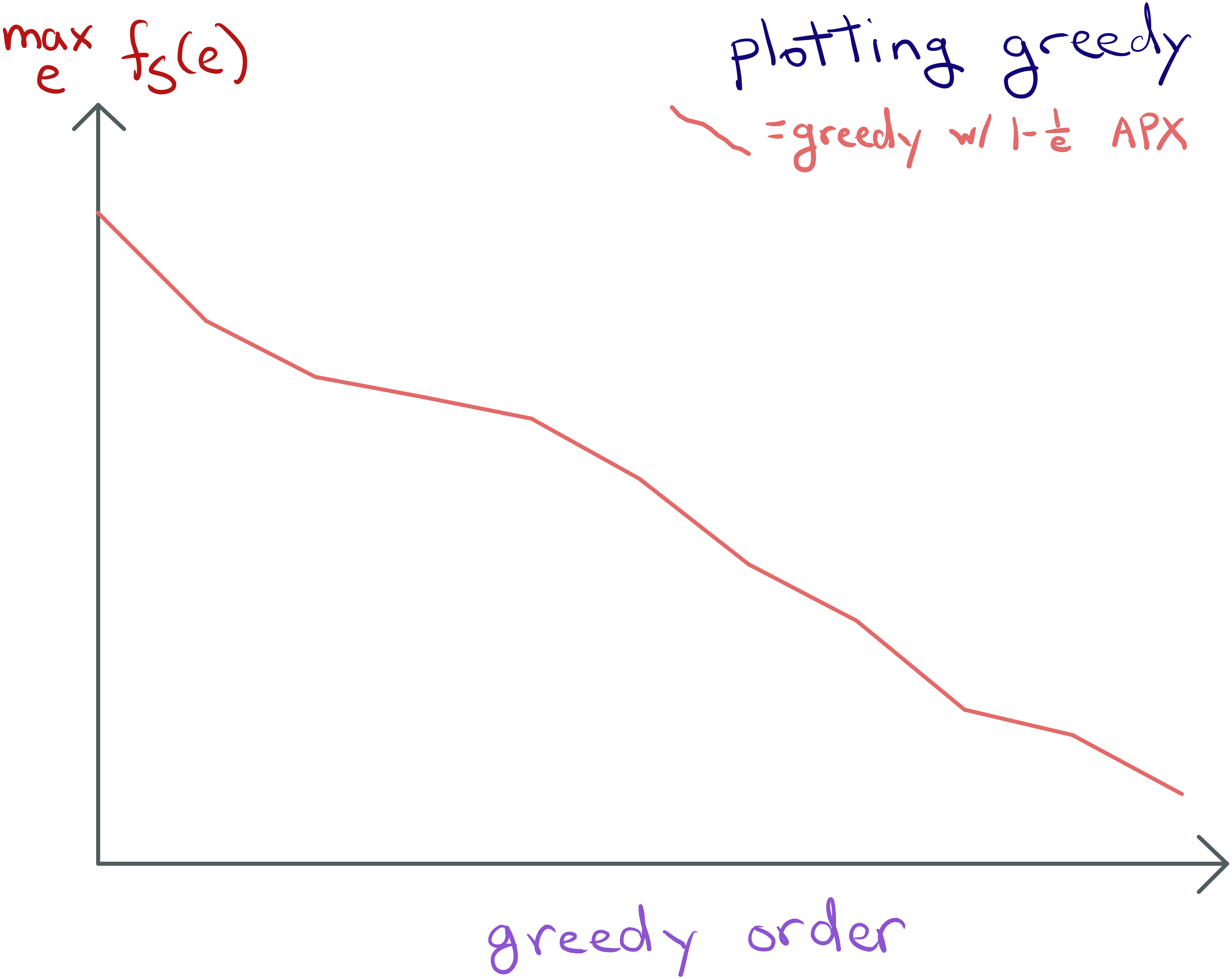
... but the elements:

negate each other w/r/t f

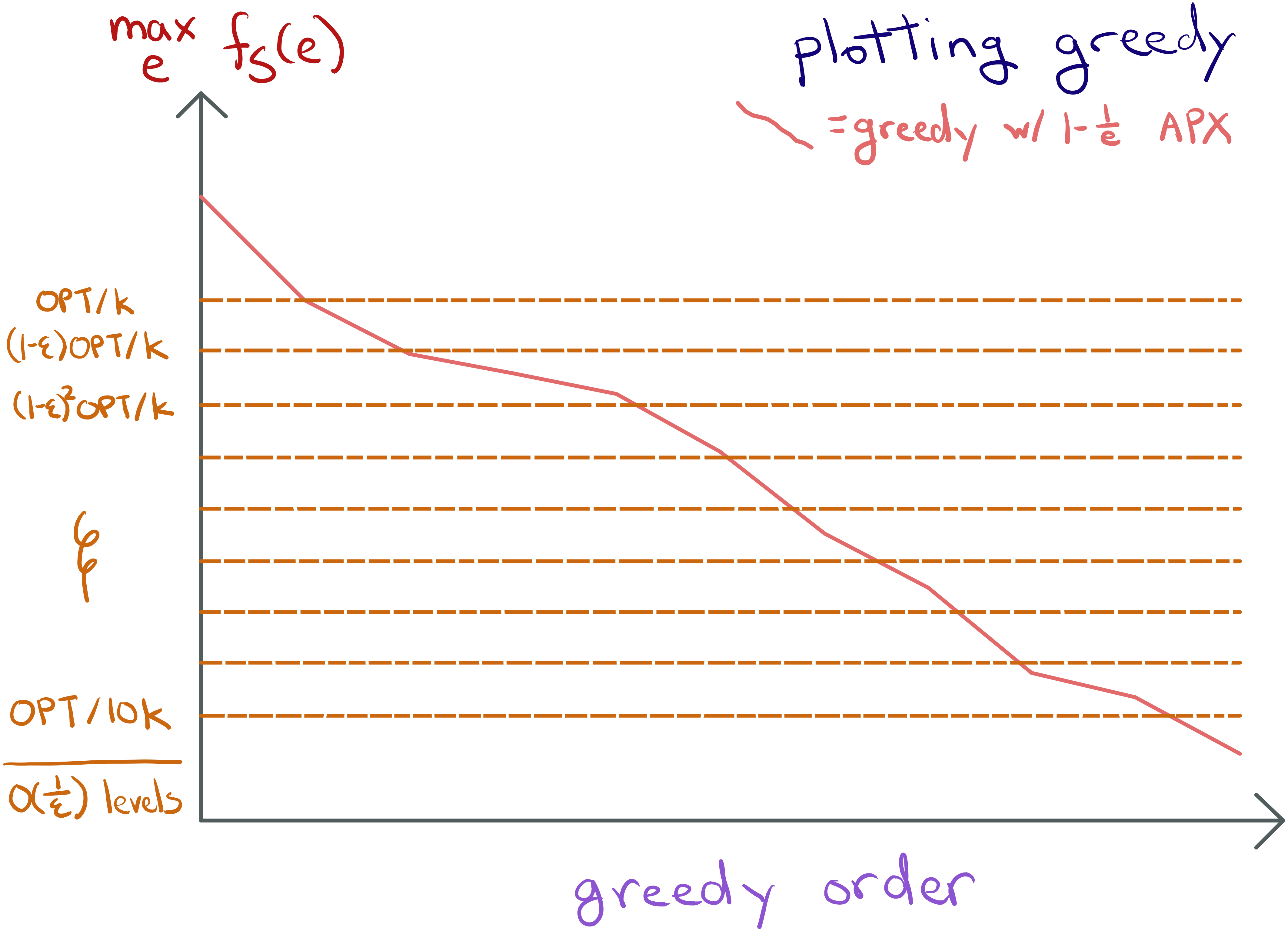
[e.g., overlapping sets in coverage]

Suppose we plot
the marginal values
of greedy

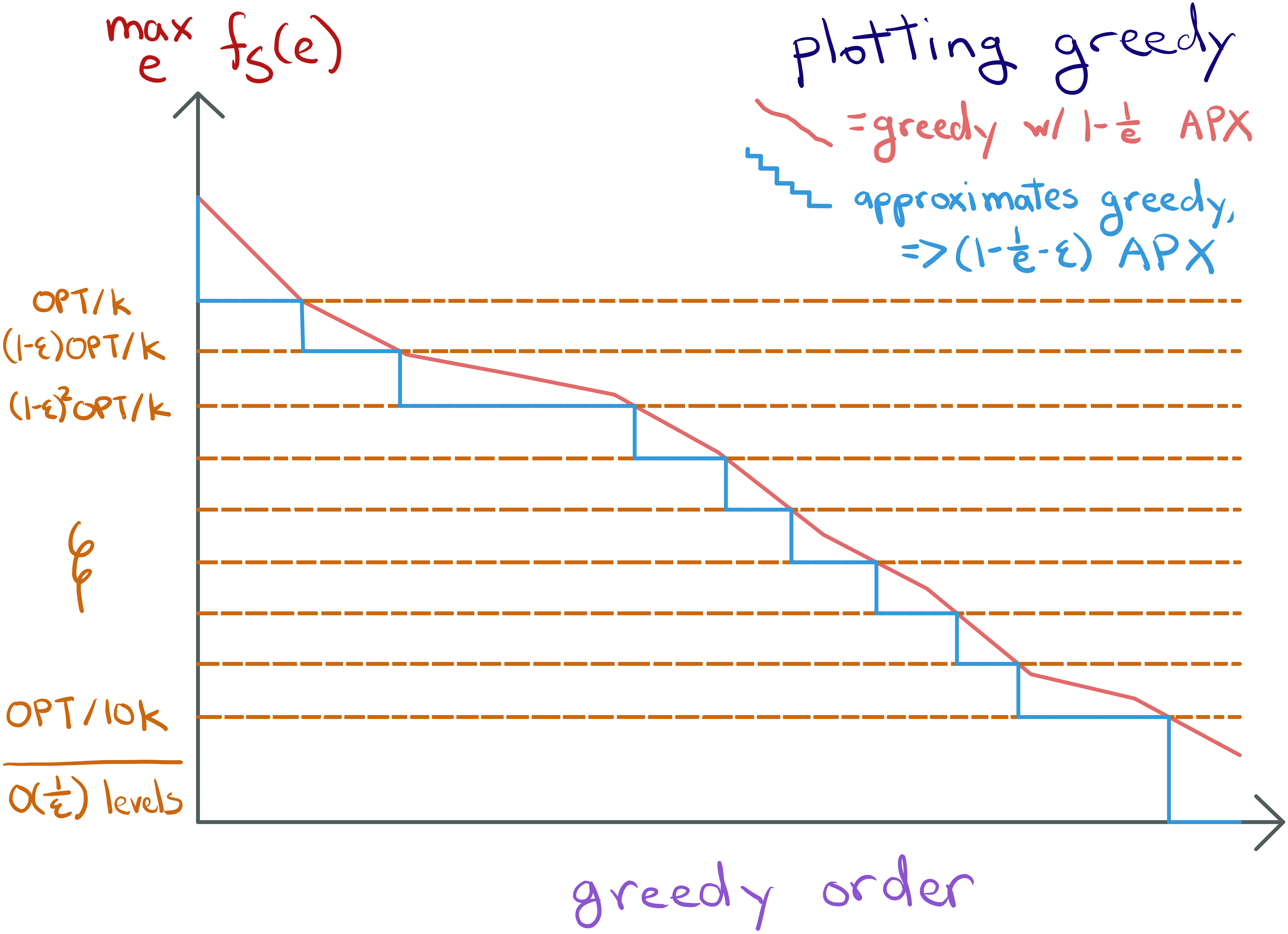
plotting greedy



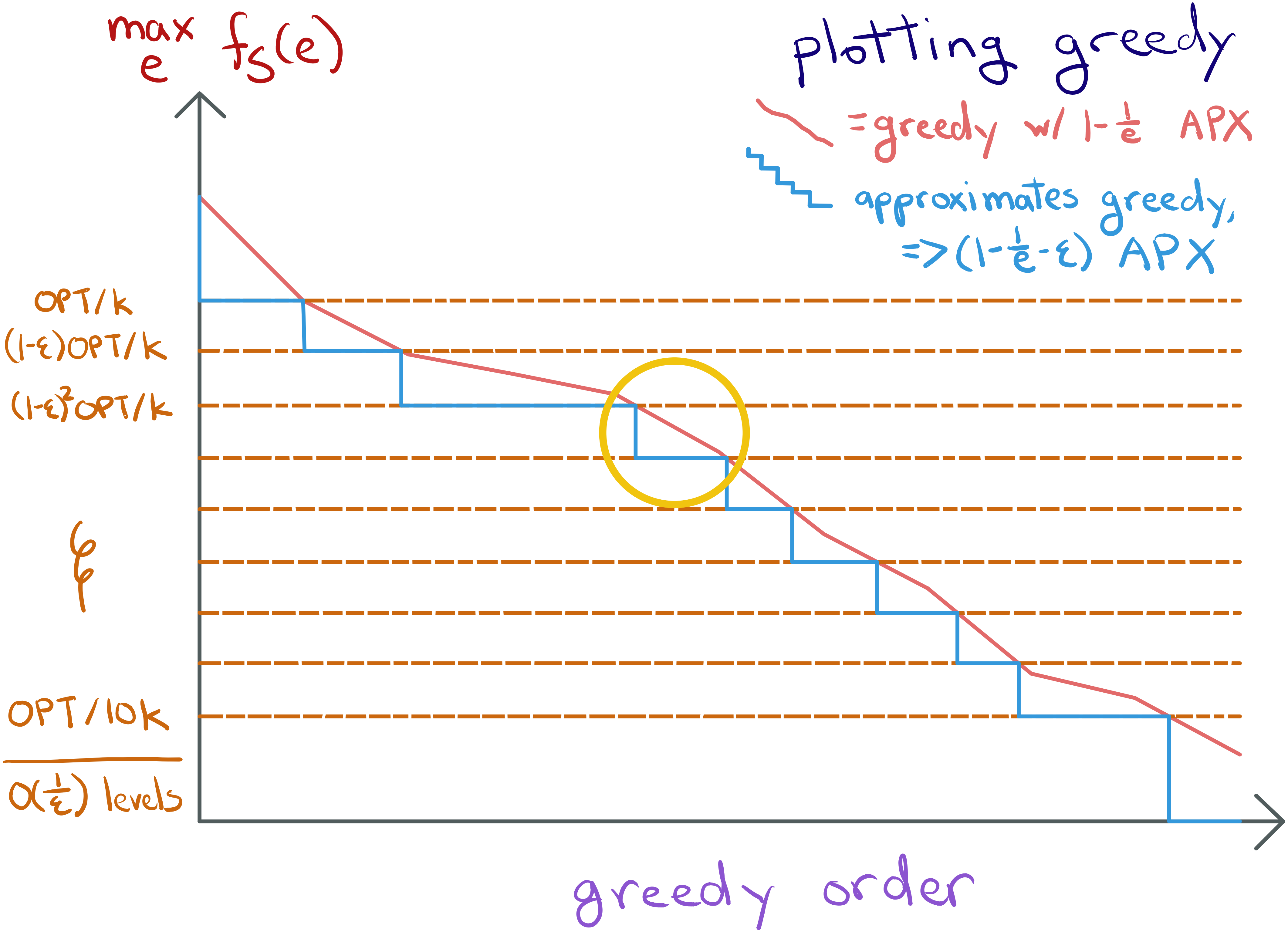
plotting greedy

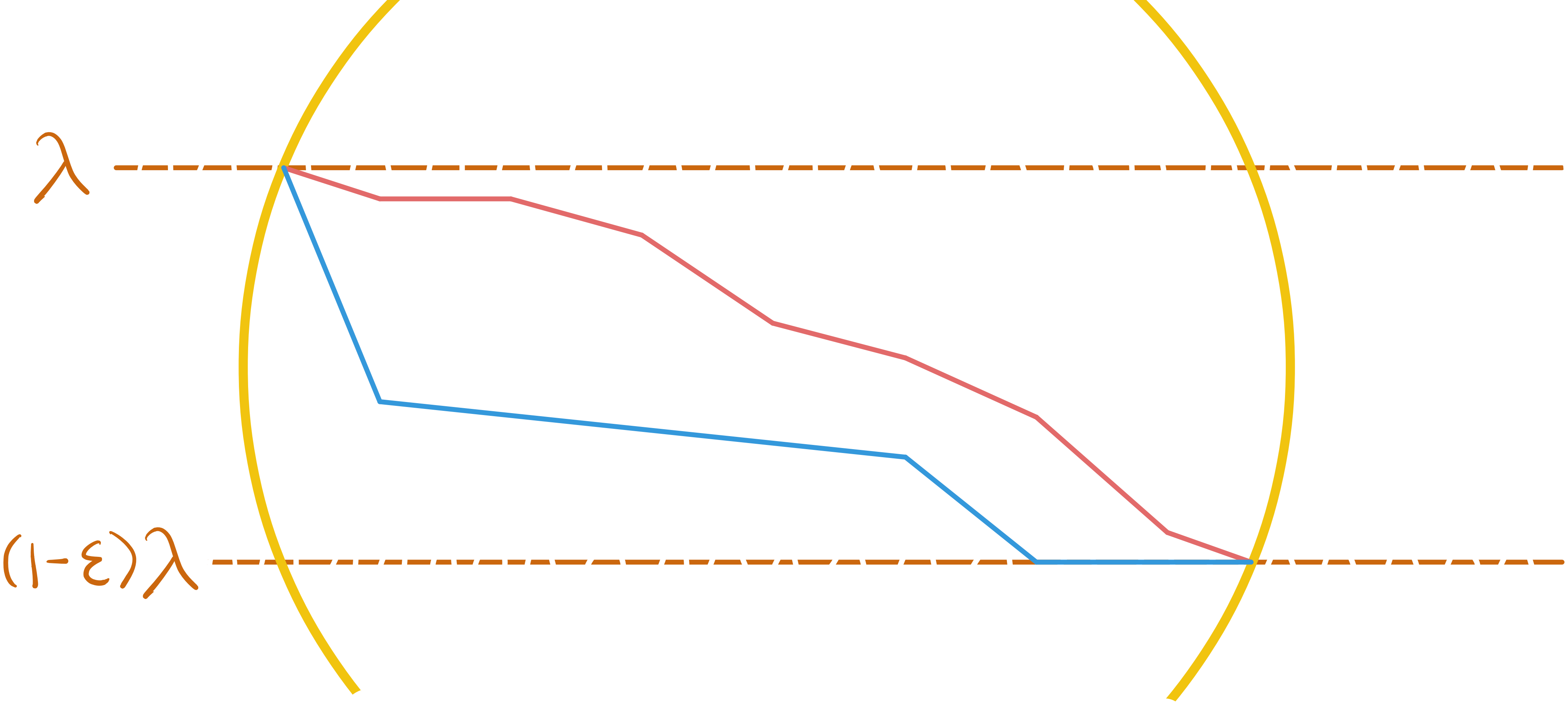


plotting greedy

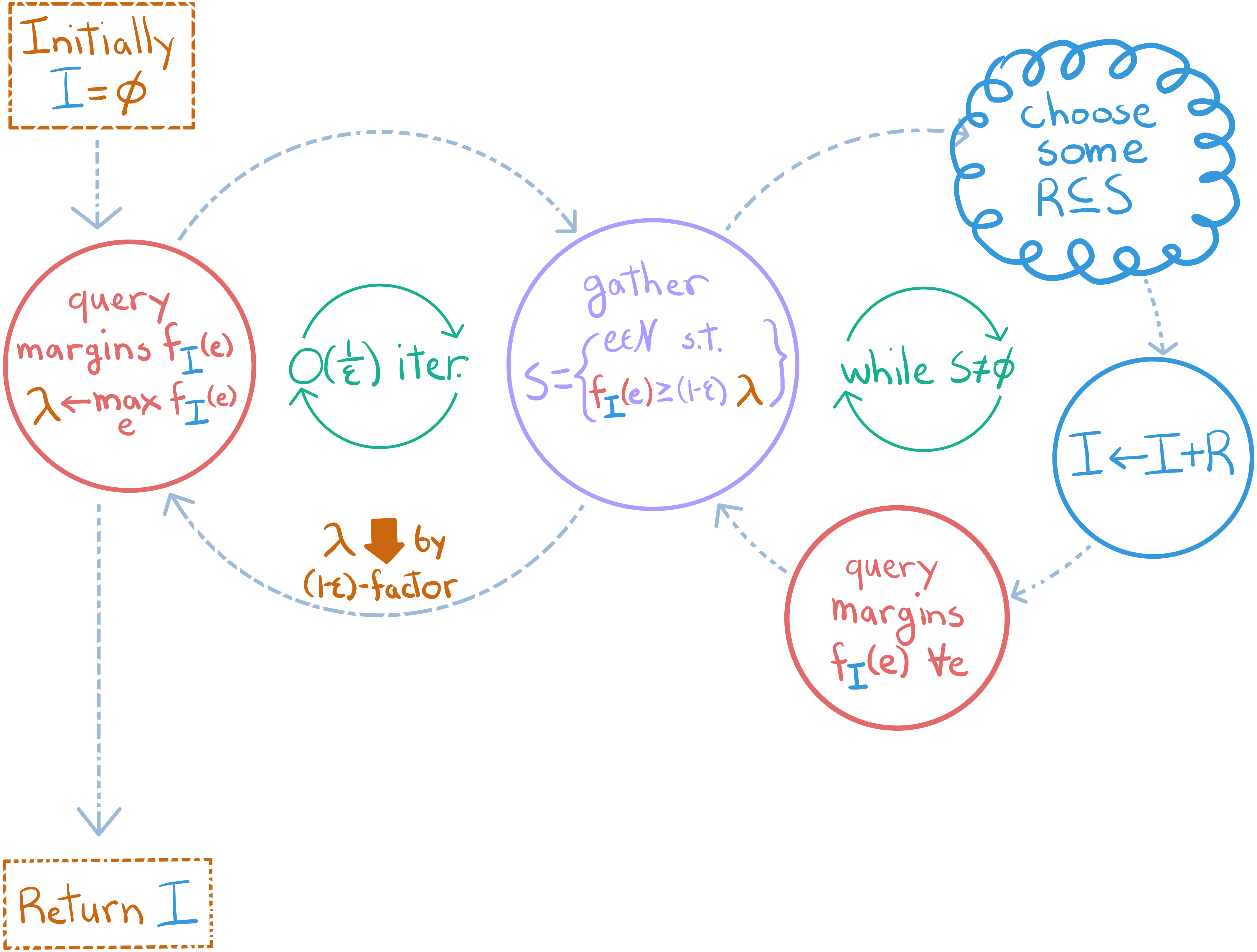


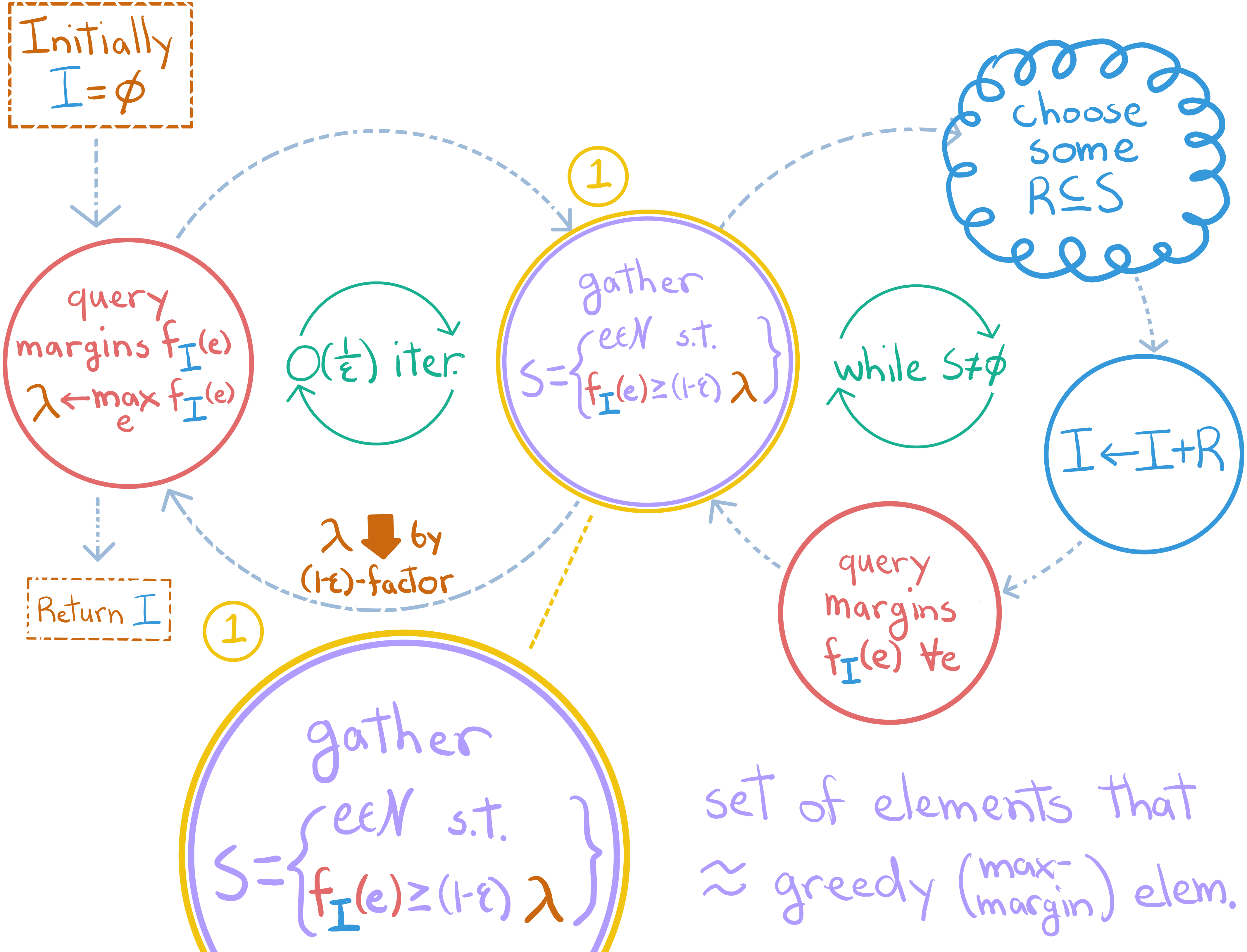
plotting greedy



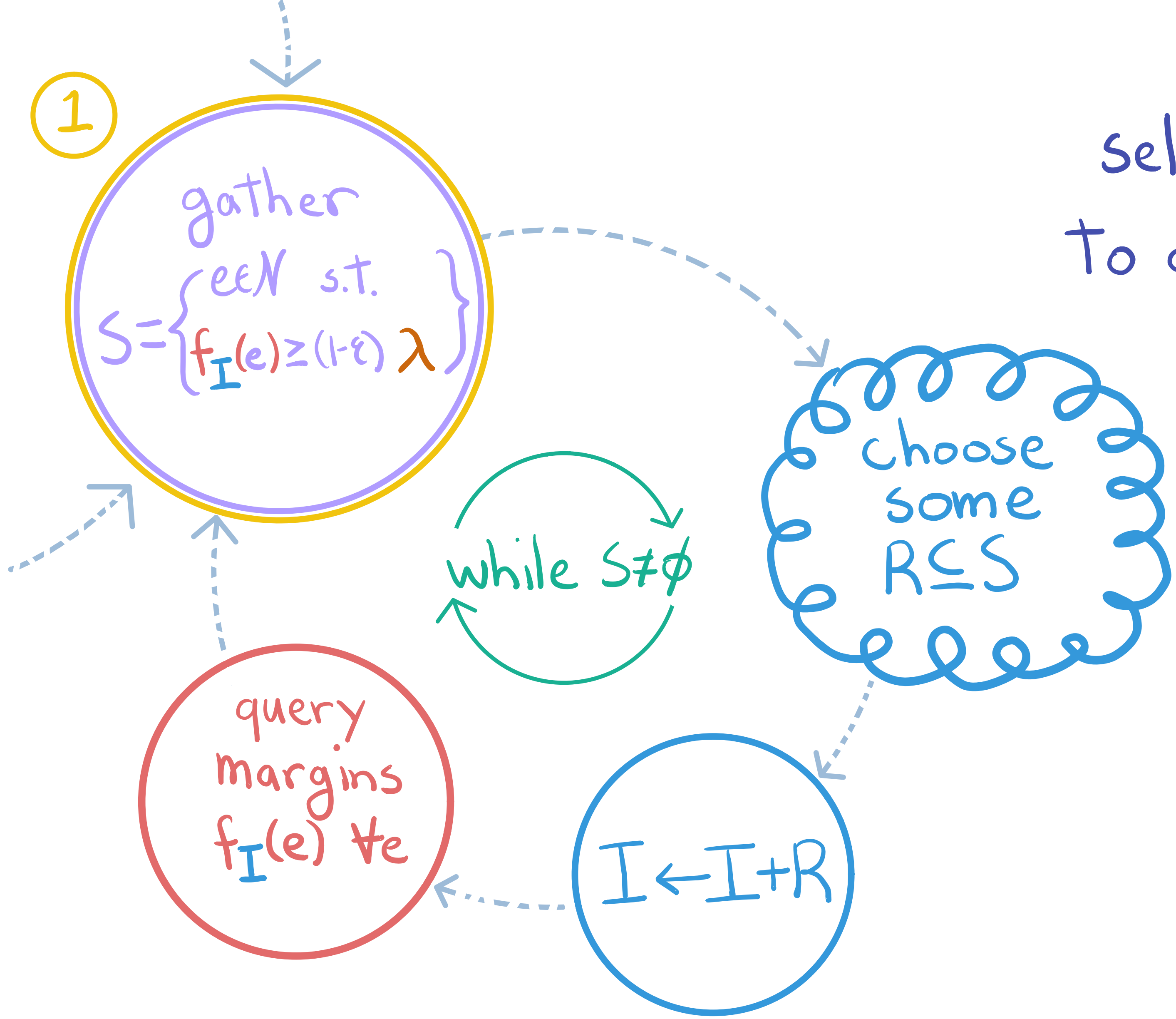


Goal: drive down max margin to $(1-\varepsilon)\lambda$ while taking elements w/ margin $\gtrsim \lambda$


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



set of elements that ≈ greedy (max-margin) elem.



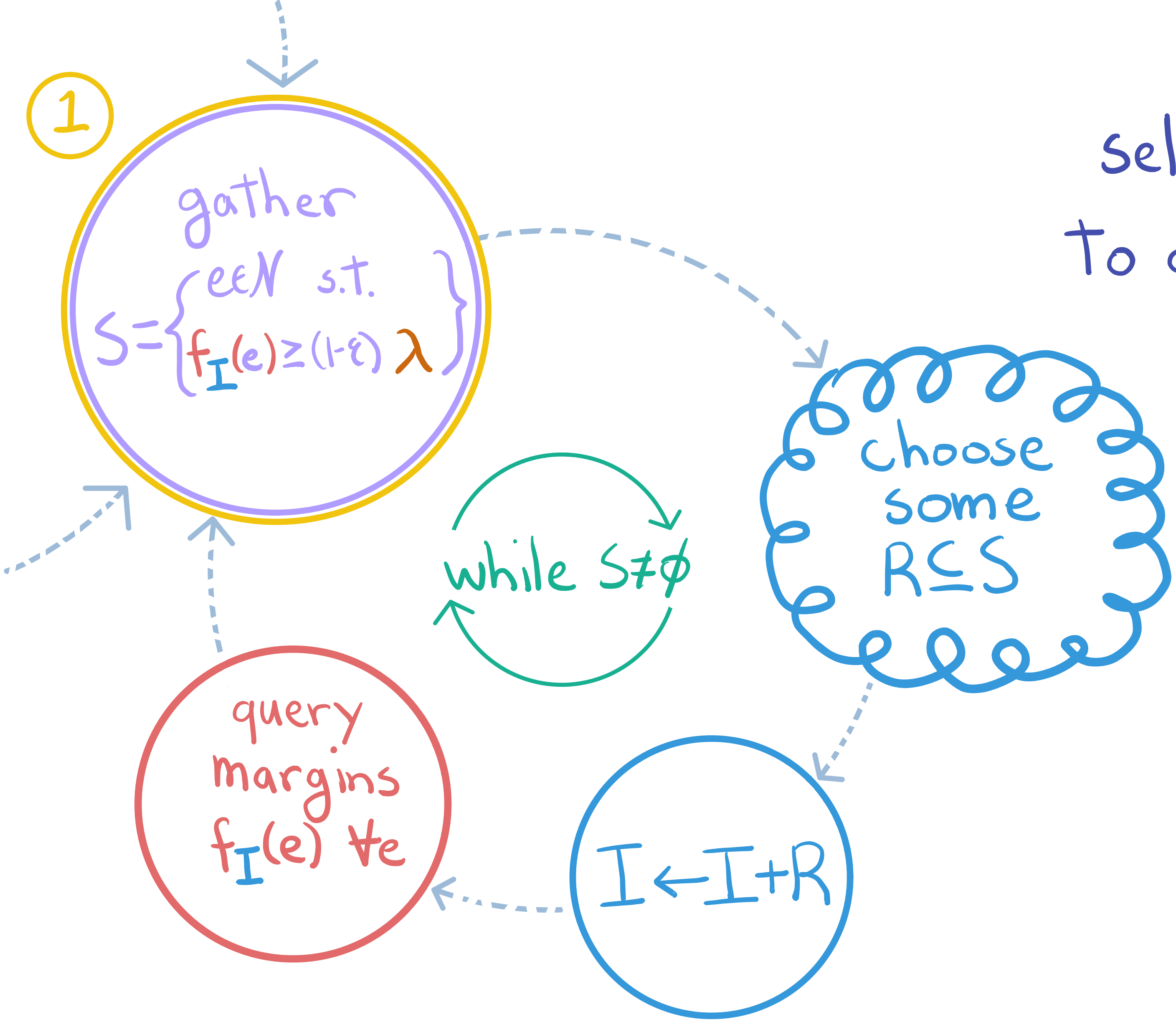
Goal

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

"good"

and

"large"



Goal

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

"good"

and

"large"

Now we're stuck again

let's take a *Continuous* point of view

need:

① continuous relaxation of f

② continuous version of greedy

"Multilinear extension"

want $F(x)$, $x \in [0,1]^n$

extending f

"Multilinear extension"

want $F(x)$, $x \in [0,1]^N$ extending f

interpret x as independent sample S

w/ $\Pr[e \in S] = x_e \quad \forall e$

"Multilinear extension"

$$\boxed{F(x) = E[f(S) \mid S \sim x]}$$

interpret x as independent sample S

w/ $\Pr[e \in S] = x_e \quad \forall e$

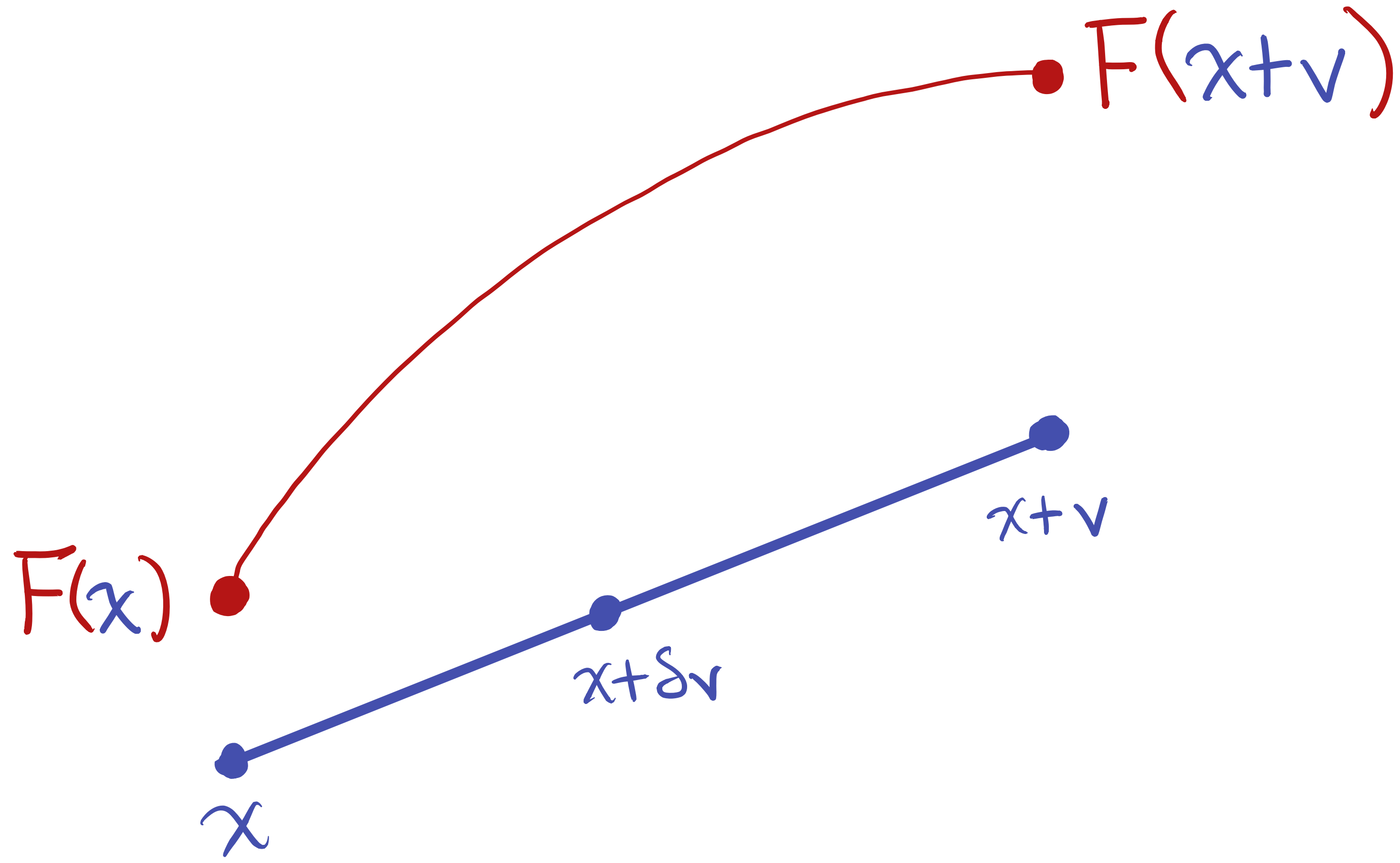
the derivative $F'(x)$

for element e,

$$F_e'(x) = \frac{\partial}{\partial x_e} F(x) = \underbrace{\lim_{\delta \to 0} \frac{F(x+\delta 1_e) - F(x)}{\delta}}_{\text{continuous analogue of marginal returns}}$$

- monotone-concave: for $x \in [0,1]^N$, $v \in \mathbb{R}^N_{\geq 0}$, $\delta > 0$

$F(x+\delta v)$ is concave in δ



new goal

maximize $F(x)$ over $x \geq 0$ s.t. $\sum_i x_i \leq k$

in parallel in the oracle model

(rounding is easy)

Initially $x = 0^N$

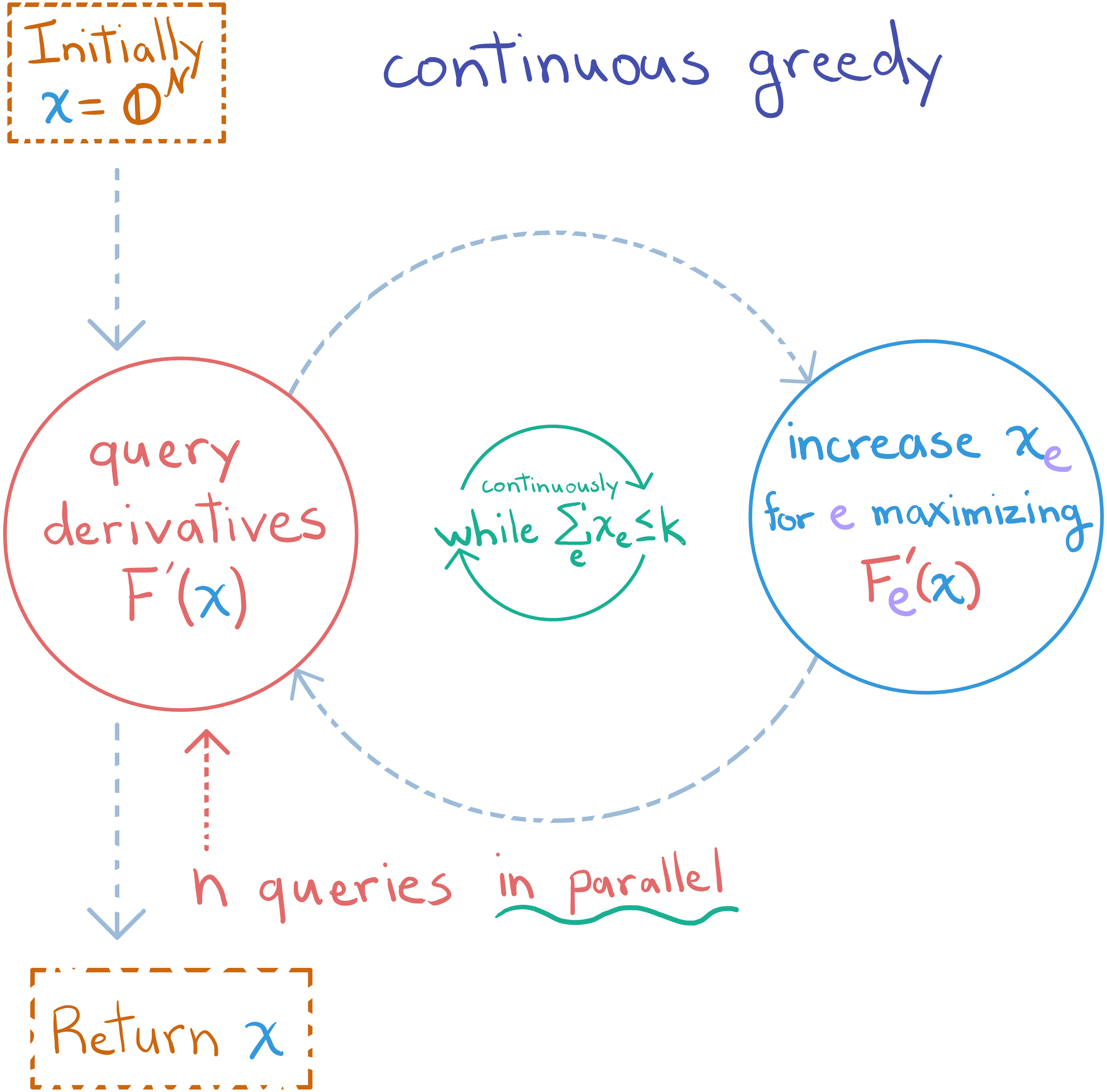
Continuous Greedy

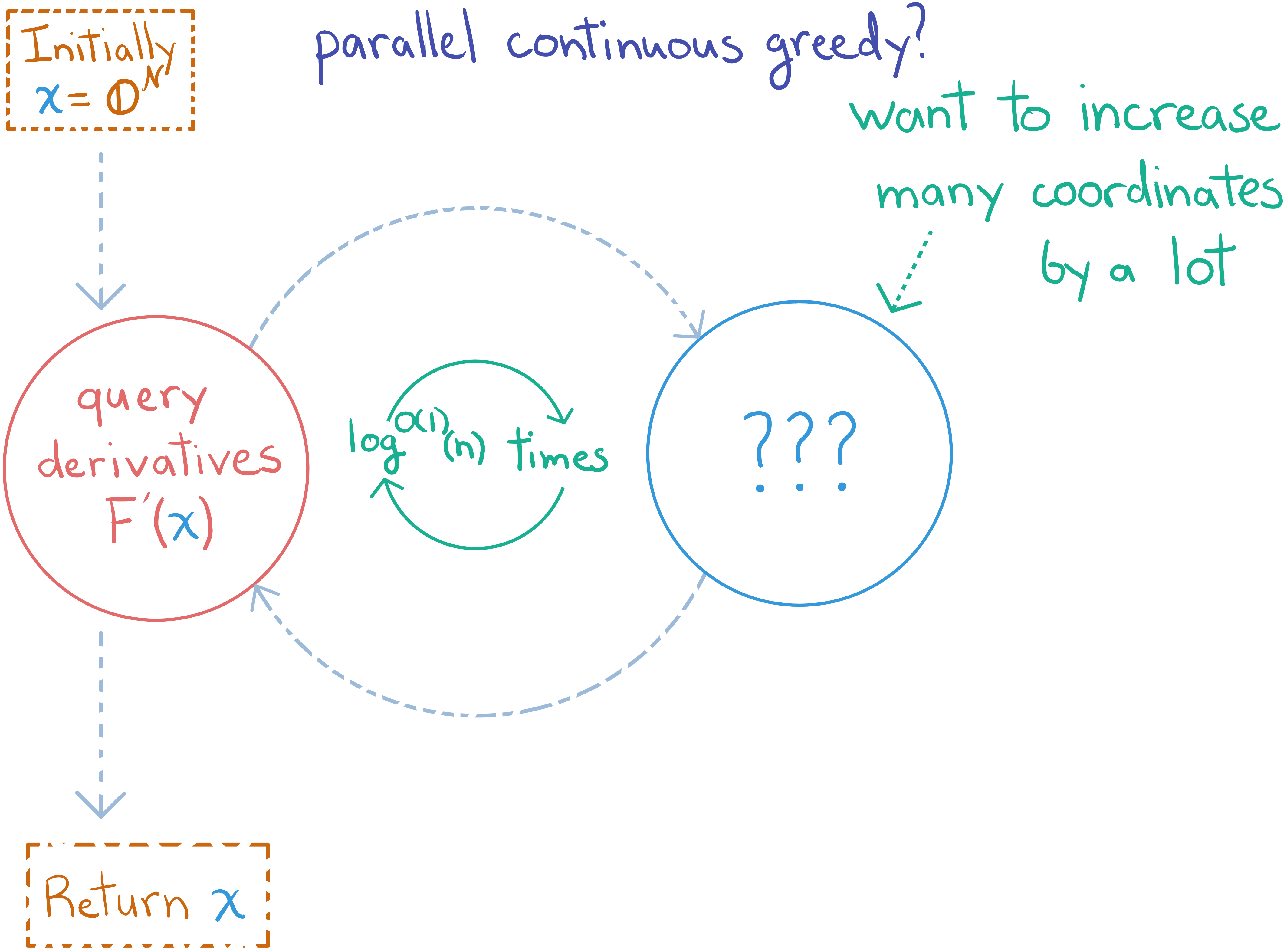
[Wolsey] [Vondrák]

[Calinescu, Chekuri, Pál, Vondrák]

Return x

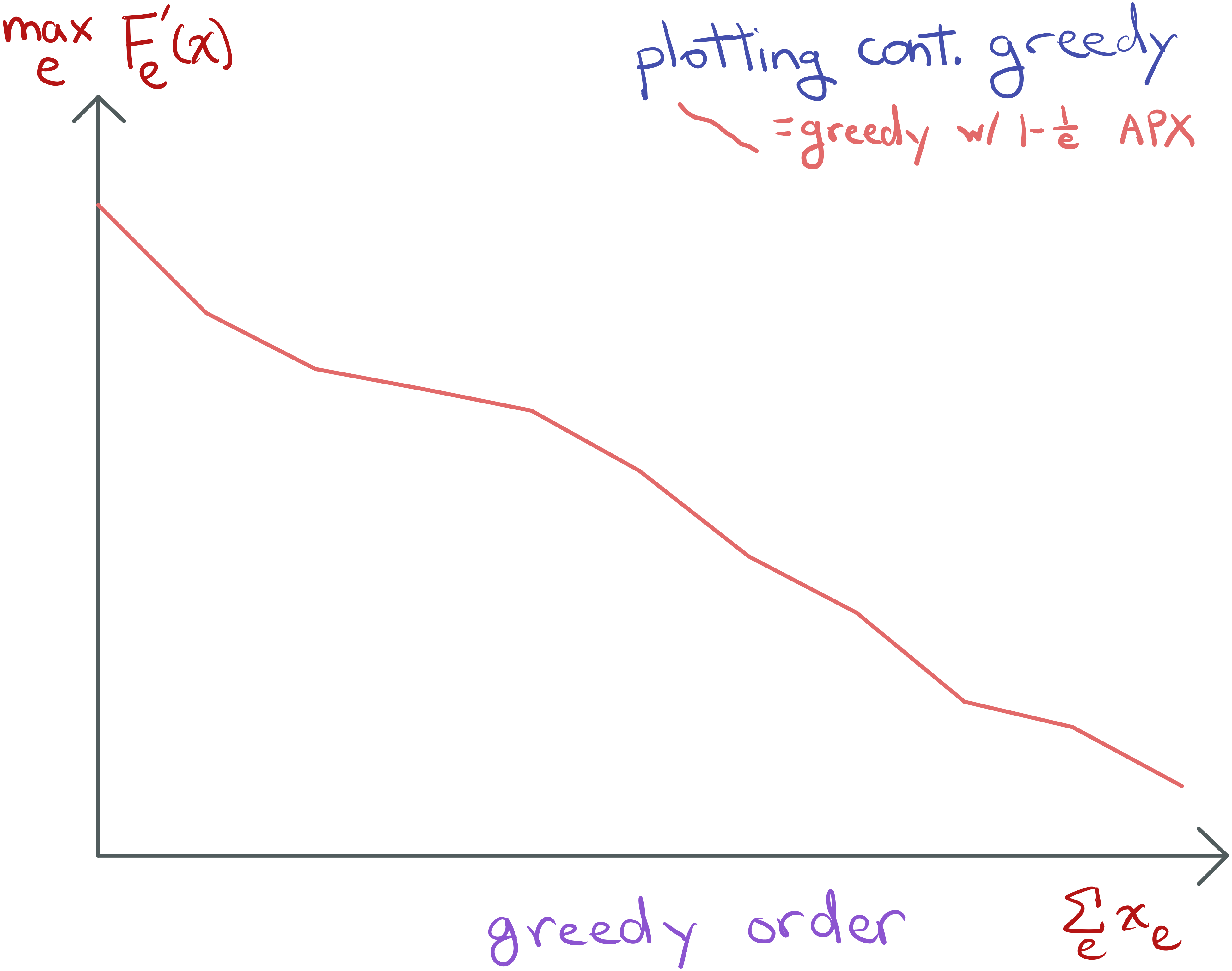
continuous greedy

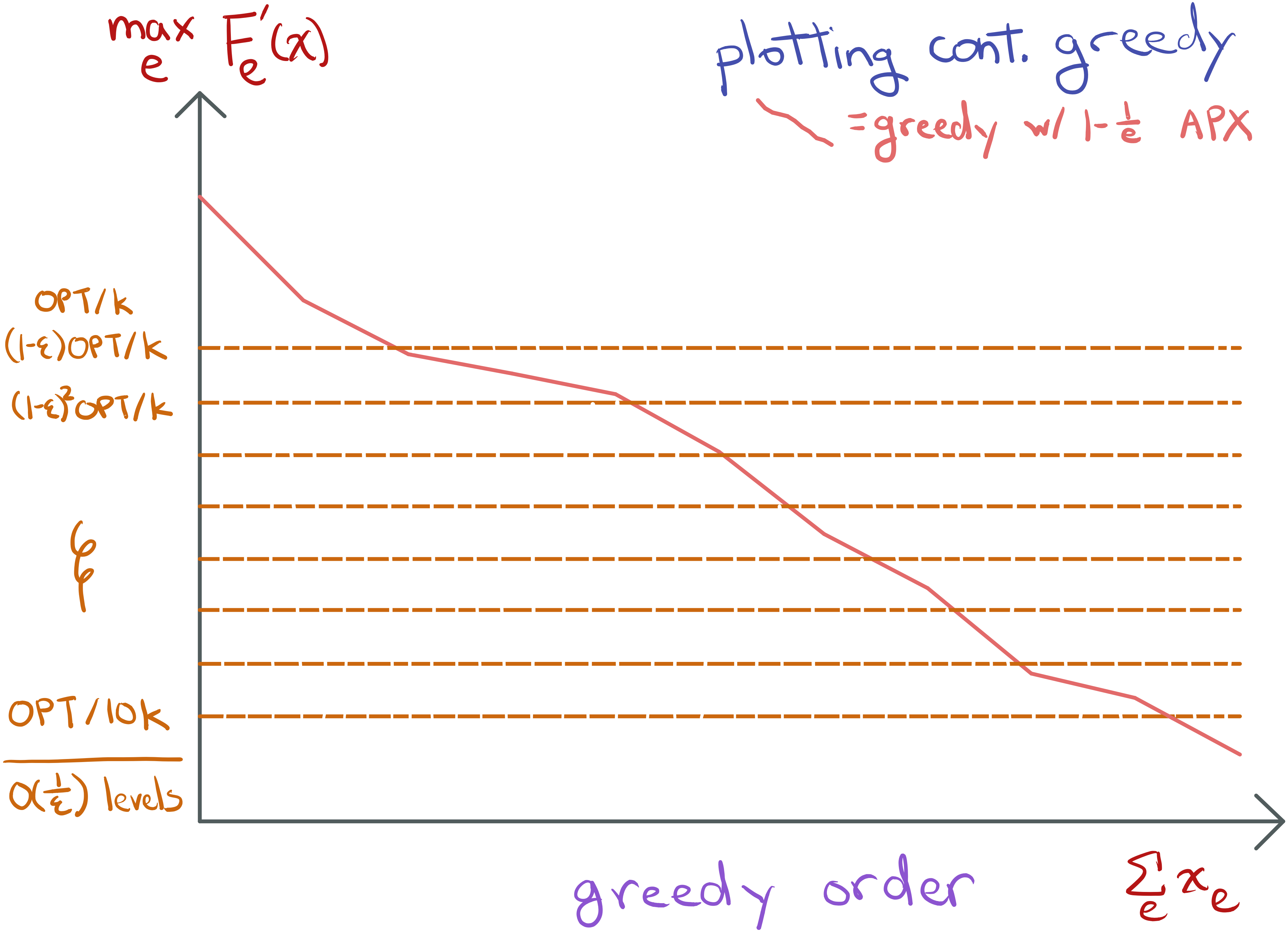


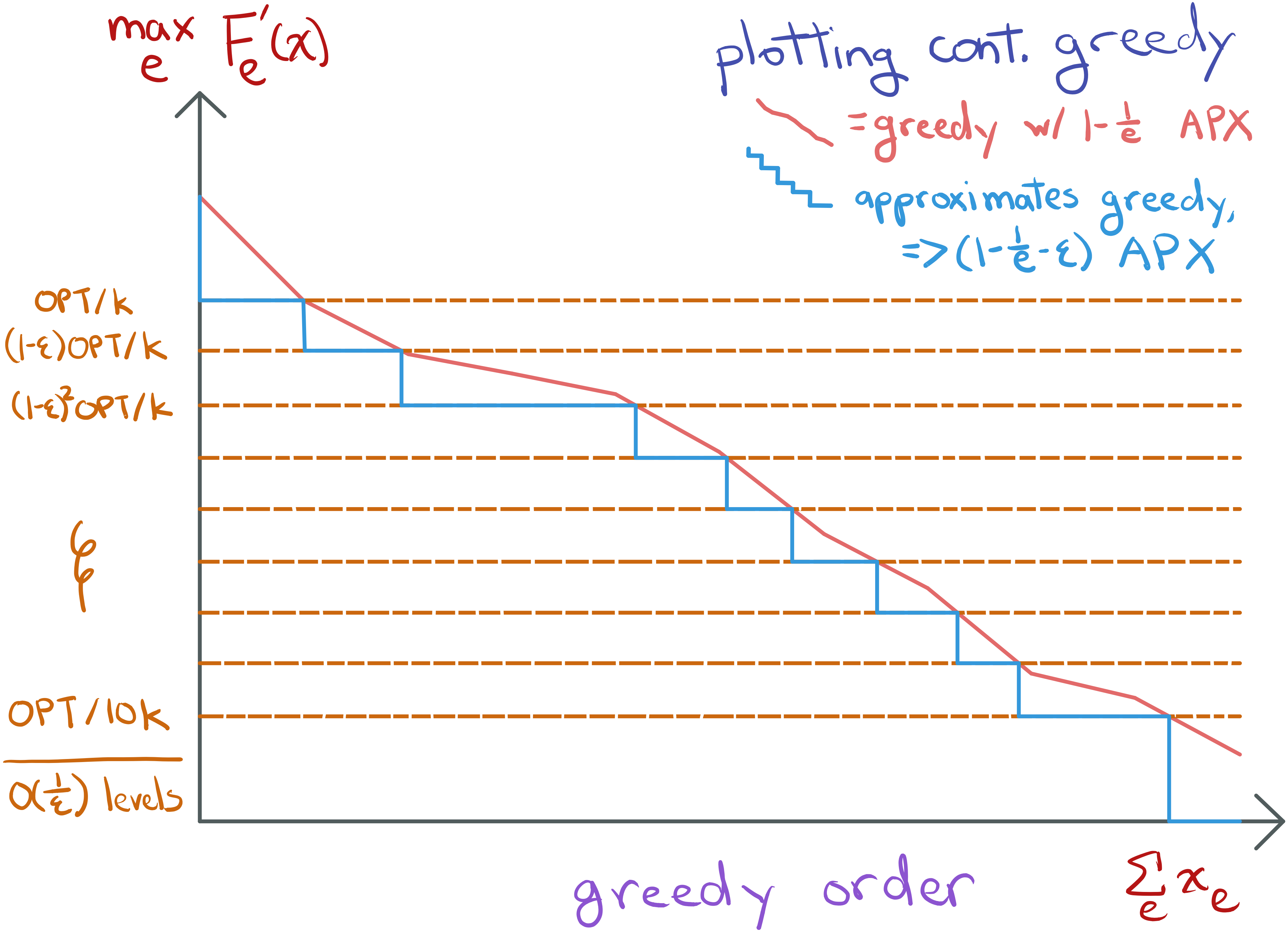

parallel continuous greedy?
Initially $x = 0^N$
want to increase many coordinates by a lot
query derivatives $F'(x)$
$\log^{O(1)}(n)$ times
???
Return x

Suppose we plot
the derivatives of
continuous greedy

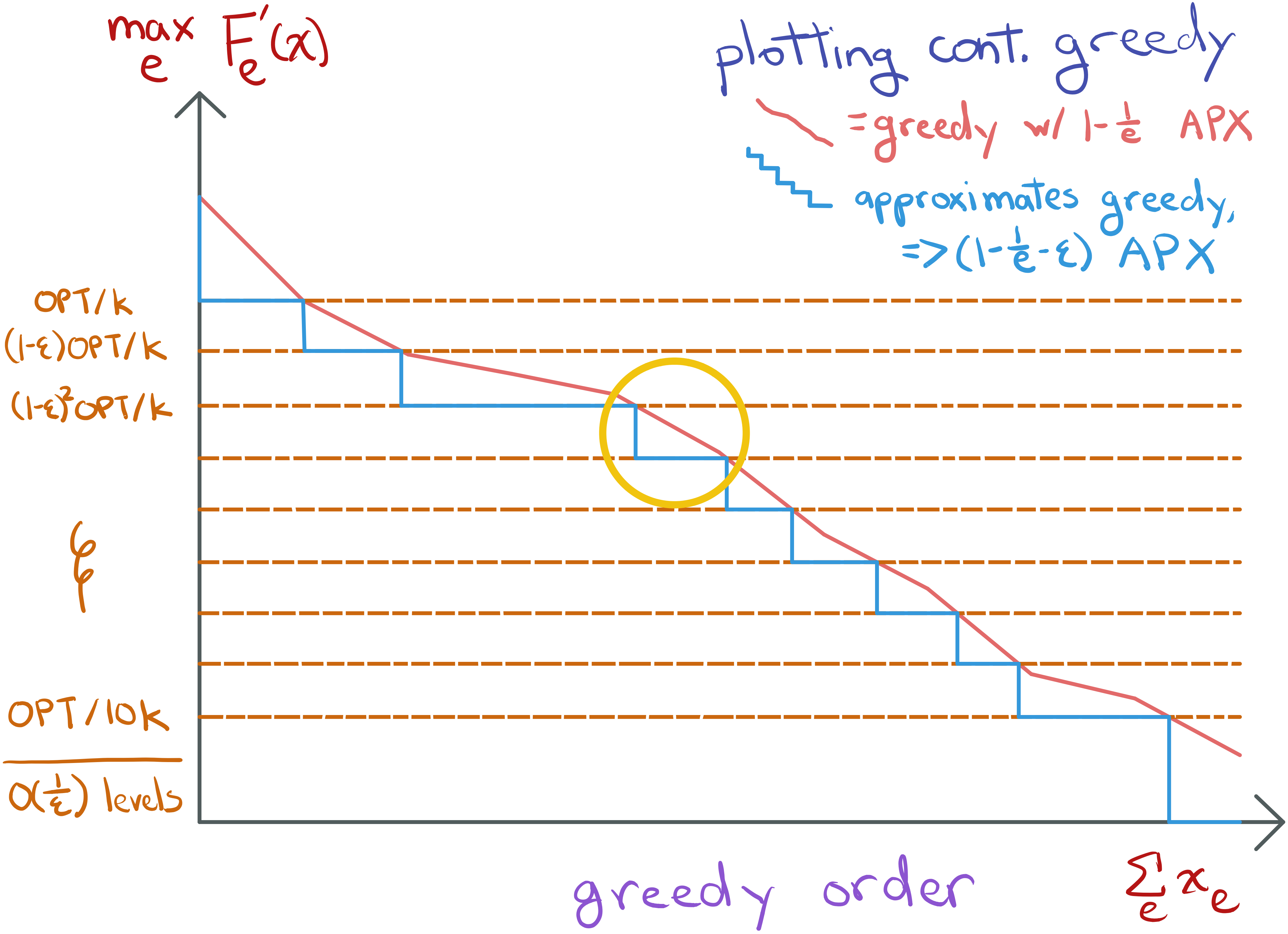
plotting cont. greedy

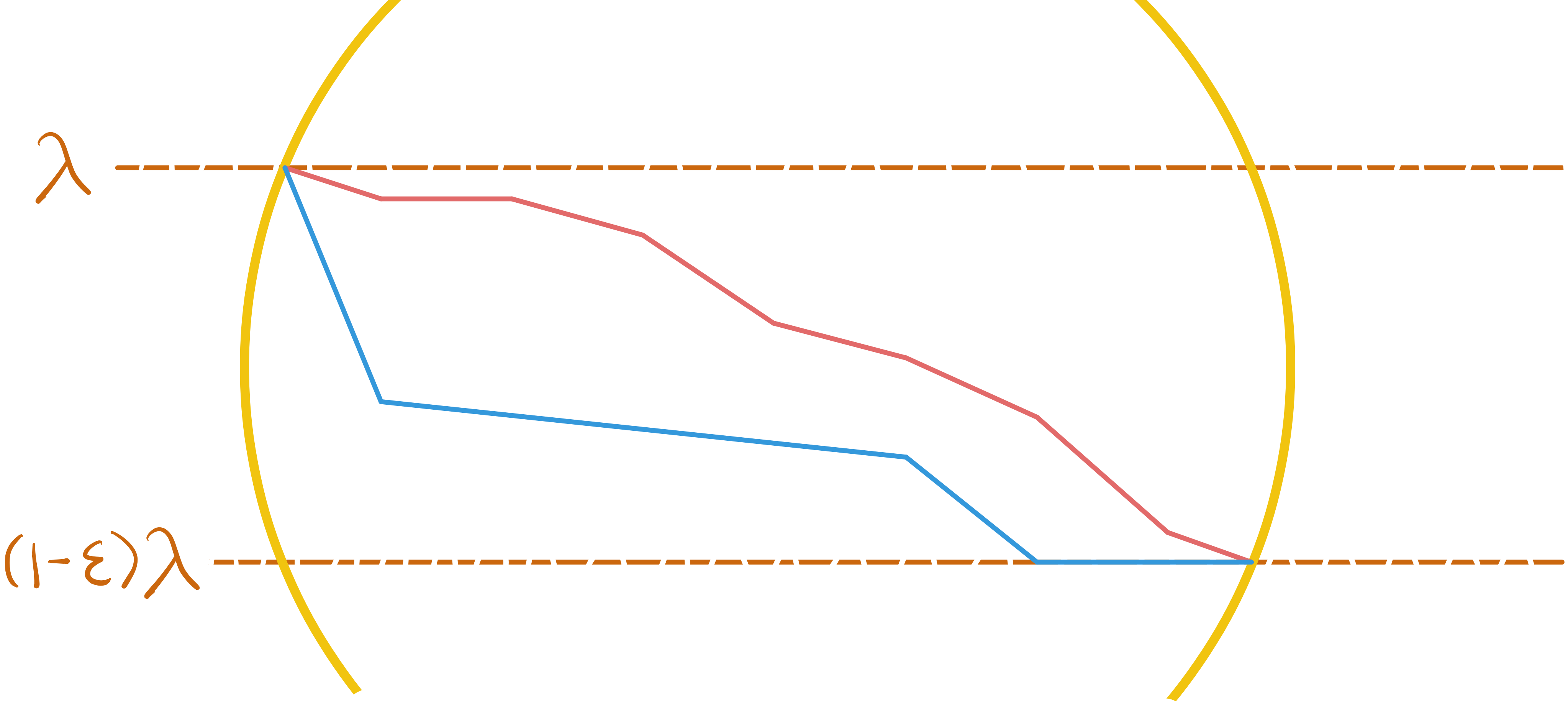



$\max_e F'_e(x)$
plotting cont. greedy
$= $ greedy w/ $1-\frac{1}{e}$ APX
OPT/k
$(1-\varepsilon)$OPT/k
$(1-\varepsilon)^2$OPT/k
φ
OPT/10k
$O(\frac{1}{\varepsilon})$ levels
greedy order
$\sum_e x_e$

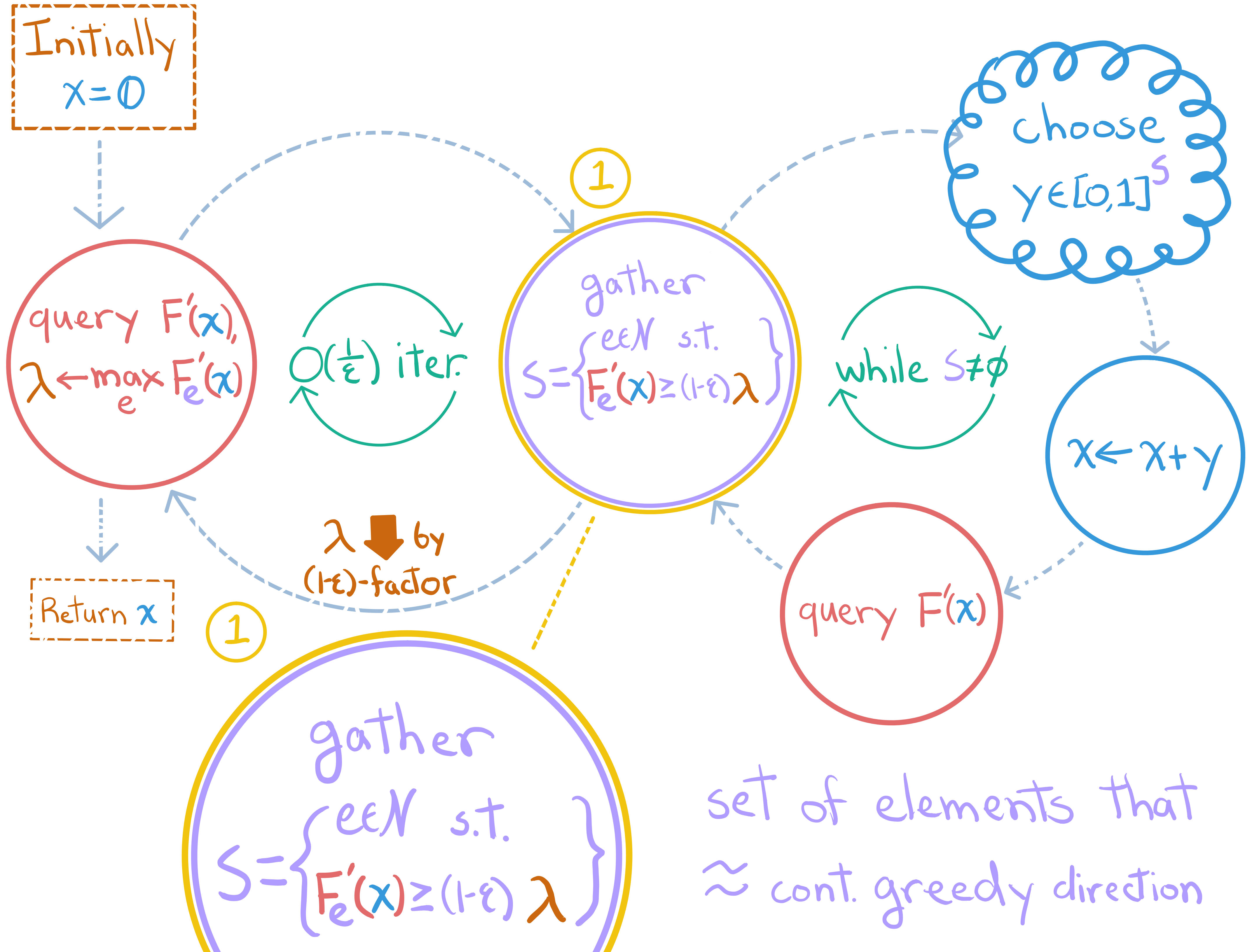

$\max_e F'_e(x)$
plotting cont. greedy
= greedy w/ $1-\frac{1}{e}$ APX
approximates greedy,
=> $(1-\frac{1}{e}-\varepsilon)$ APX
OPT/k
$(1-\varepsilon)$OPT/k
$(1-\varepsilon)^2$OPT/k
φ
OPT/10k
$O(\frac{1}{\varepsilon})$ levels
greedy order
$\sum_e x_e$

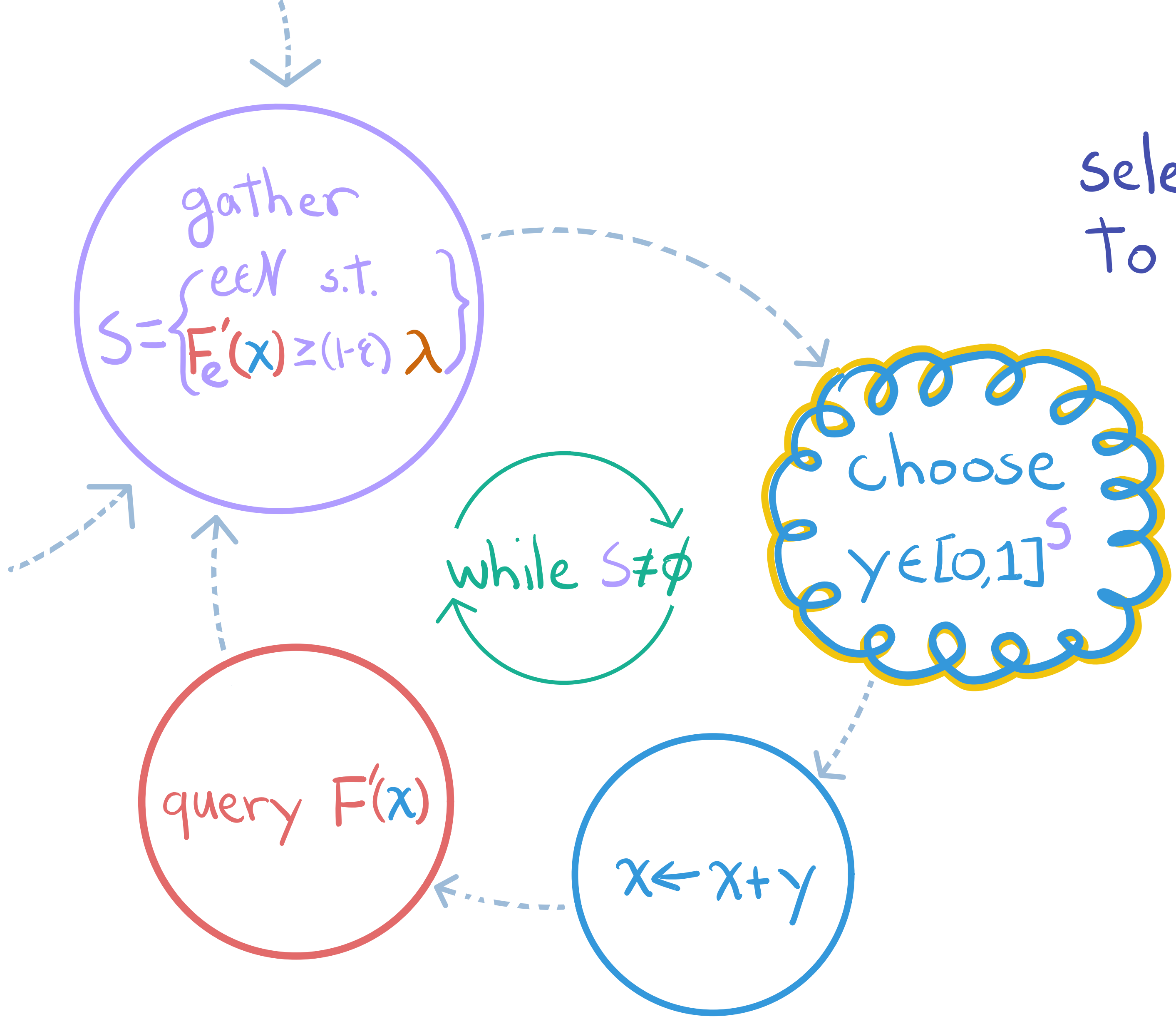
plotting cont. greedy





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


Initially
$x = 0$
query $F'(x)$,
$\lambda \leftarrow \max_e F'_e(x)$
$O(\frac{1}{\varepsilon})$ iter.
①
gather
$S = \{e \in N$ s.t. $F'_e(x) \geq (1-\varepsilon)\lambda\}$
while $S \neq \phi$
choose
$y \in [0,1]^S$
$x \leftarrow x + y$
query $F'(x)$
λ ⬇ by
$(1-\varepsilon)$-factor
Return x
①
gather
$S = \{e \in N$ s.t. $F'_e(x) \geq (1-\varepsilon)\lambda\}$
set of elements that
$\approx$ cont. greedy direction



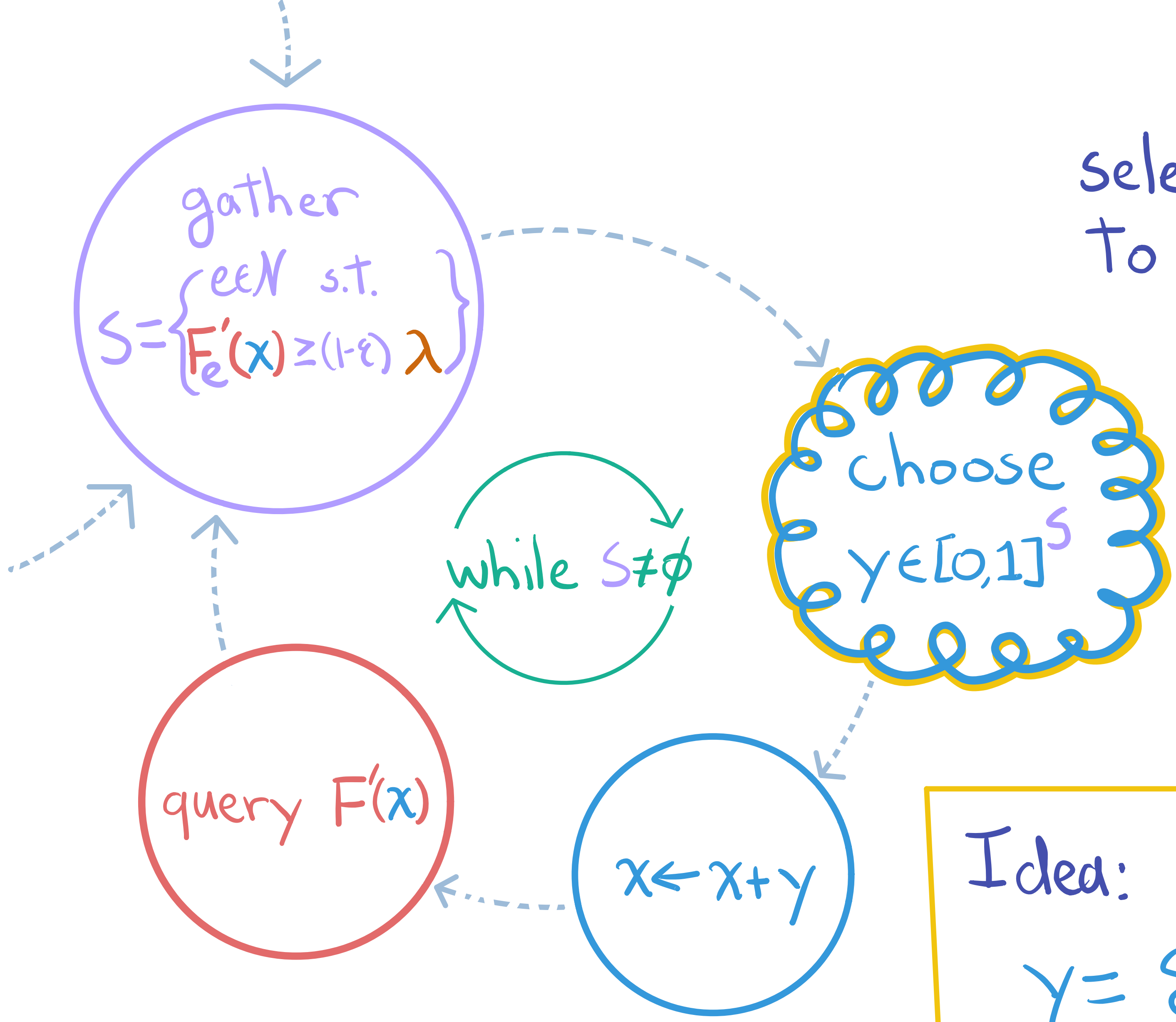
Goal

select a vector $y \in [0,1]^S$ to add to x that is both

"good"

and

"large"



Goal

select a vector $y \in [0,1]^S$ to add to x that is both

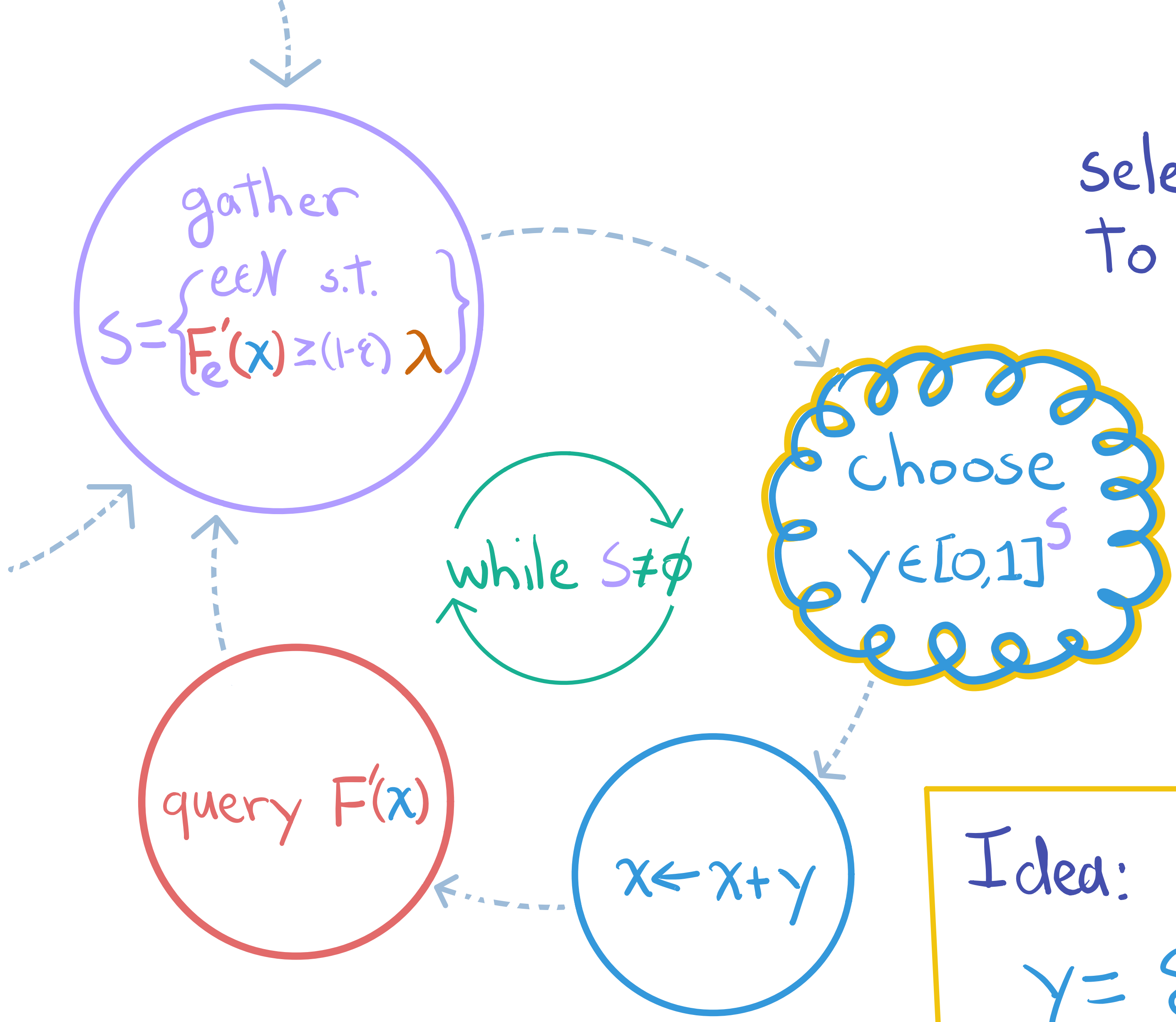
"good"

and

"large"

Idea: choose

$y = \delta \mathbb{1}_S$ for some $\delta > 0$



Goal

select a vector $y\in[0,1]^S$ to add to x that is both "good" and "large"

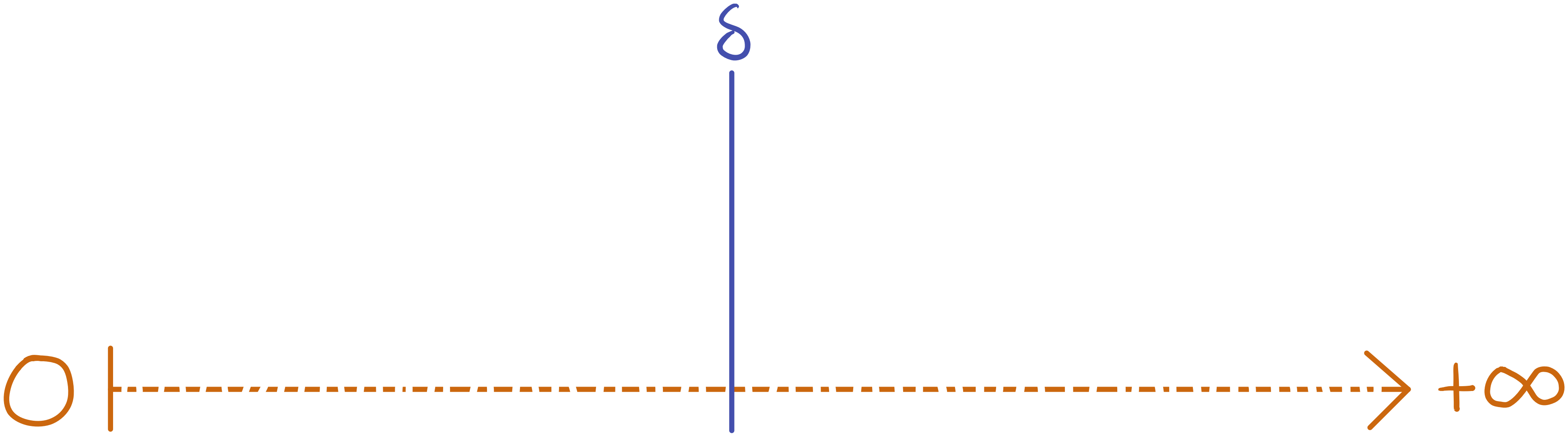
Idea: choose $y=\delta\mathbb{1}_S$ for some $\delta>0$

how to choose δ?

choosing δ

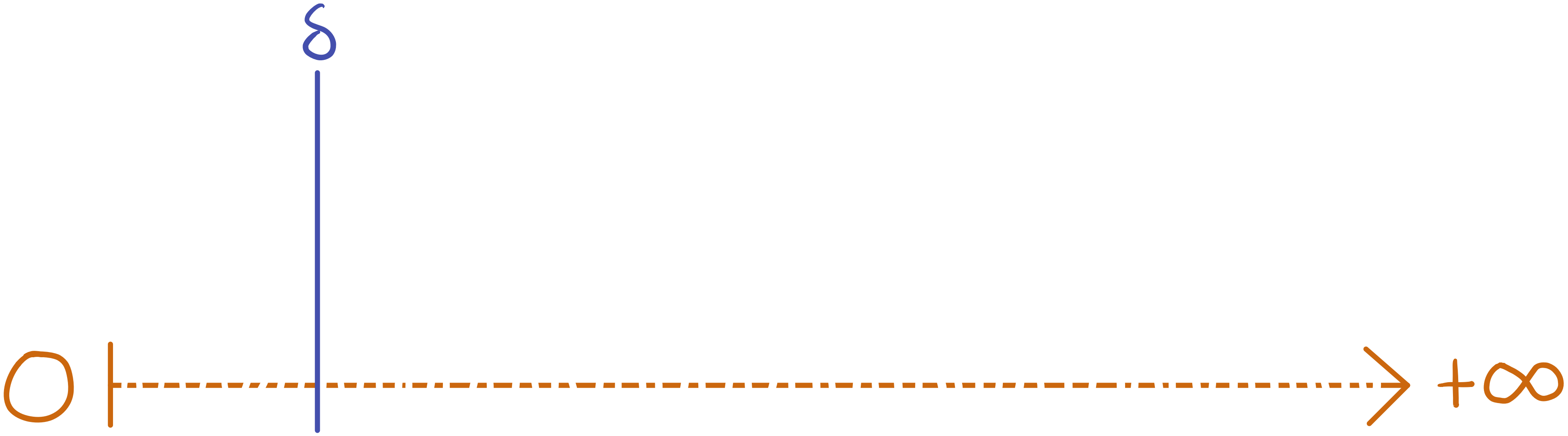
$$x_e \leftarrow x_e + \delta \quad \forall e \in S$$

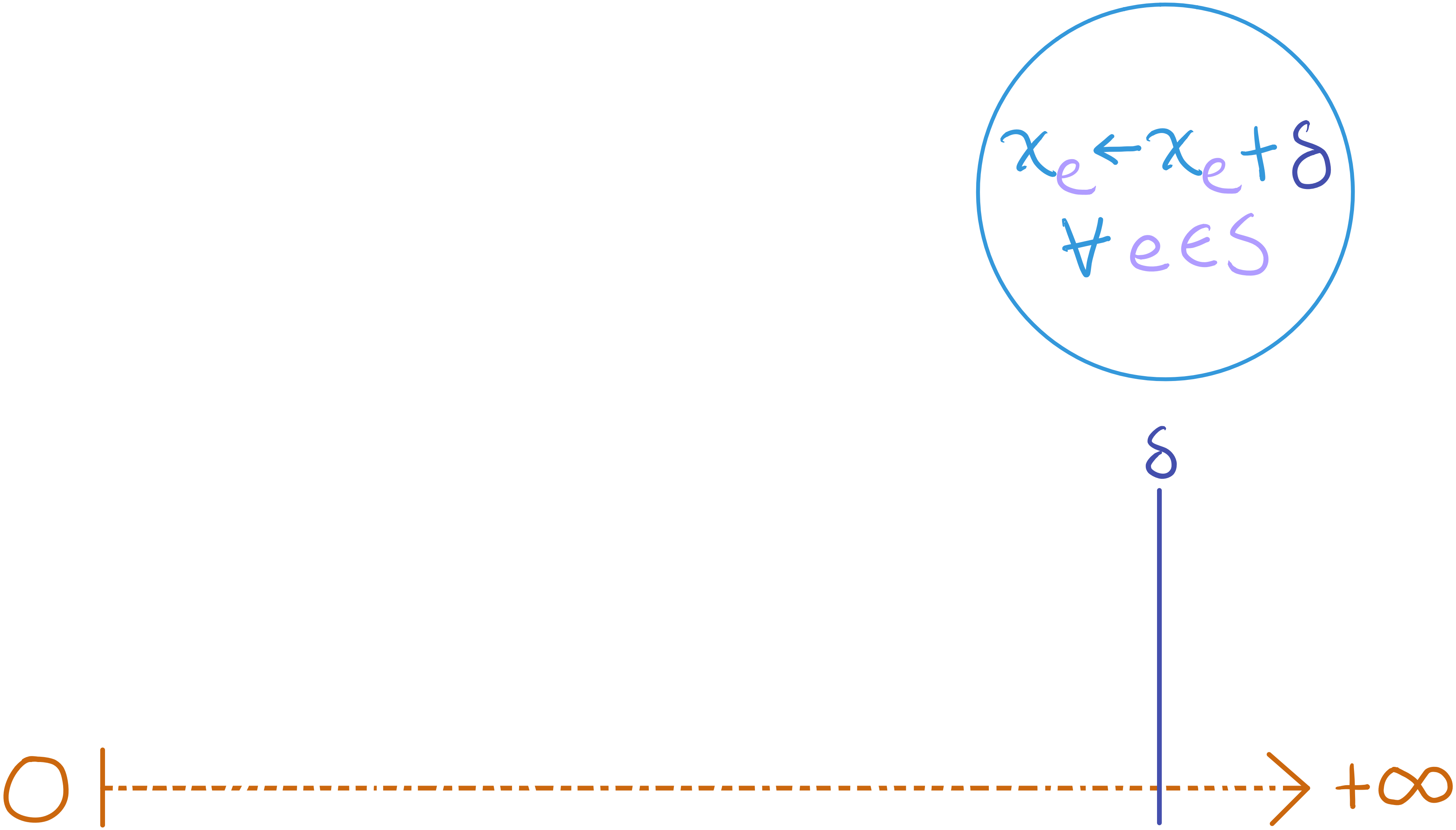
$$S = \{\text{good } e\}$$

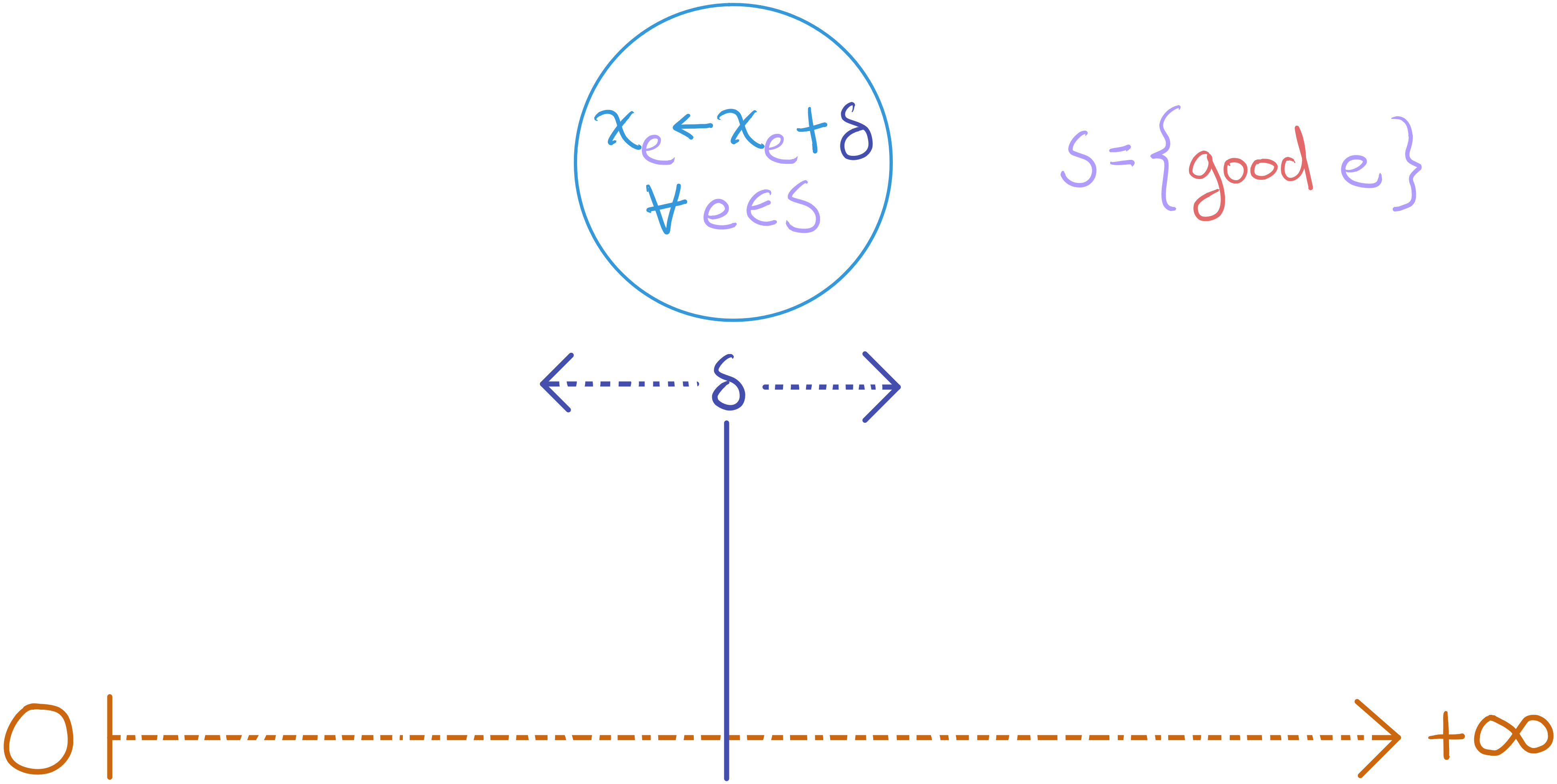


$$x_e \leftarrow x_e + \delta$$
$$\forall e \in S$$

$$S = \{\text{good } e\}$$



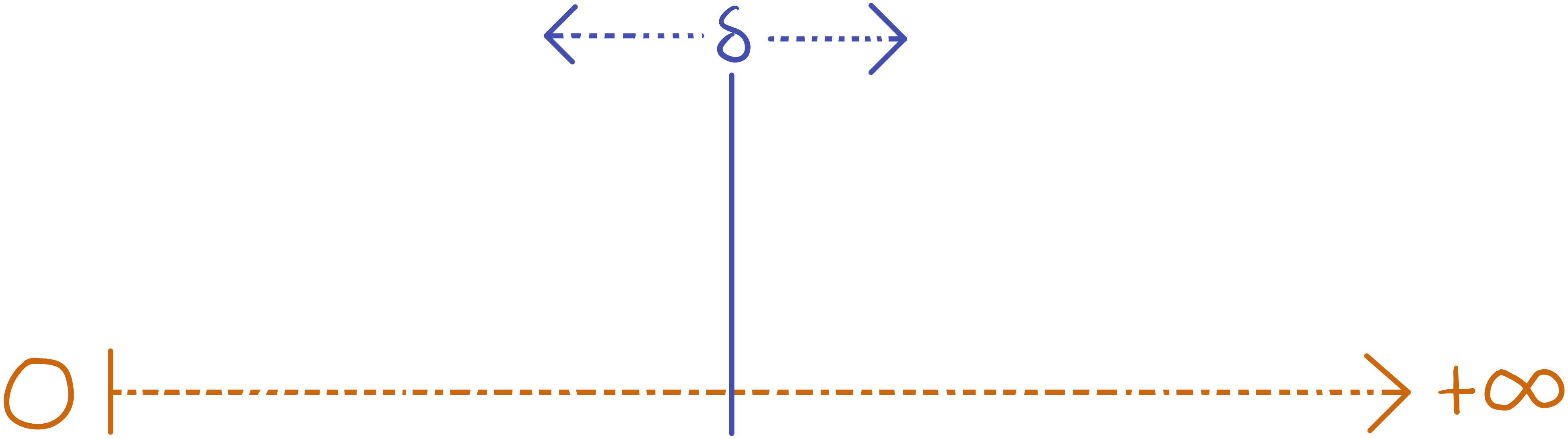

$x_e \leftarrow x_e + \delta$
$\forall e \in S$
δ
0
+∞


$x_e \leftarrow x_e + \delta$
$\forall e \in S$
$S = \{\text{good } e\}$
δ
0
$+\infty$

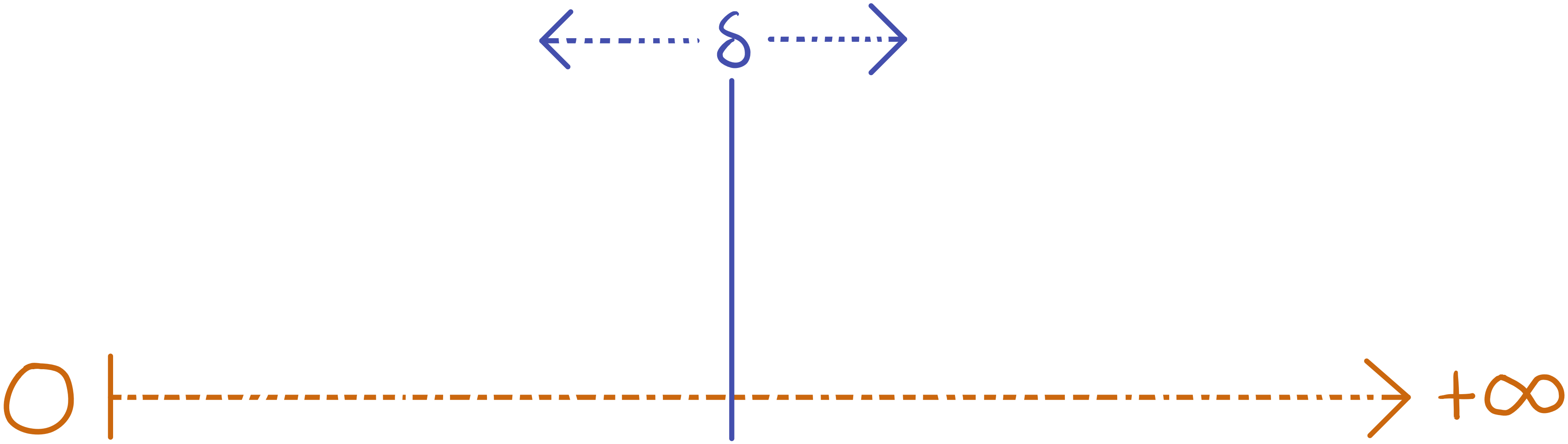
we want $\dfrac{\text{bang}}{\text{buck}} = \dfrac{F(x+\delta S)-F(x)}{\delta|S|}$ ⬆

$$x_e \leftarrow x_e + \delta \quad \forall e \in S$$

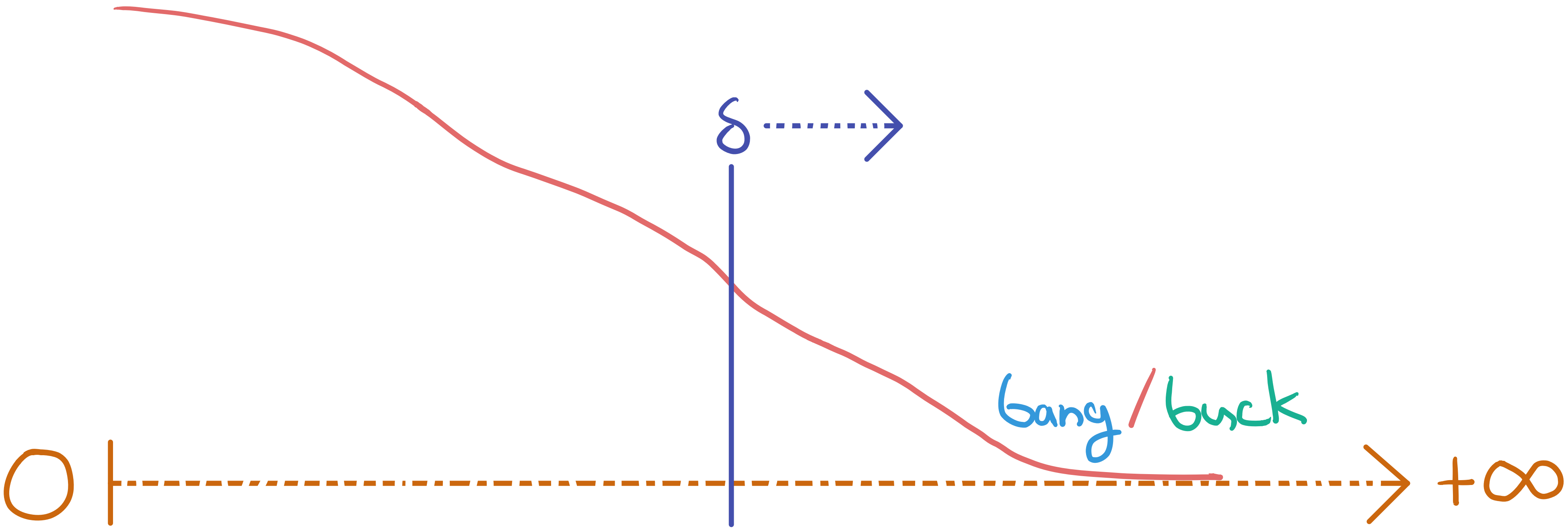
$S = \{\text{good } e\}$



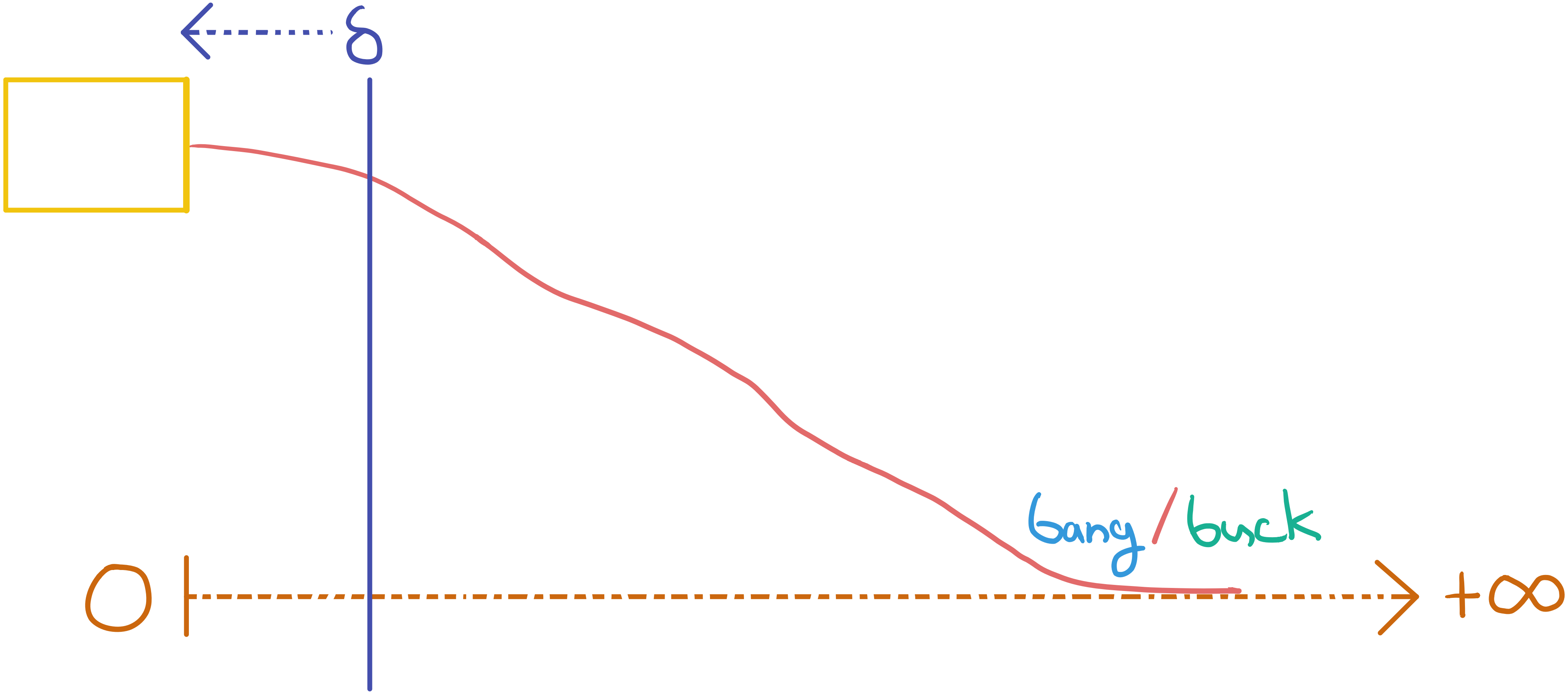
$$\frac{\text{bang}}{\text{buck}} = \frac{F(x+\delta S)-F(x)}{\delta|S|} \begin{matrix} \leftarrow \cdots \\ \leftarrow \cdots \end{matrix} \frac{\text{concave in } \delta}{\text{linear in } \delta}$$



$$\downarrow \text{ in } \delta \left\{ \frac{\text{bang}}{\text{buck}} = \frac{F(x+\delta S)-F(x)}{\delta|S|} \right. \begin{matrix} \longleftarrow \\ \\ \longleftarrow \end{matrix} \frac{\text{concave in } \delta}{\text{linear in } \delta}$$

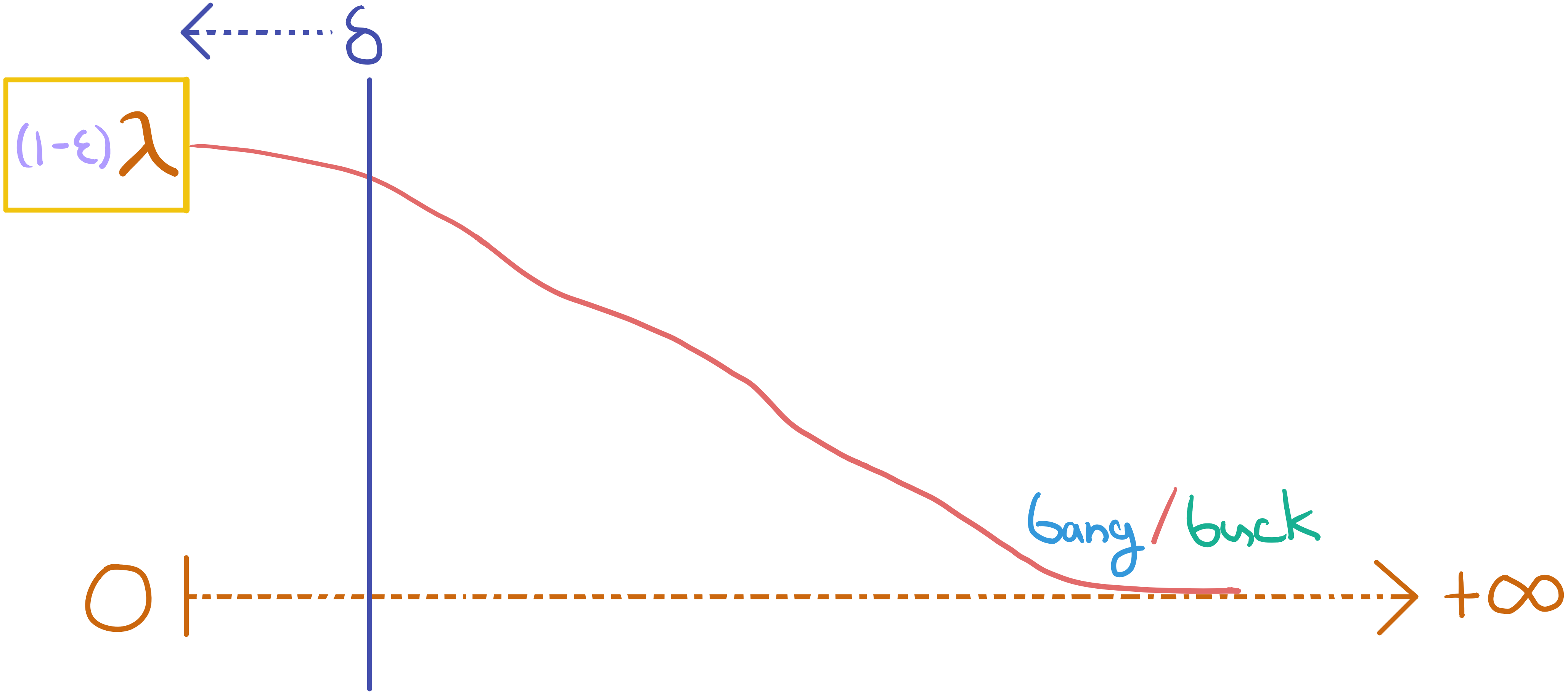


$$\lim_{\delta \downarrow 0} \frac{F(x+\delta S)-F(x)}{\delta |S|} =$$

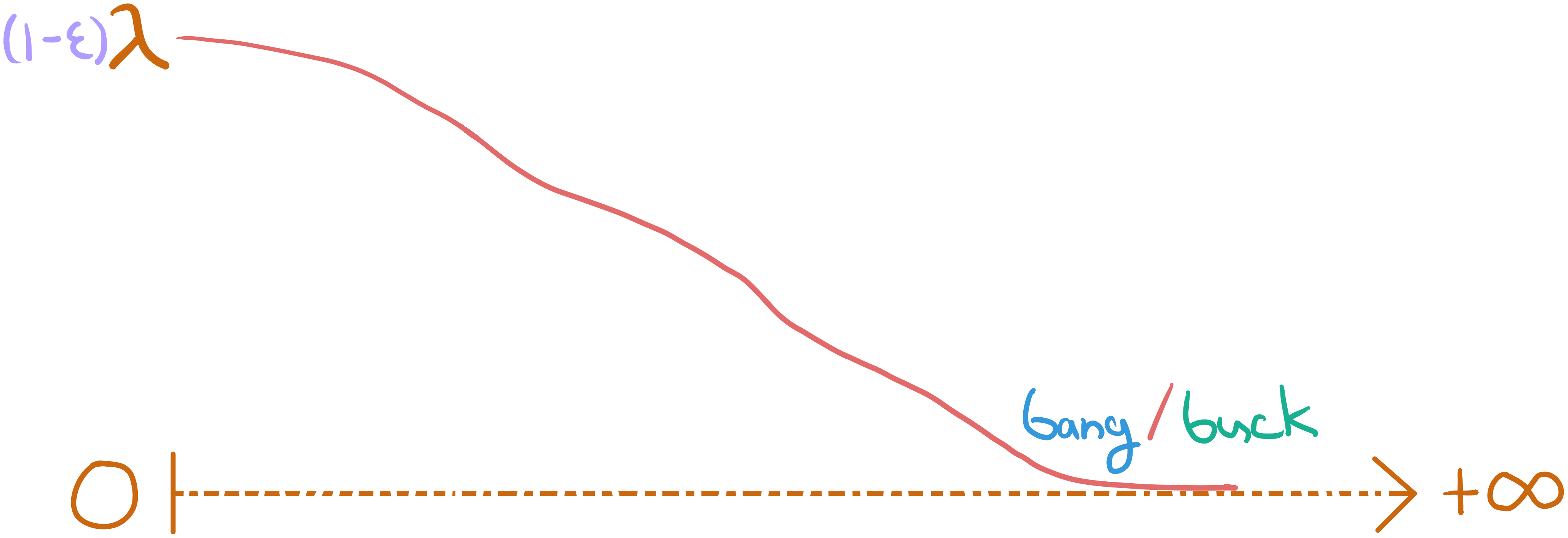


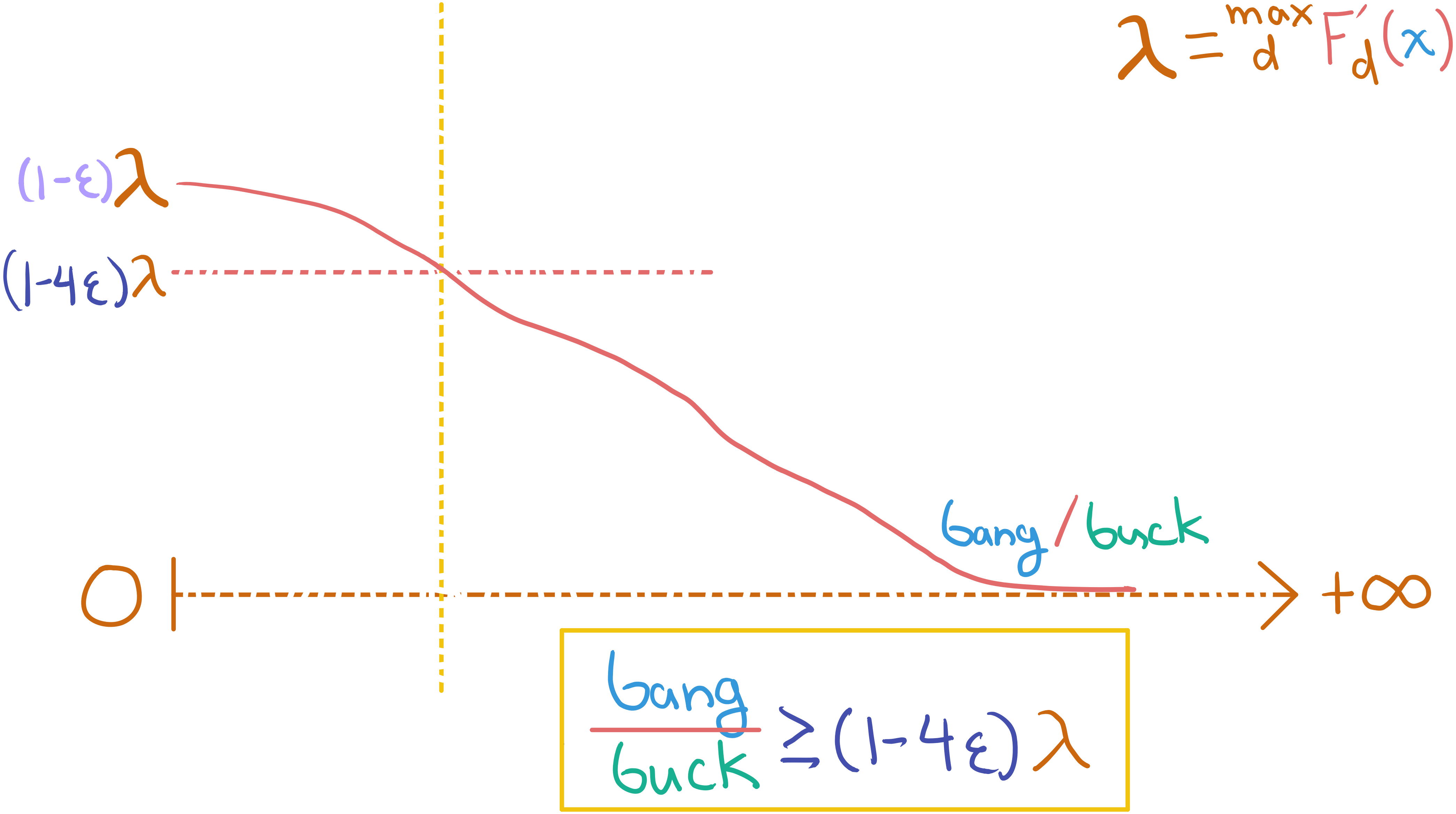
$$\lim_{\delta \downarrow 0} \frac{F(x+\delta S)-F(x)}{\delta |S|} = \frac{1}{|S|} \langle F'(x), S \rangle \geq (1-\varepsilon)\lambda$$

$$\lambda = \max_d F'_d(x)$$



$$\lambda = \max_d F_d'(x)$$




$\lambda = \max_d F'_d(x)$
$(1-\varepsilon)\lambda$
$(1-4\varepsilon)\lambda$
bang/buck
0
$+\infty$
$\boxed{\frac{\text{bang}}{\text{buck}} \geq (1-4\varepsilon)\lambda}$

$$\lambda = \max_d F_d'(x)$$

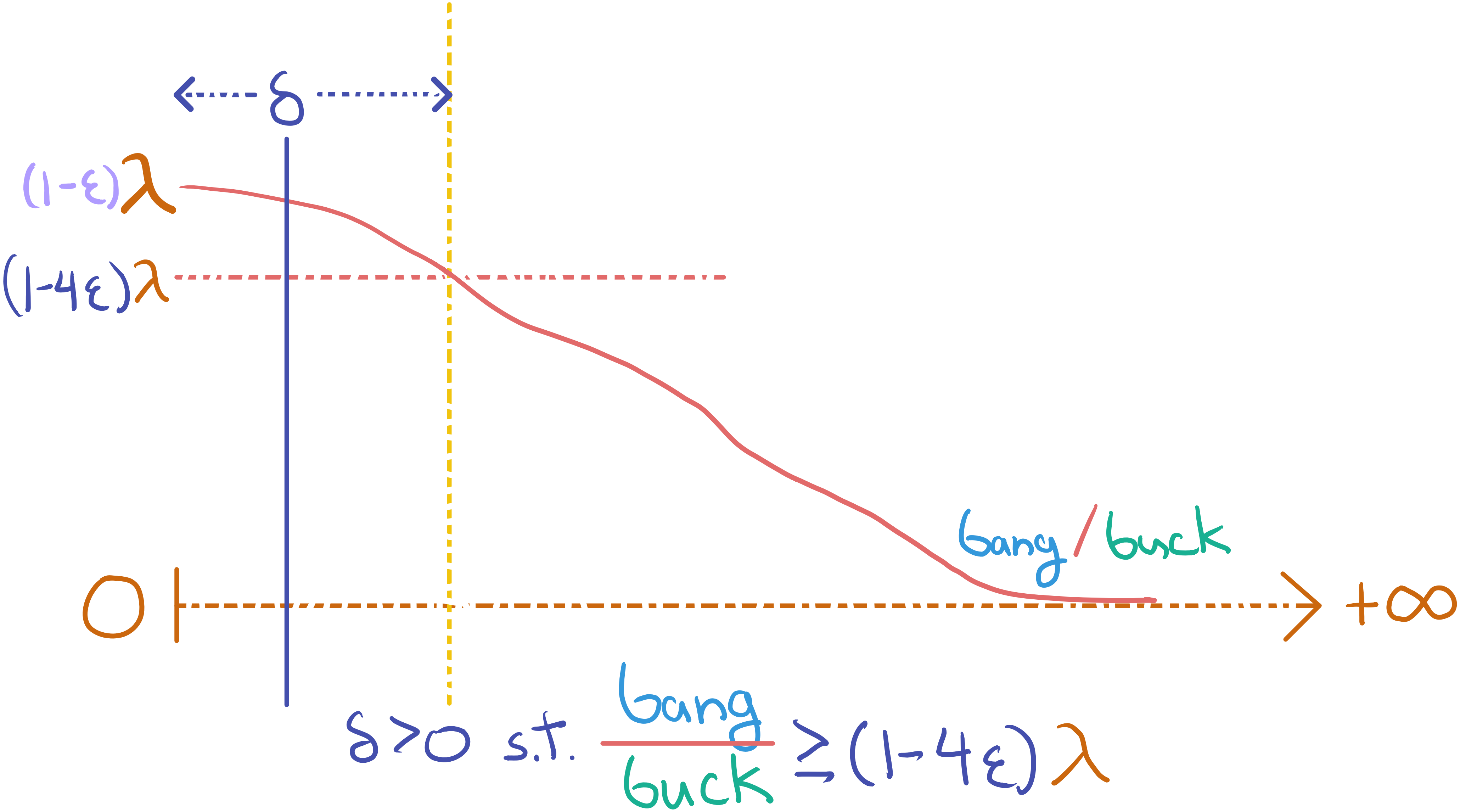
$$\delta > 0 \text{ s.t. } \boxed{\frac{\text{bang}}{\text{buck}} \geq (1-4\varepsilon)\lambda}$$

$$\boxed{\frac{\text{bang}}{\text{buck}} \geq (1-4\varepsilon)\lambda}$$

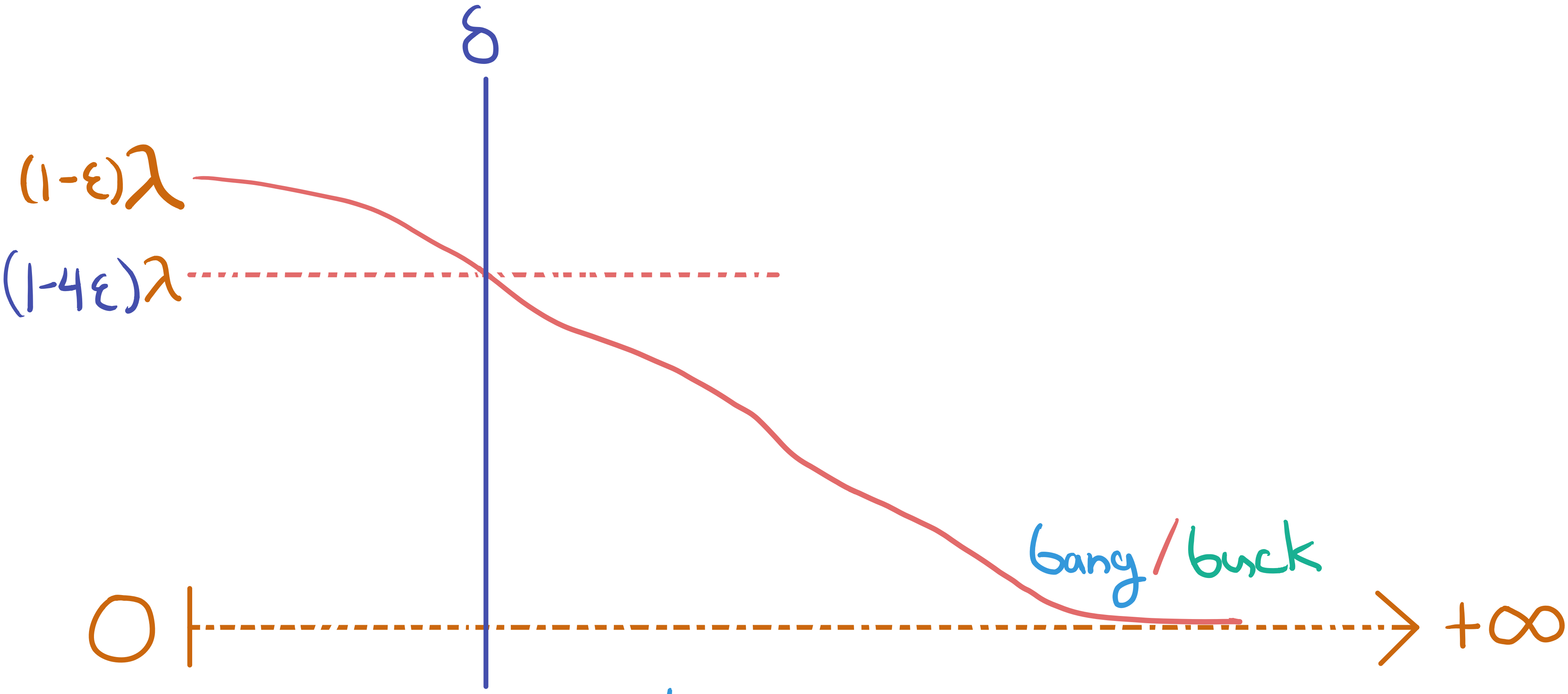
$$\frac{\text{bang}}{\text{buck}} \gtrsim \lambda \gtrsim \text{continuous greedy's } \frac{\text{bang}}{\text{buck}}$$

$\Rightarrow \gtrsim$ Same $(1-\frac{1}{e})$-APX guarantee

as continuous greedy

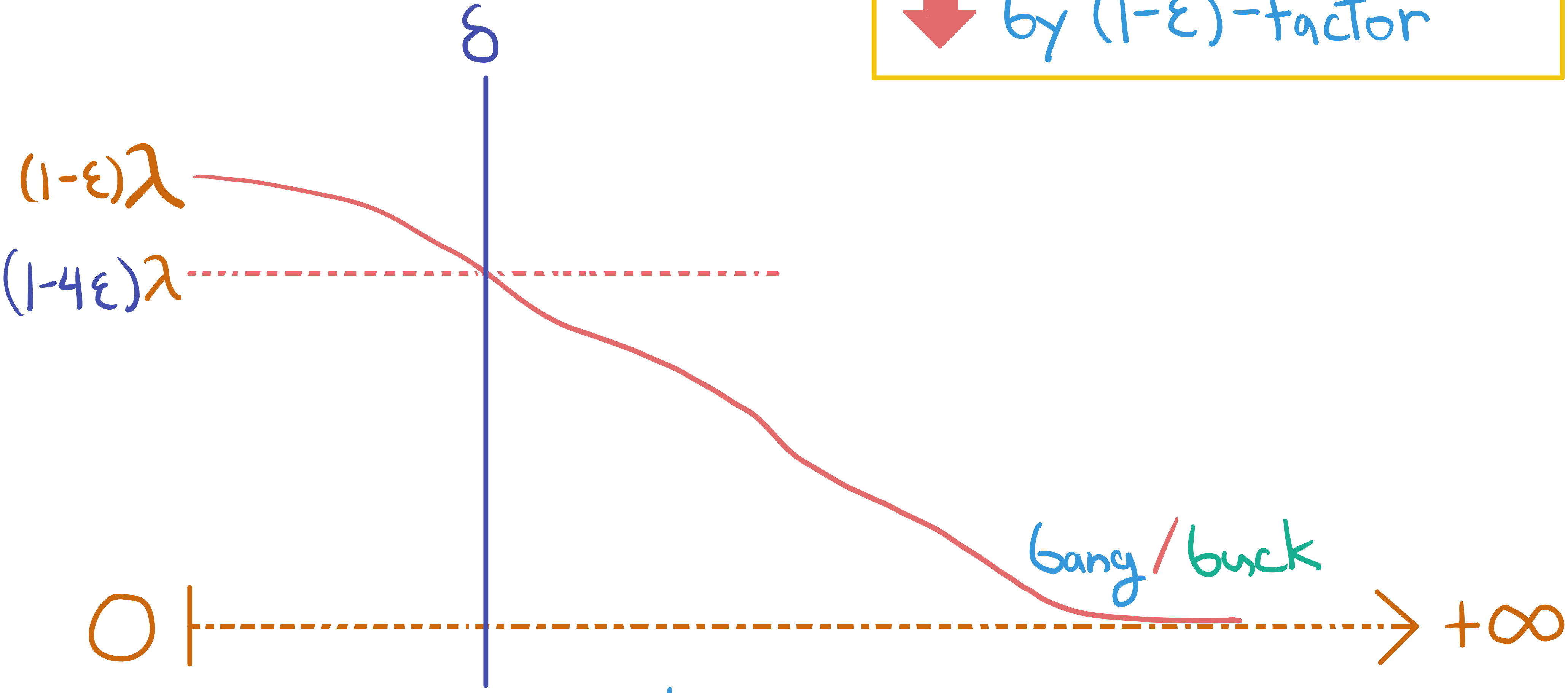


$$\delta > 0 \text{ s.t. } \frac{\text{bang}}{\text{buck}} \geq (1-4\varepsilon)\lambda$$

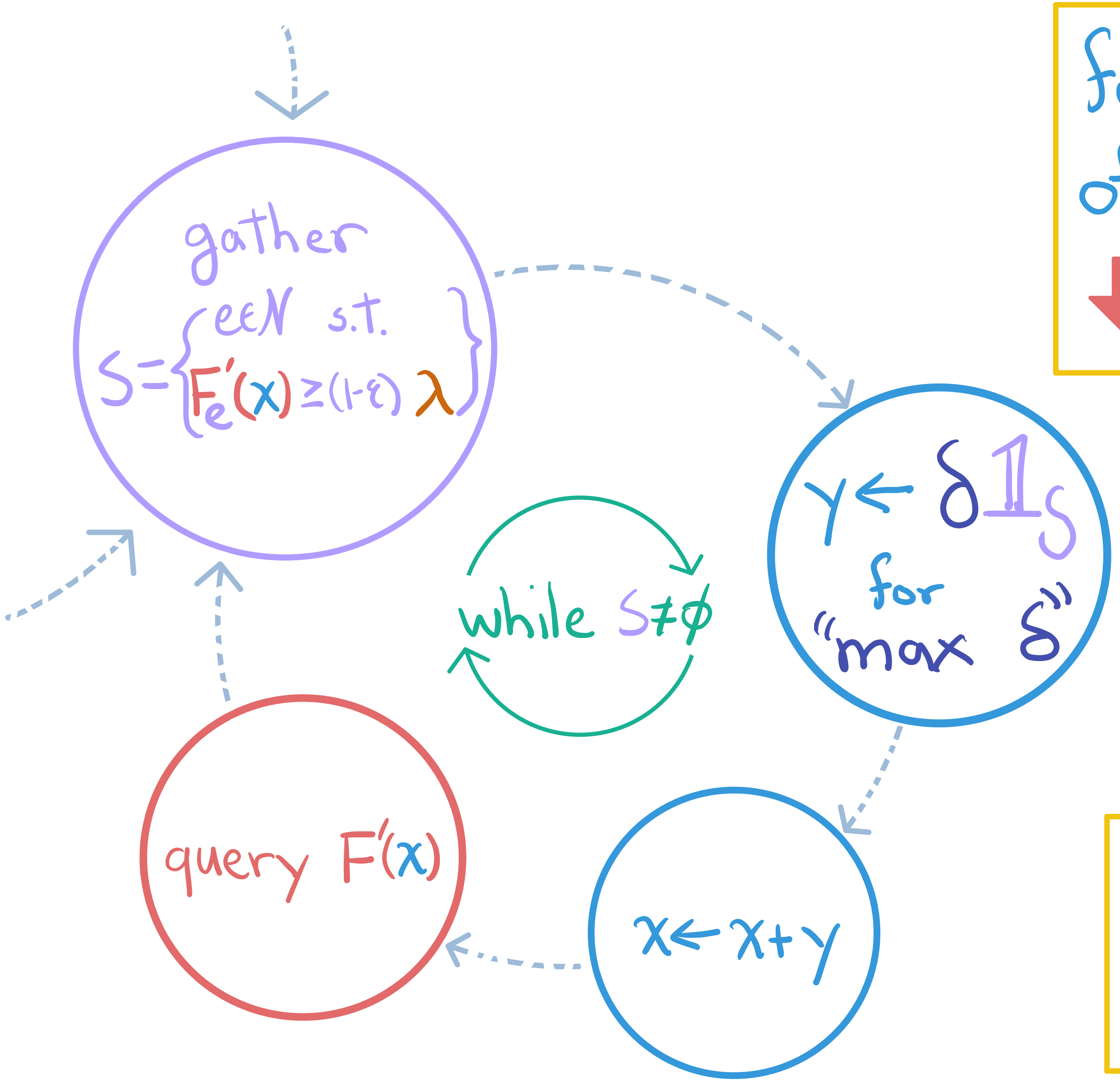


$$\boxed{\max\ \delta > 0}\ \text{s.t.}\ \frac{\text{bang}}{\text{buck}} \geq (1-4\varepsilon)\lambda$$

for max δ, ε-fraction of S has derivatives ⬇ by $(1-\varepsilon)$-factor

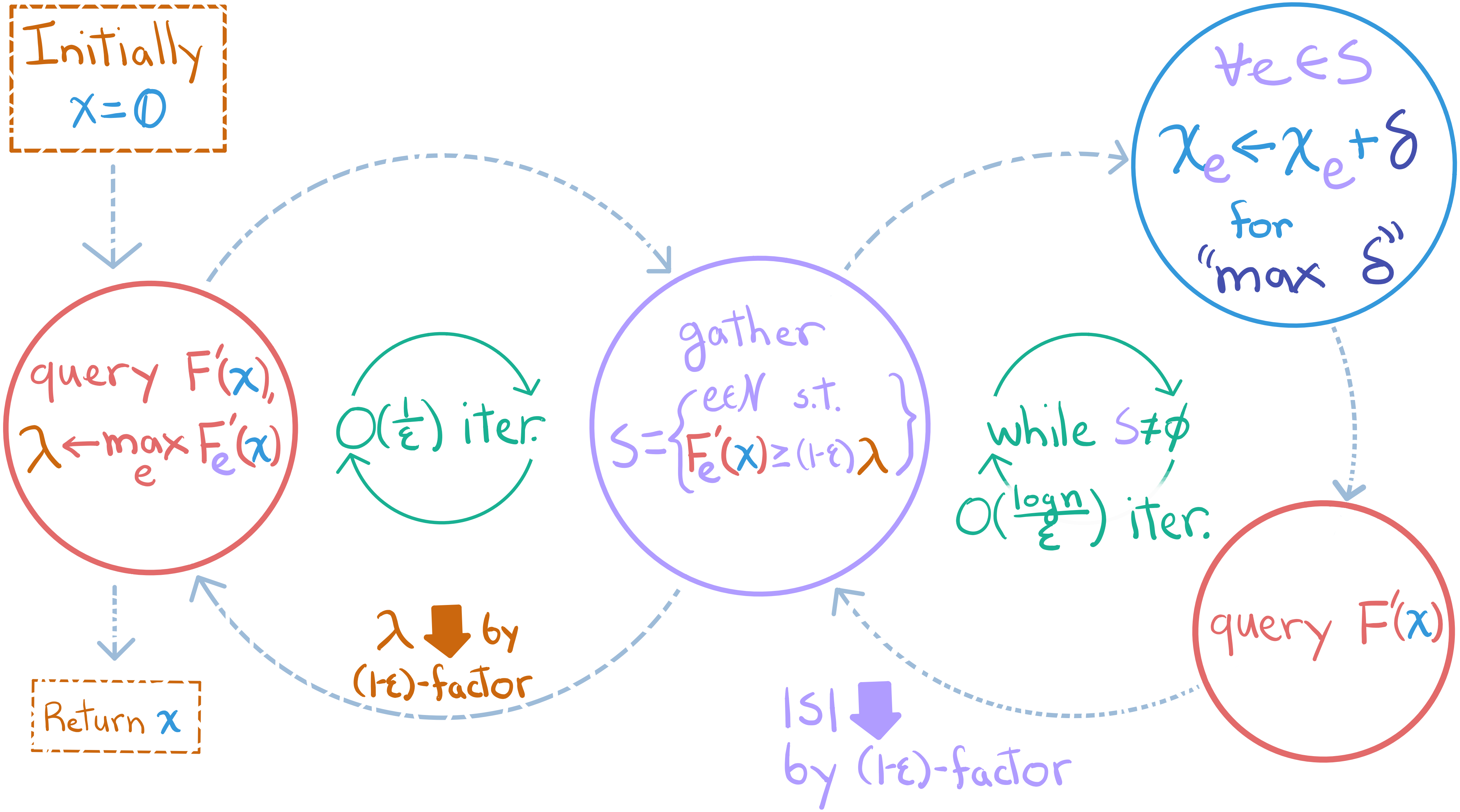


$\boxed{\max\ \delta > 0}$ s.t. $\frac{\text{bang}}{\text{buck}} \geq (1-4\varepsilon)\lambda$

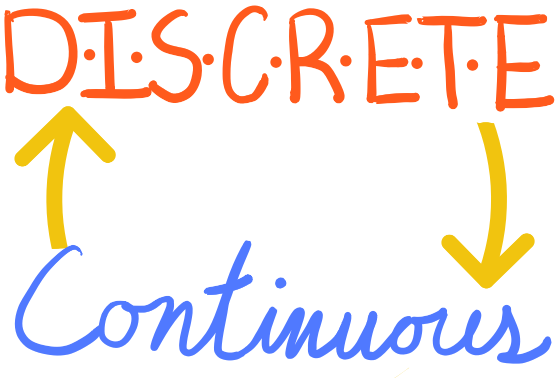


for max δ, ε-fraction of S has derivatives ⬇ by $(1-\varepsilon)$-factor

after $O(\frac{\log n}{\varepsilon})$ iter., $S = \emptyset$



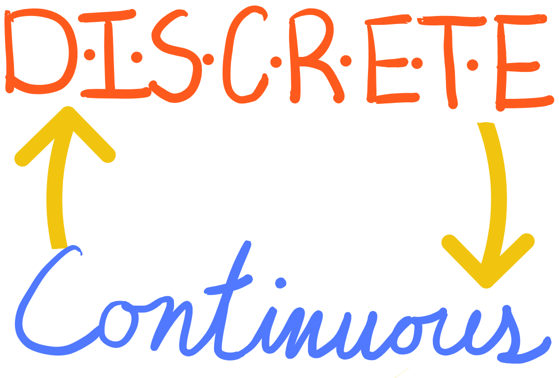
$$O\left(\frac{1}{\varepsilon}\right) \times O\left(\frac{\log n}{\varepsilon}\right) = O\left(\frac{\log n}{\varepsilon^2}\right) \text{ depth}$$



DISCRETE

Continuous

can we get a DISCRETE algorithm?



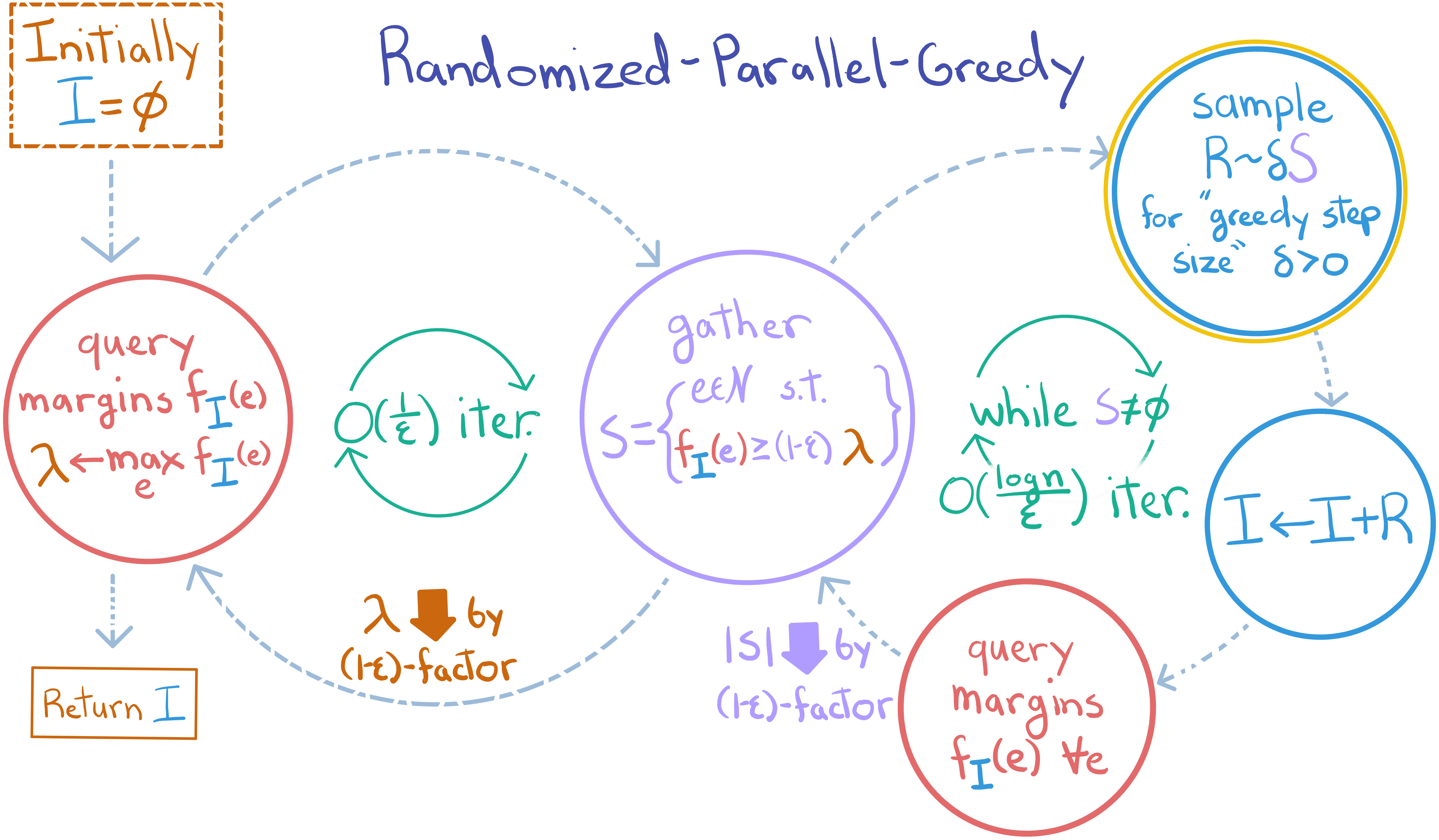
D·I·S·C·R·E·T·E

Continuous

can we get a D·I·S·C·R·E·T·E algorithm?

yes: round (i.e. sample) online

Randomized-Parallel-Greedy



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

Parallelizing greedy for cardinality and beyond

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

[Fisher, Nemhauser, Wolsey]

DISCRETE

Continuous

- continuous approach
- calculus "flattens" submodularity
- "greedy step size" δ

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

general packing constraints in parallel

matroid constraints in parallel

2nd half of talk

Given:

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

  $(\mathcal{I} \subseteq 2^N)$

Goal

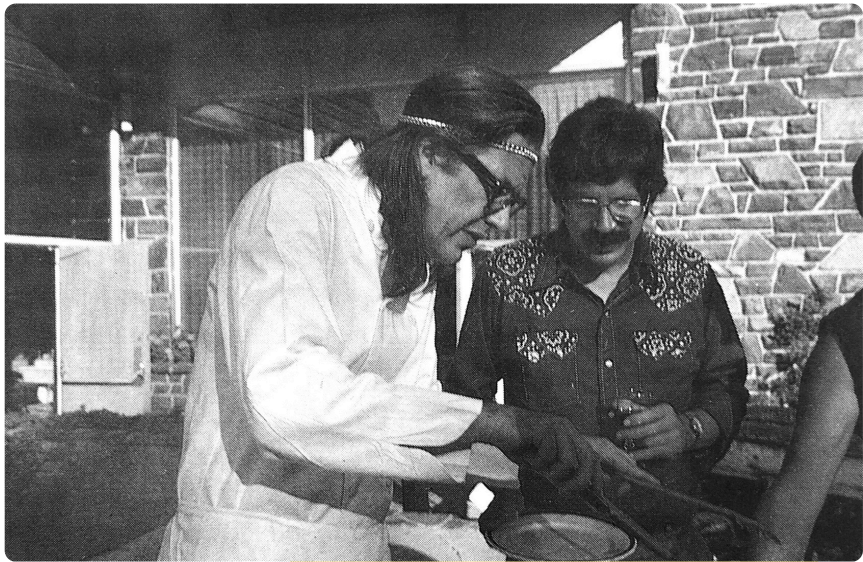
maximize $f(I)$ s.t. $I \in \mathcal{I}$

in parallel in the oracle model

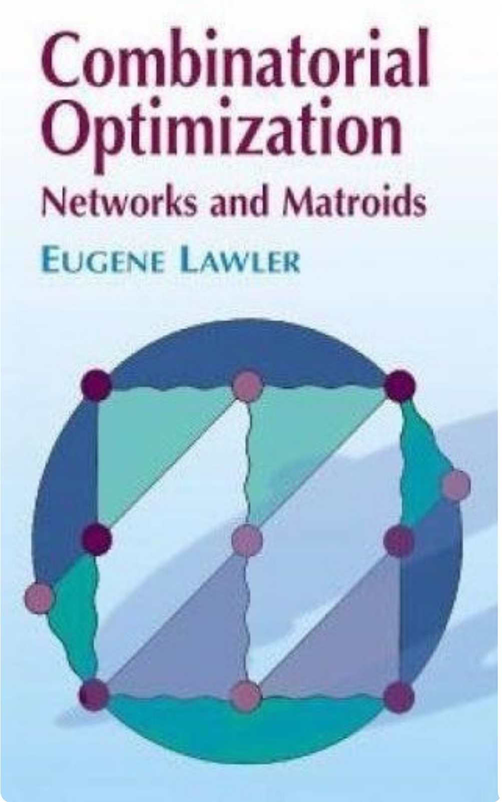
why matroids?

①

matroids play a central & unifying role in combinatorial optimization



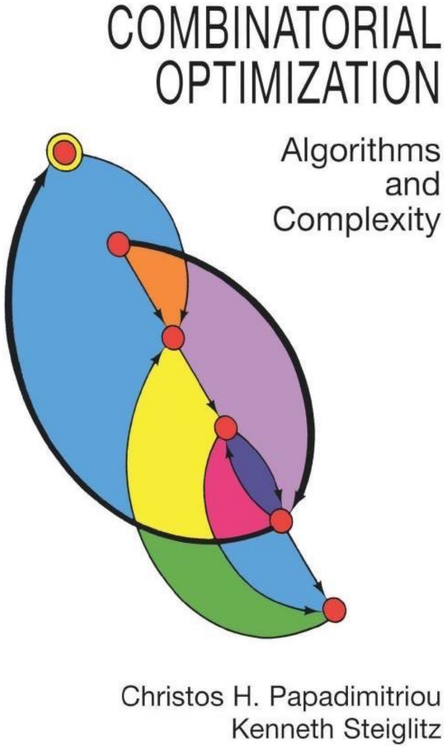
Edmonds & Lawler, 1977



Combinatorial Optimization

Networks and Matroids

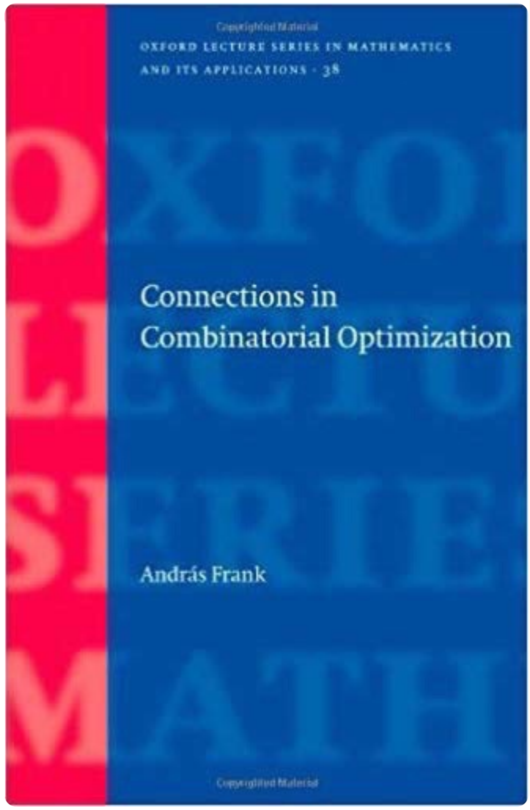
EUGENE LAWLER



COMBINATORIAL OPTIMIZATION

Algorithms and Complexity

Christos H. Papadimitriou
Kenneth Steiglitz



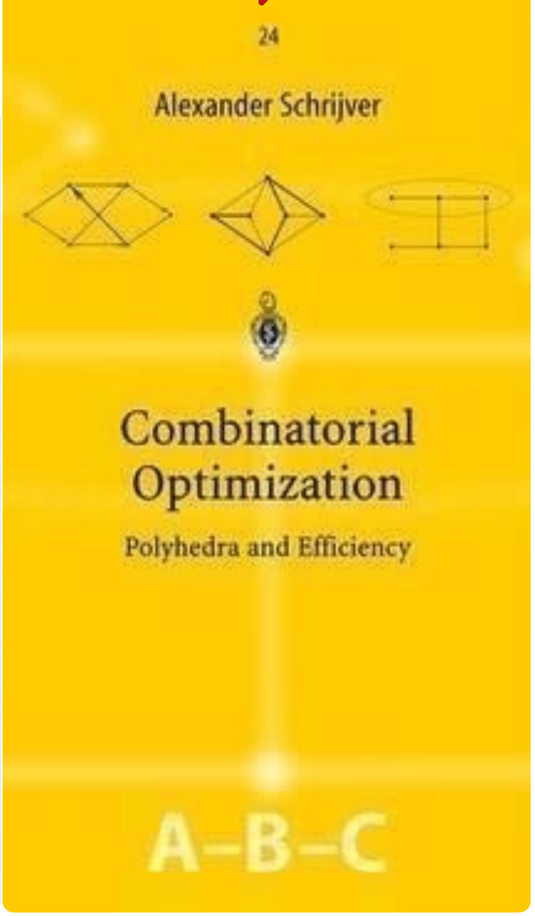
Copyrighted Material

OXFORD LECTURE SERIES IN MATHEMATICS AND ITS APPLICATIONS · 38

Connections in Combinatorial Optimization

András Frank

Copyrighted Material



24

Alexander Schrijver

Combinatorial Optimization

Polyhedra and Efficiency

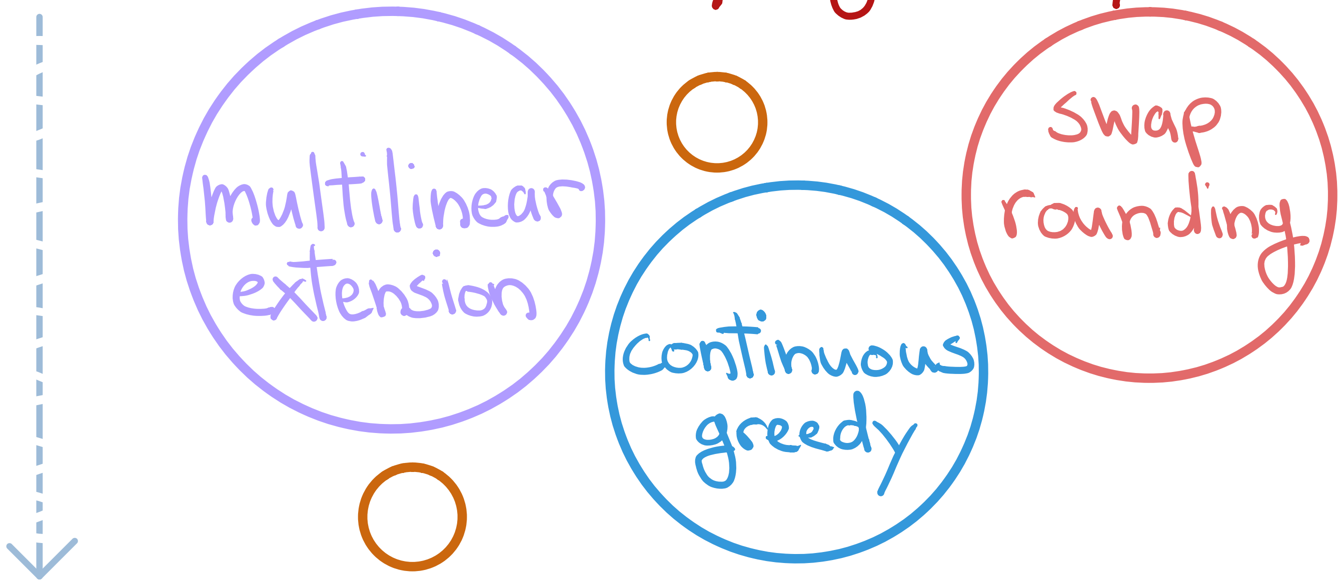
A–B–C

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

1978: $\frac{1}{2}$ APX by greedy



2007: $1-\frac{1}{e}$ APX for monotone

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

1978: ½ APX by greedy

2007: $1-\frac{1}{e}$ APX for monotone

2011: $\frac{1}{e}$ APX for nonnegative

③ generality

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

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

$\mathcal{N}$ "ground set"

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

s.t. (1)

(2)

(3)

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

$\mathcal{N}$ "ground set"

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

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

(2)

(3)

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

$\mathcal{N}$ "ground set"

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

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

(2) hereditary:

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

(3)

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

$\mathcal{N}$ "ground set"

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

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

(2) hereditary:

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

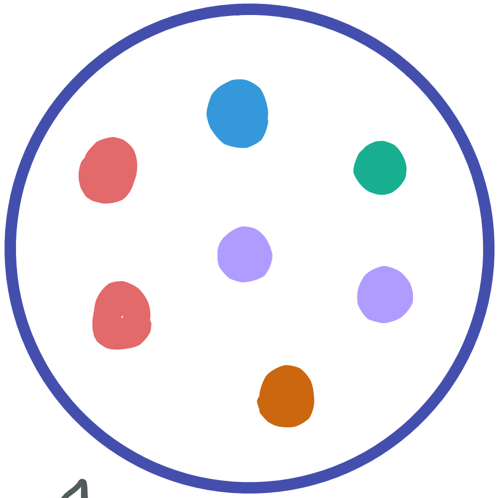
(3) extendability: $S, T \in \mathcal{I},\ |S| < |T|$

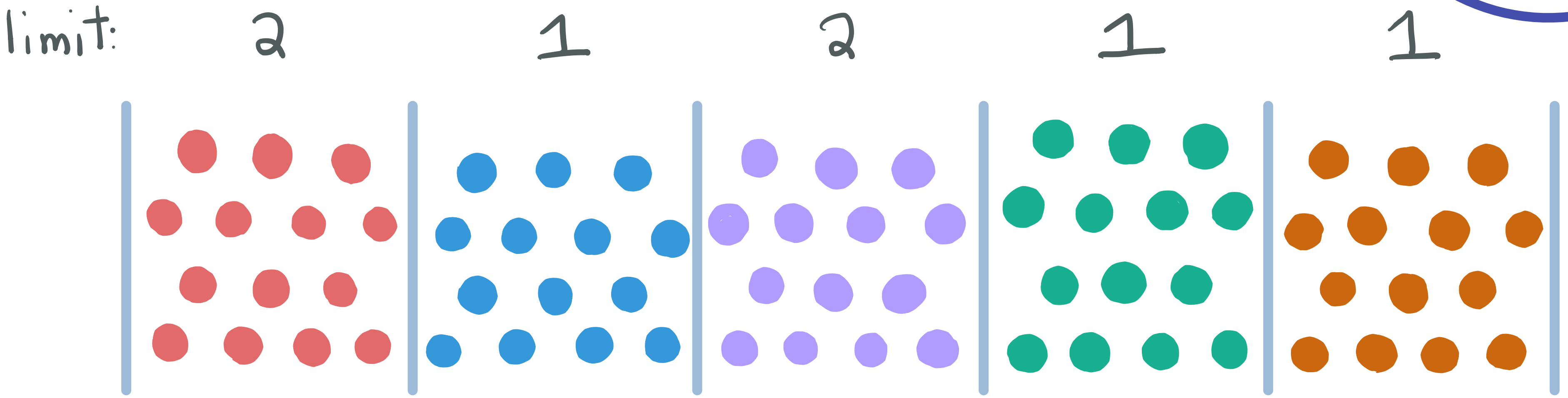
$\Rightarrow S + e \in \mathcal{I}$ for some $e \in T \setminus S$

e.g. "partition matroid"

$\mathcal{N}$ = set of elements w/ colors

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

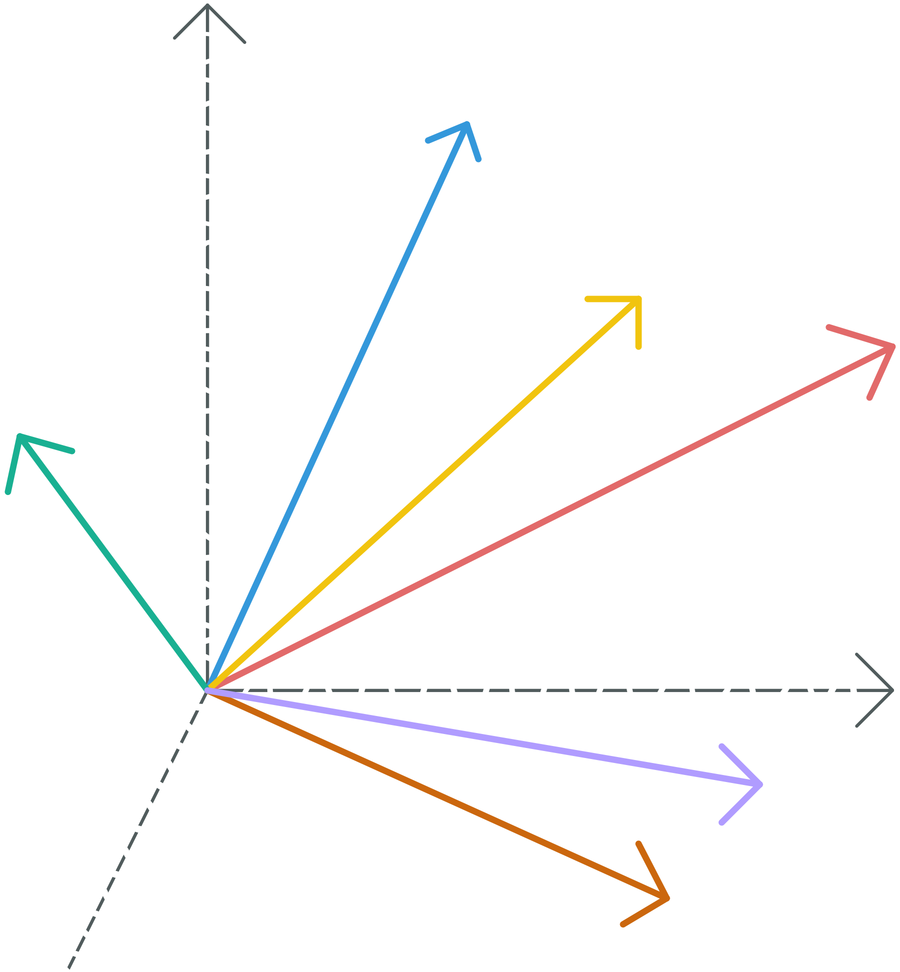




e.g. "linear matroid"

$\mathcal{N}$ = set of vectors in vector space

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

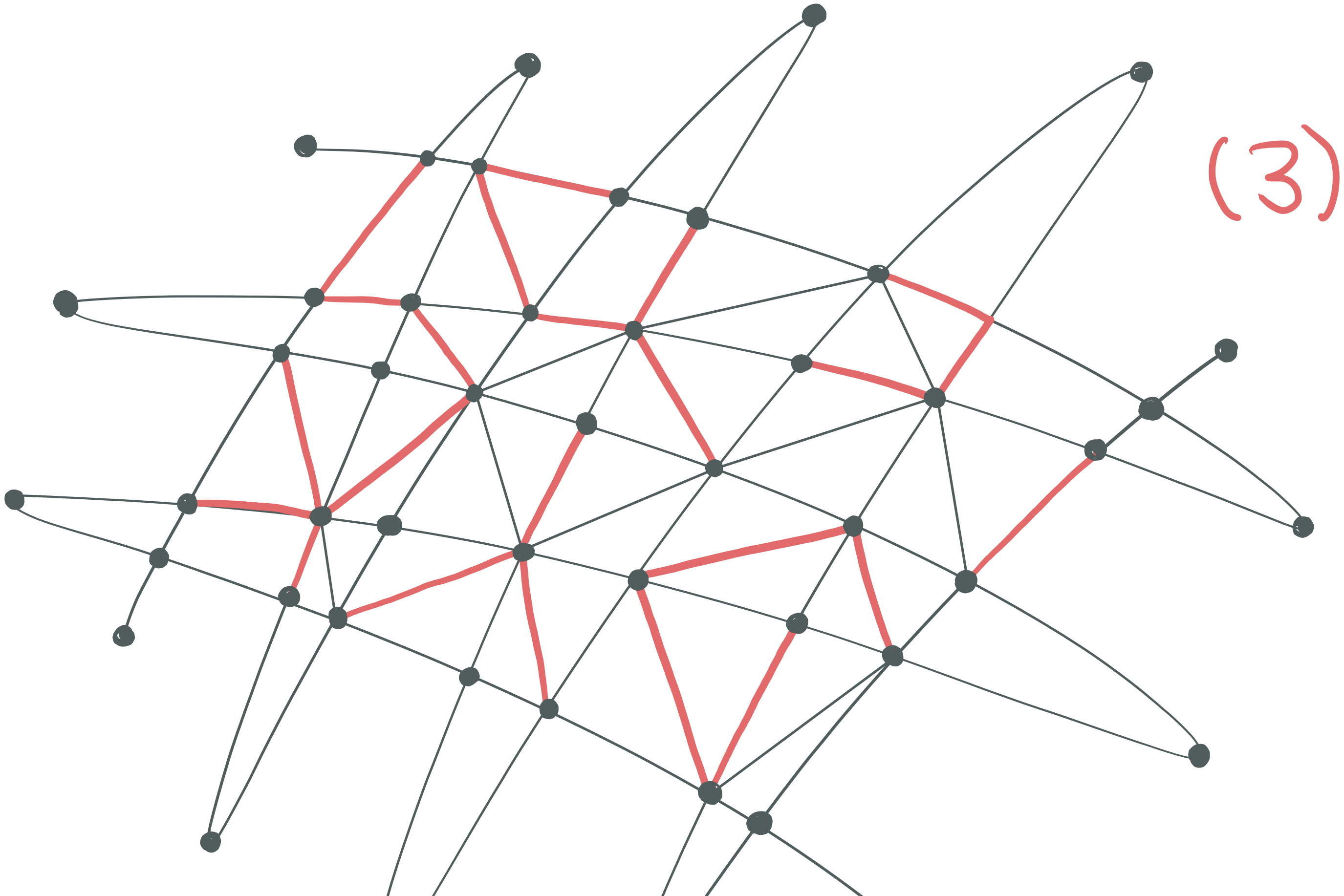


(3) extendability:

one can always extend an independent set of vectors to a base e.g. "graphic matroid"

$\mathcal{N}$ = edges of a fixed graph

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



(3) extendability:

every forest in connected graph extends to a spanning tree

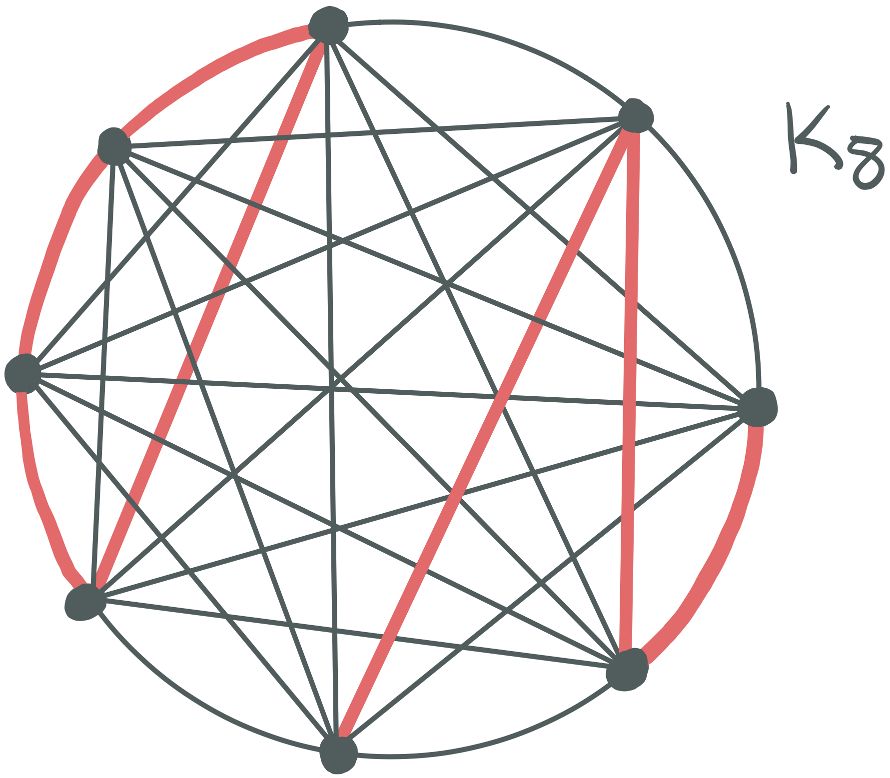
Matroid terminology

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

max. cardinality of any independent set in S

e.g. graphic matroid

rank(S) $= n - \#\left\{\begin{array}{c}\text{connected} \\ \text{components} \\ \text{in } S\end{array}\right\}$



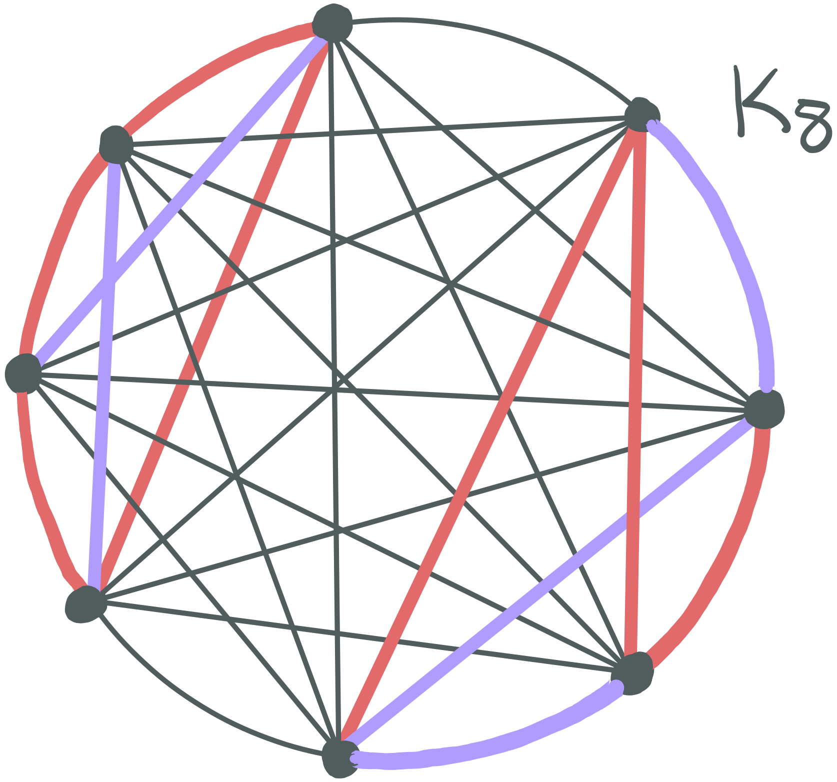
rank = 8-2=6

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

weaker* than rank oracle

PRAM rank oracles known for linear matroids, graphic matroids

Given:

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

Goal

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

in parallel in the oracle model

$1/2$-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

Initially
$S=\emptyset$

Greedy algorithm

(for matroids)

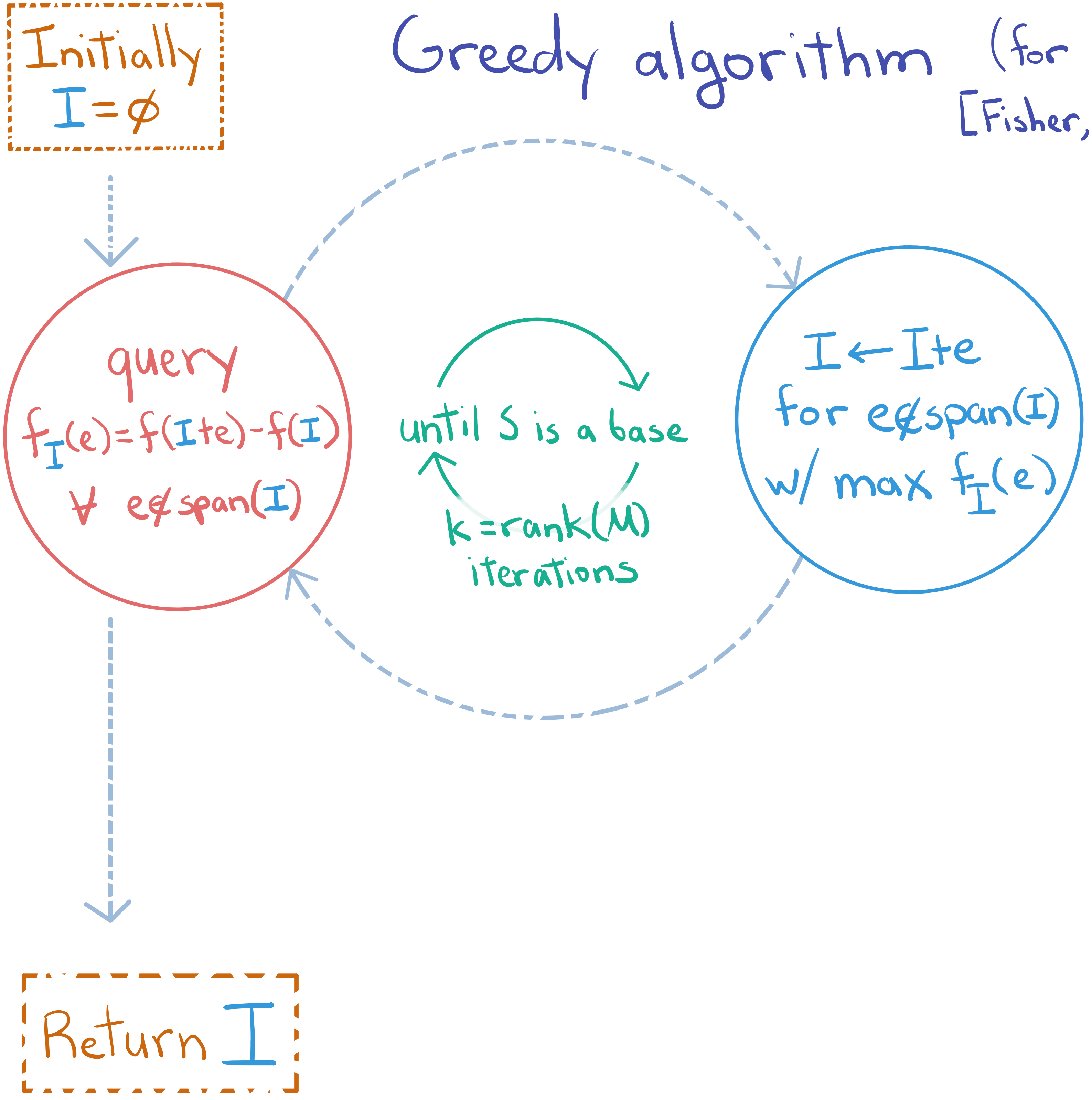
[Fisher, Nemhauser, Wolsey]

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

Return S

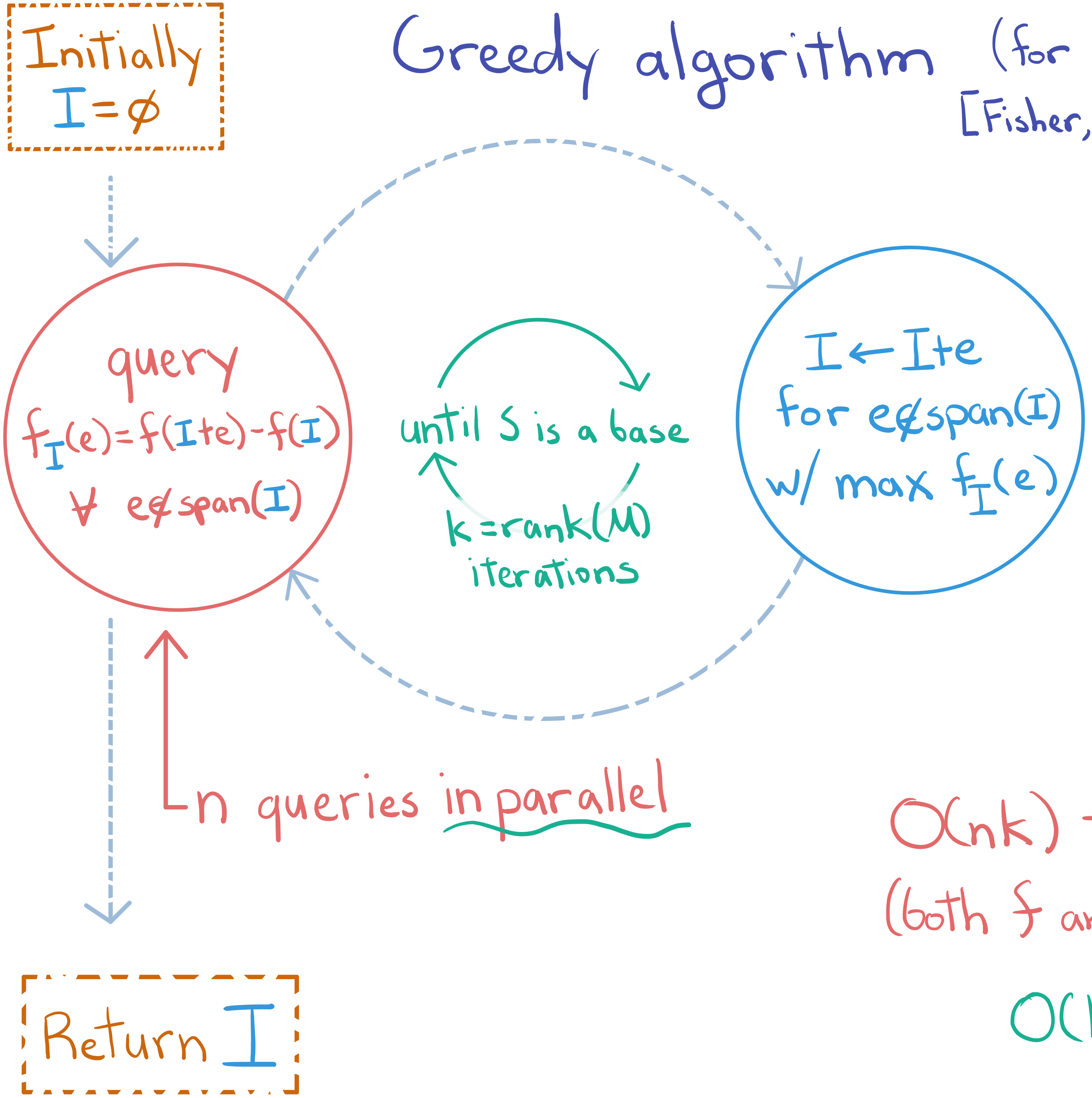
Greedy algorithm (for matroids)

[Fisher, Nemhauser, Wolsey]



Greedy algorithm (for matroids)

[Fisher, Nemhauser, Wolsey]



Initially $I=\emptyset$

query

$f_I(e)=f(I+e)-f(I)$

$\forall\ e\notin \mathrm{span}(I)$

until S is a base

$k=\mathrm{rank}(M)$ iterations

$I\leftarrow I+e$

for $e\notin \mathrm{span}(I)$

w/ max $f_I(e)$

n queries in parallel

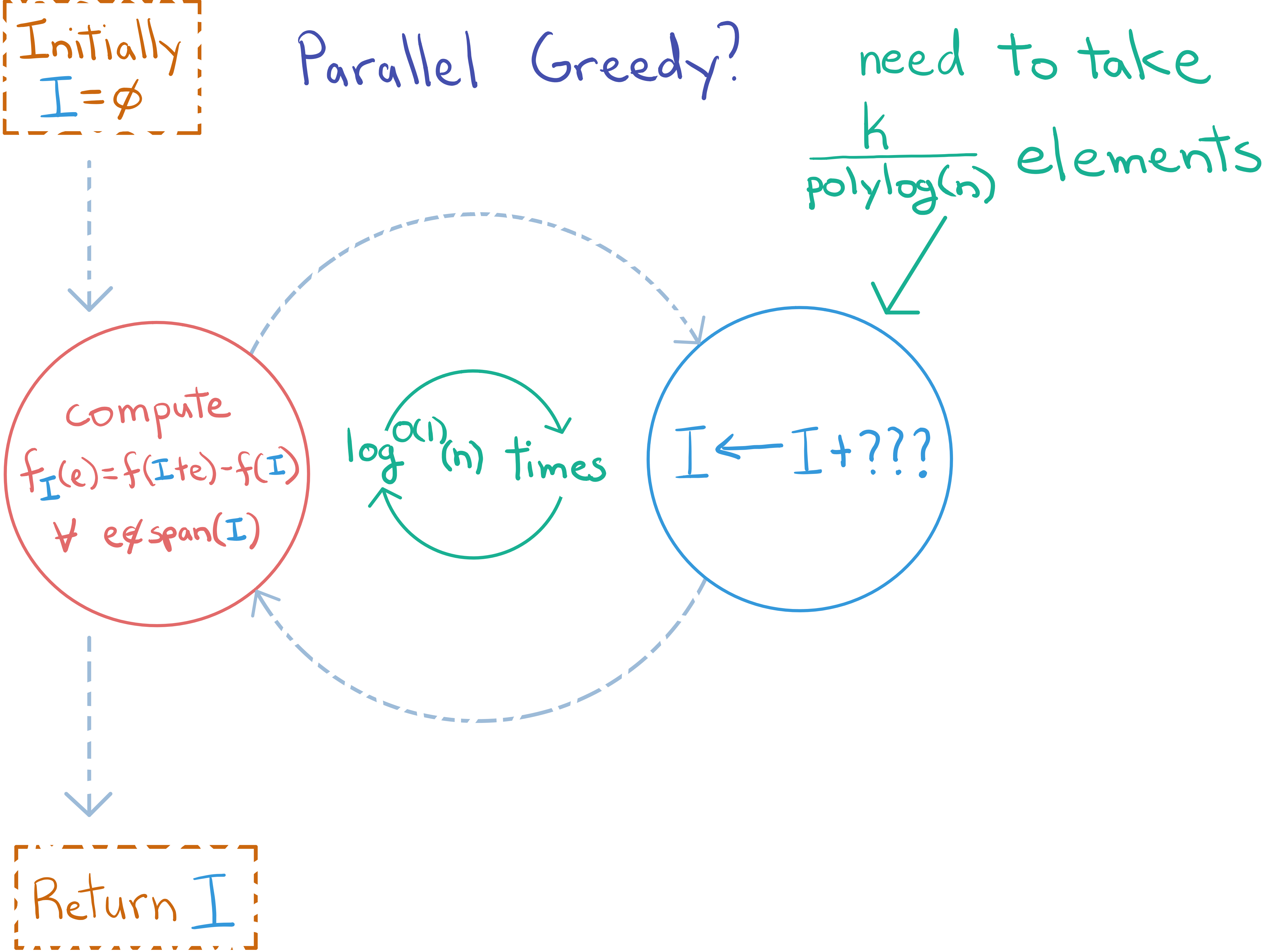
Return I

$O(nk)$ total queries

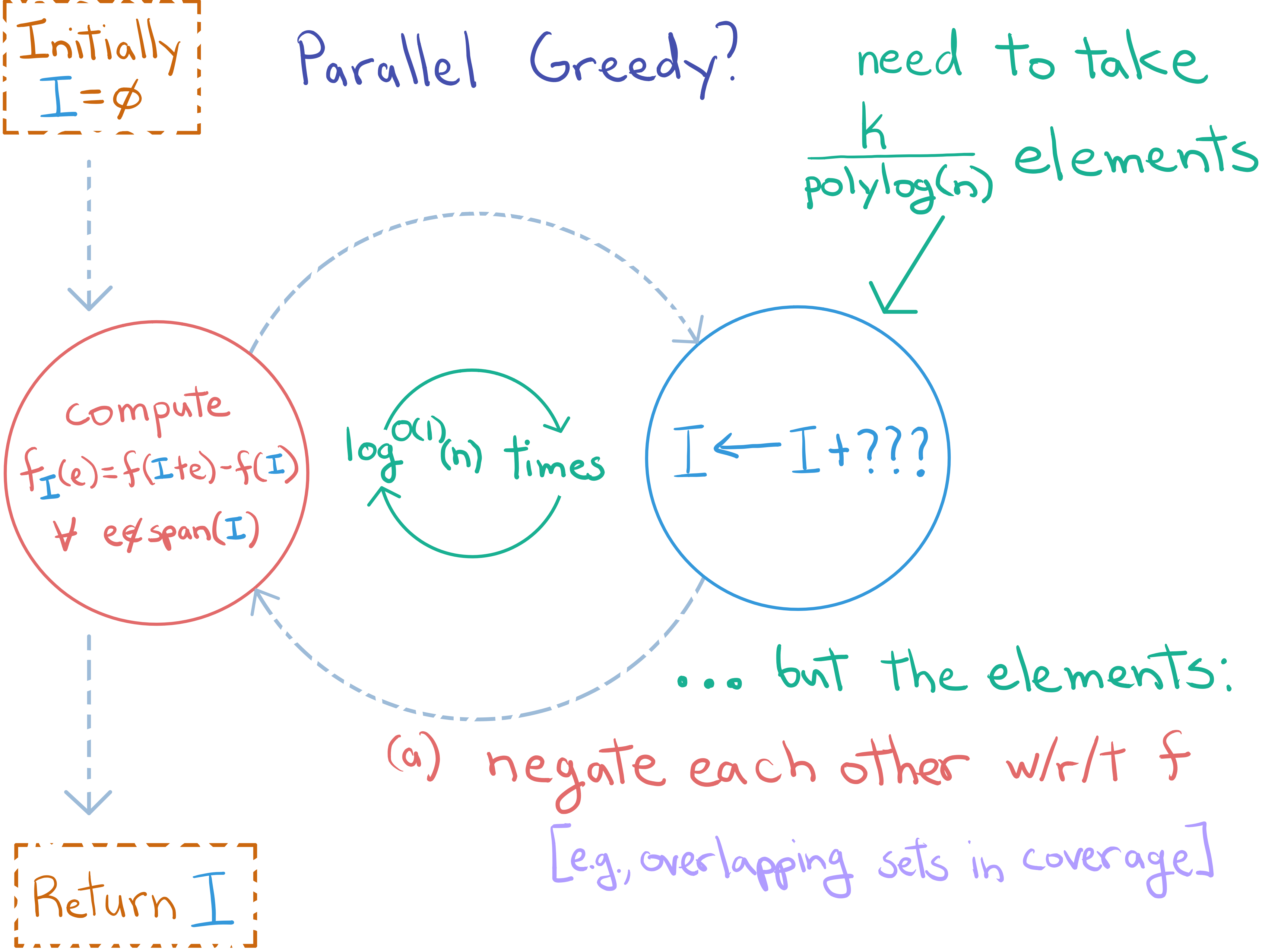
(both f and span oracles)

$O(k)$ adaptivity

Parallel Greedy?



Parallel Greedy?



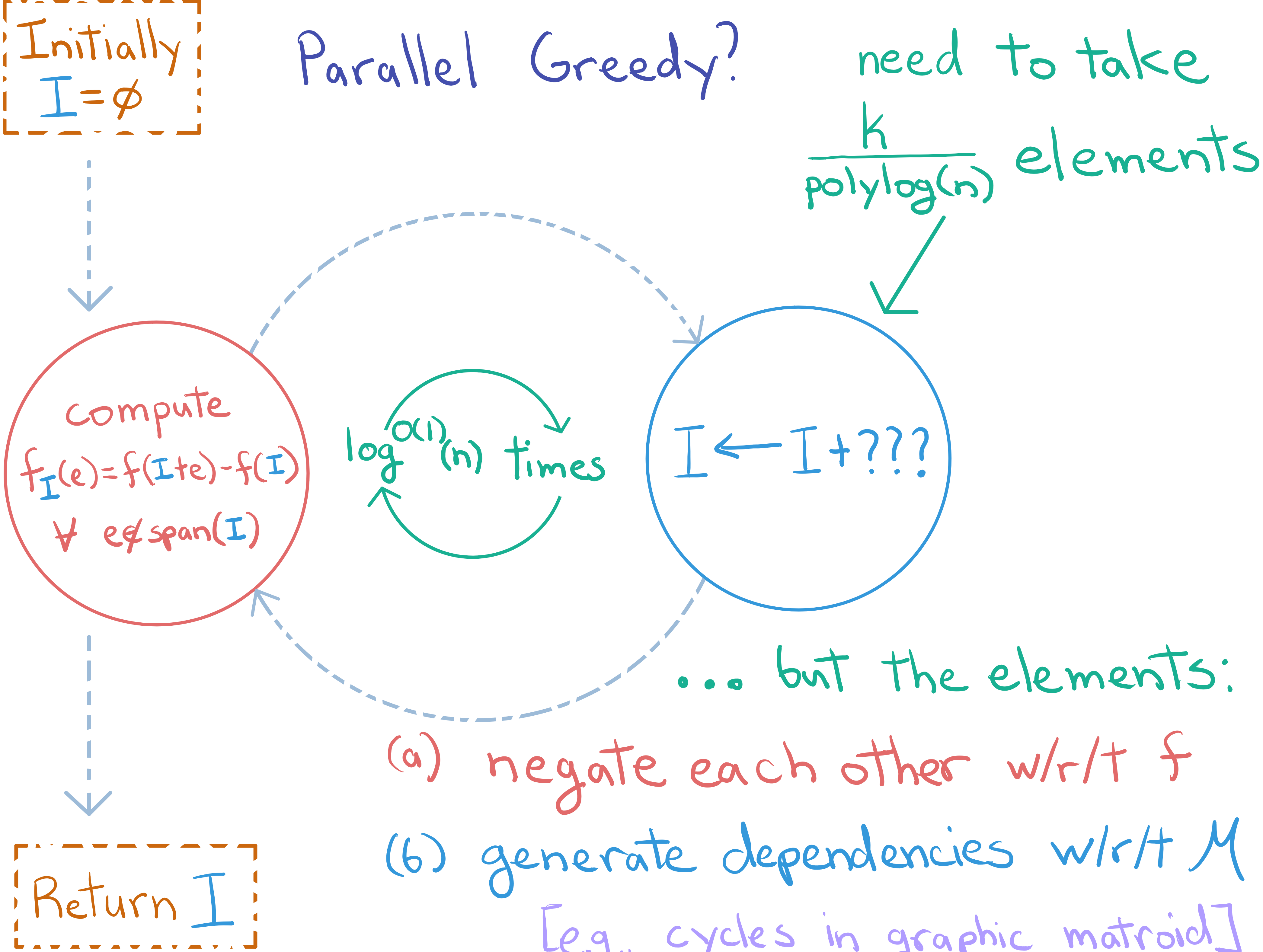
need to take $\frac{k}{\mathrm{polylog}(n)}$ elements

... but the elements:

(a) negate each other w/r/t f

[e.g., overlapping sets in coverage]

Parallel Greedy?



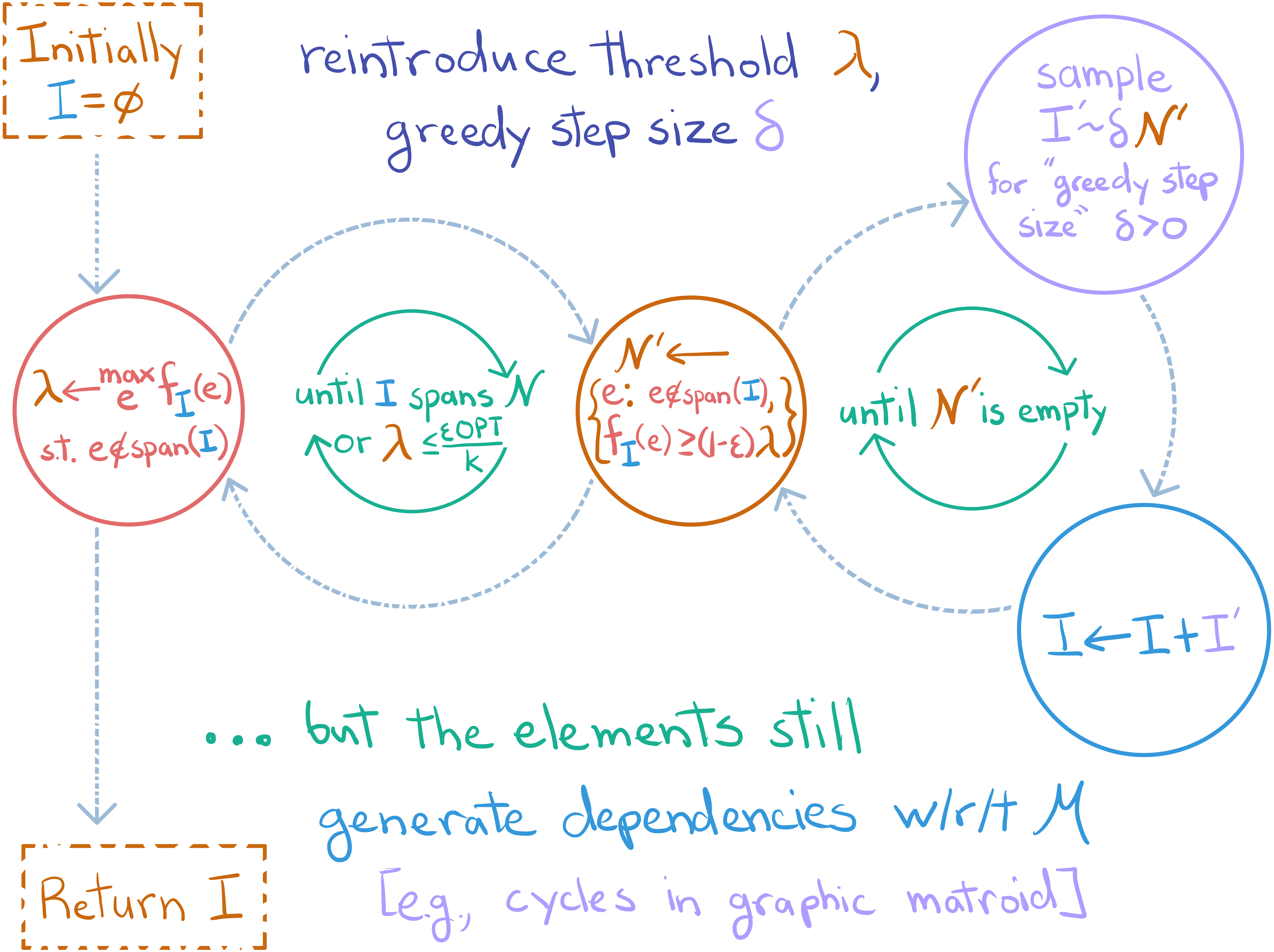
... but the elements:

(a) negate each other w/r/t f

(b) generate dependencies w/r/t $\mathcal{M}$

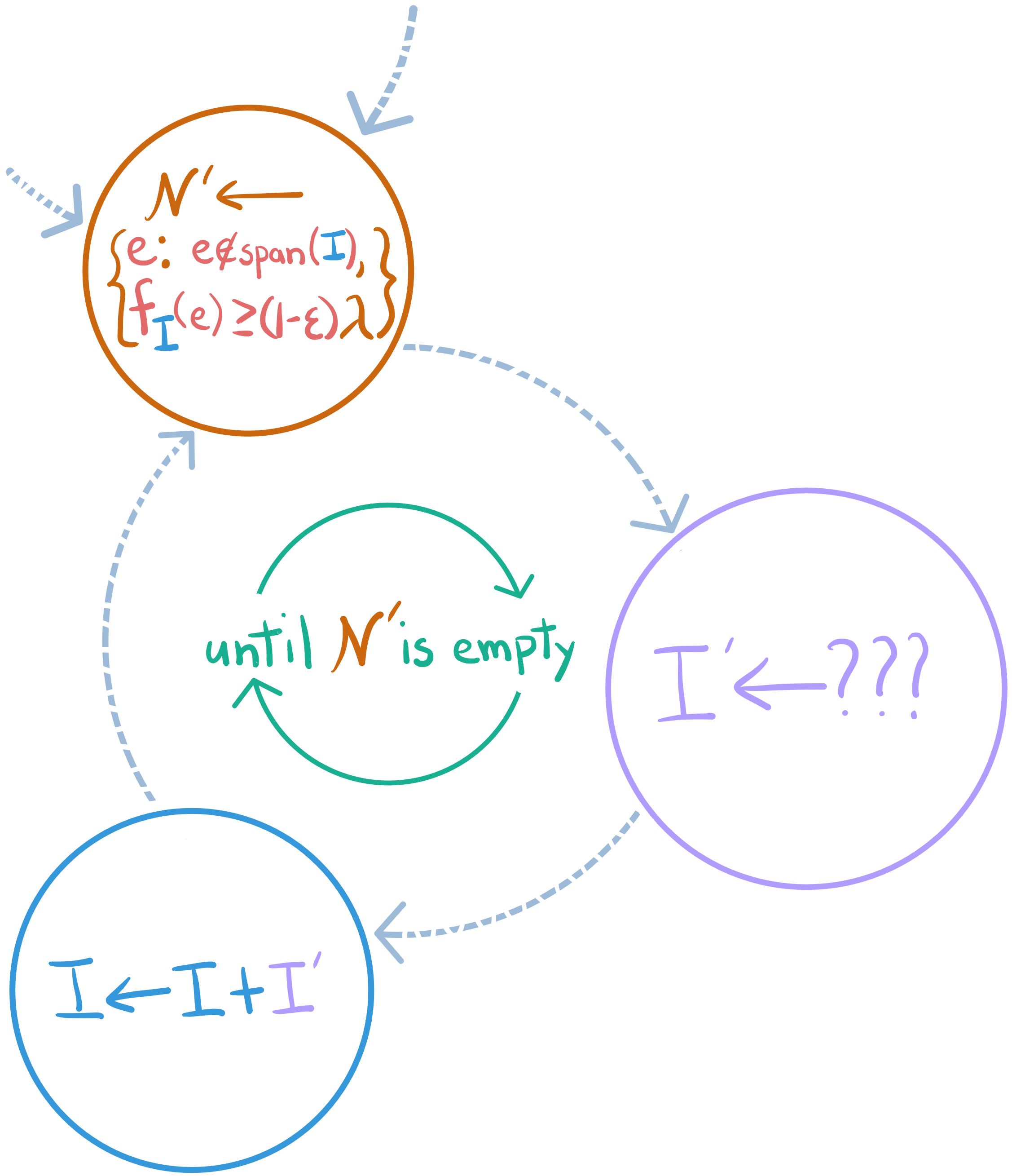
[e.g., cycles in graphic matroid]

reintroduce threshold λ, greedy step size δ



... but the elements still generate dependencies w/r/t $\mathcal{M}$

[e.g., cycles in graphic matroid]



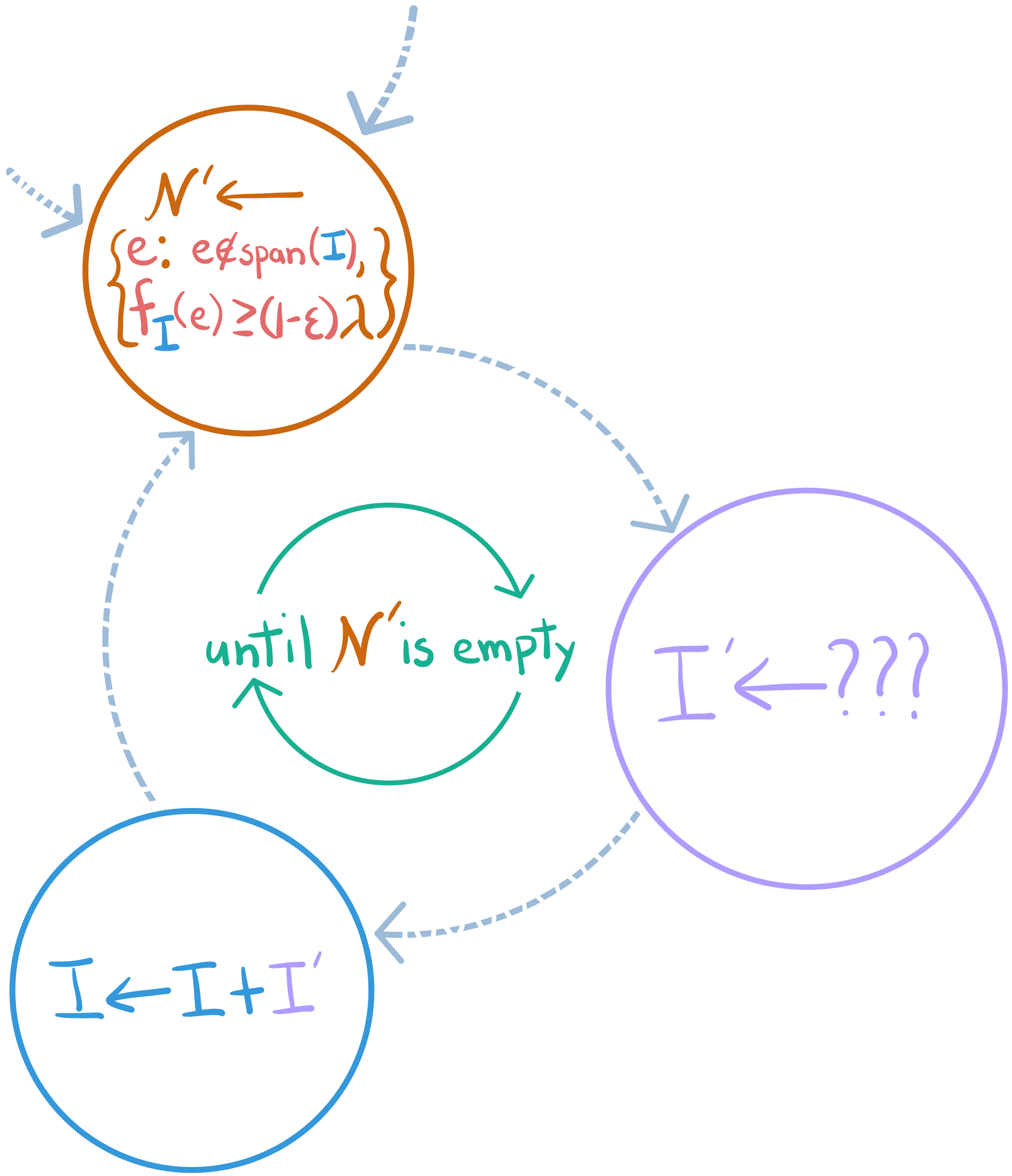
Goal

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 $f_I(I') \simeq \lambda|I'|$

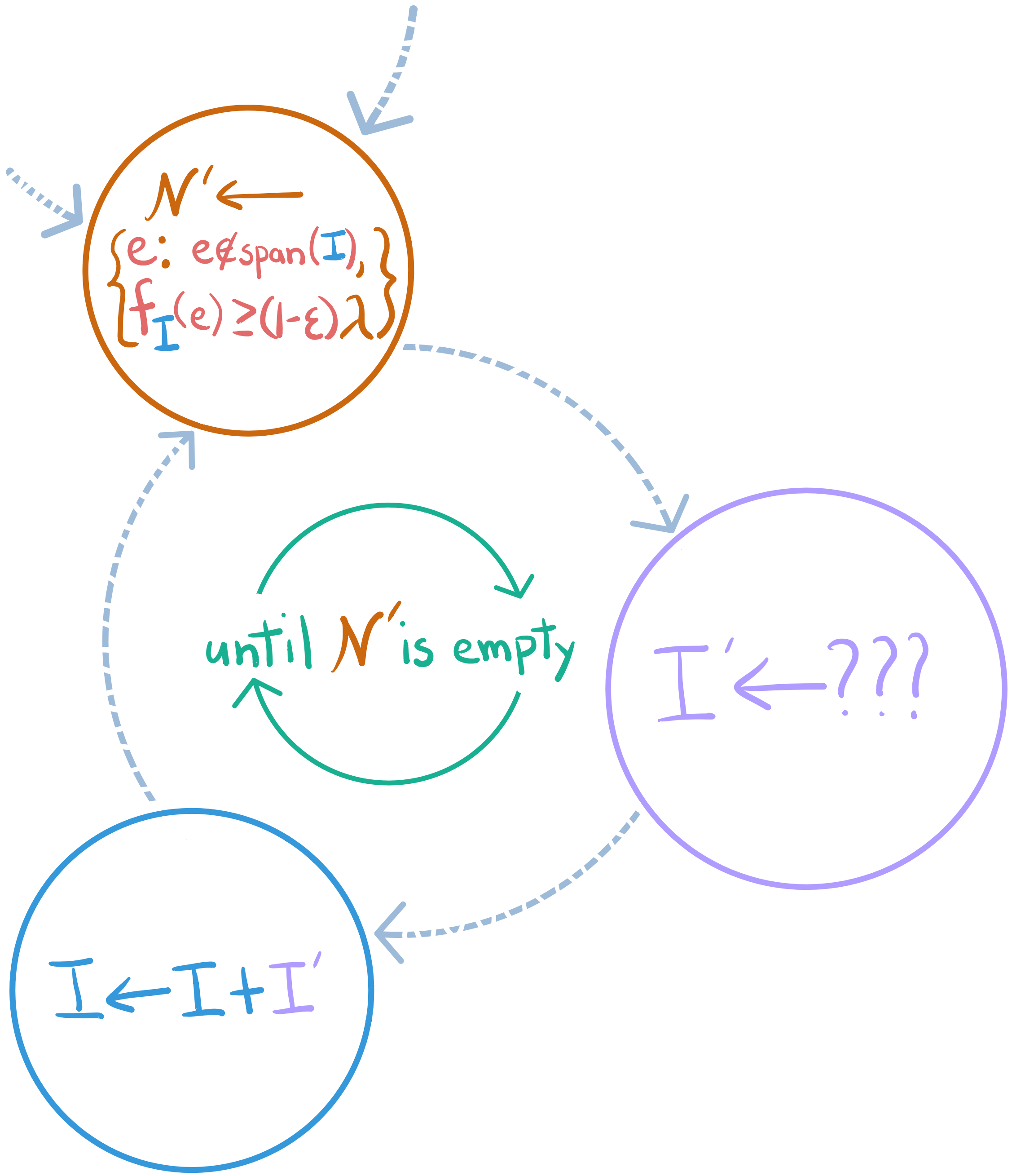
(b) $|N'|$ ⬇ a lot w/r/t $I+I'$

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.

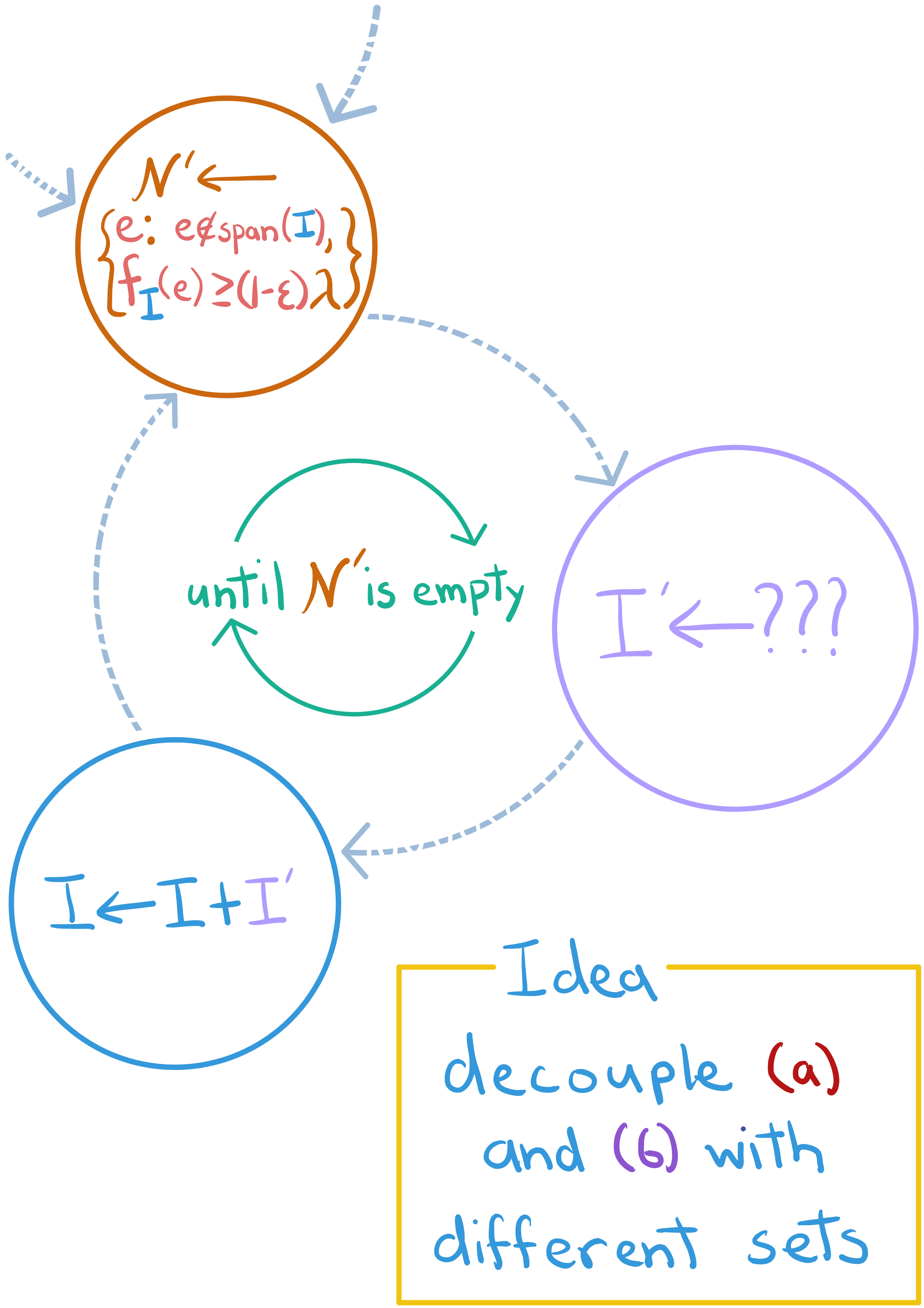
(a) $I + I'$ is independent and $f_I(I') \simeq \lambda |I'|$

(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) $\not\Leftrightarrow$ (b)



Goal

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

(a) $I + I'$ is independent

and $f_I(I') \simeq \lambda |I'|$

(b) $|N'|$ ⬇ a lot w/r/t $I + I'$

by combination of:

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

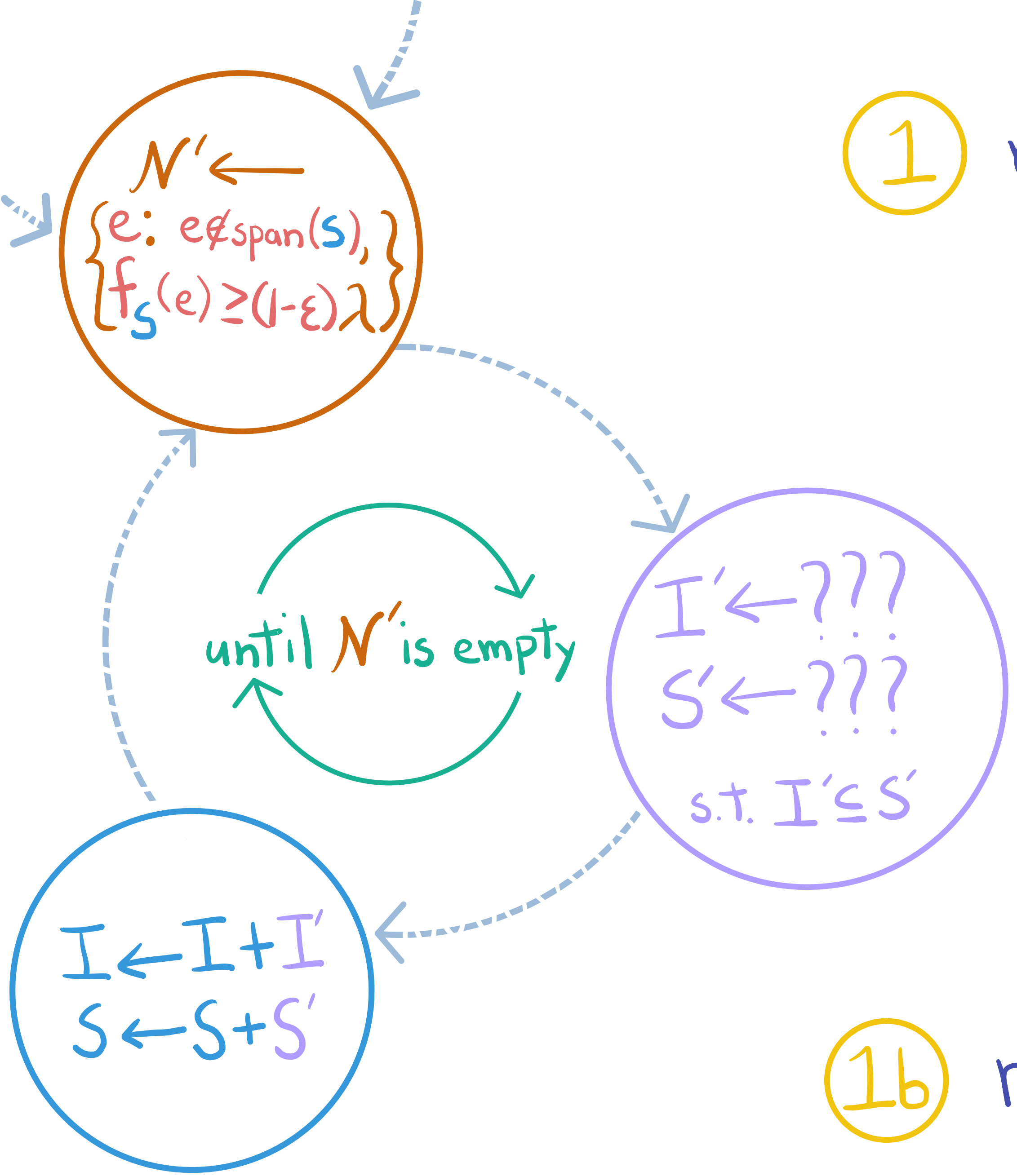
(ii) $f_{I+I'}(e) \ll f_I(e)$

for many $e \in N'$

Idea

decouple (a) and (b) with different sets

But (a) =X= (b)



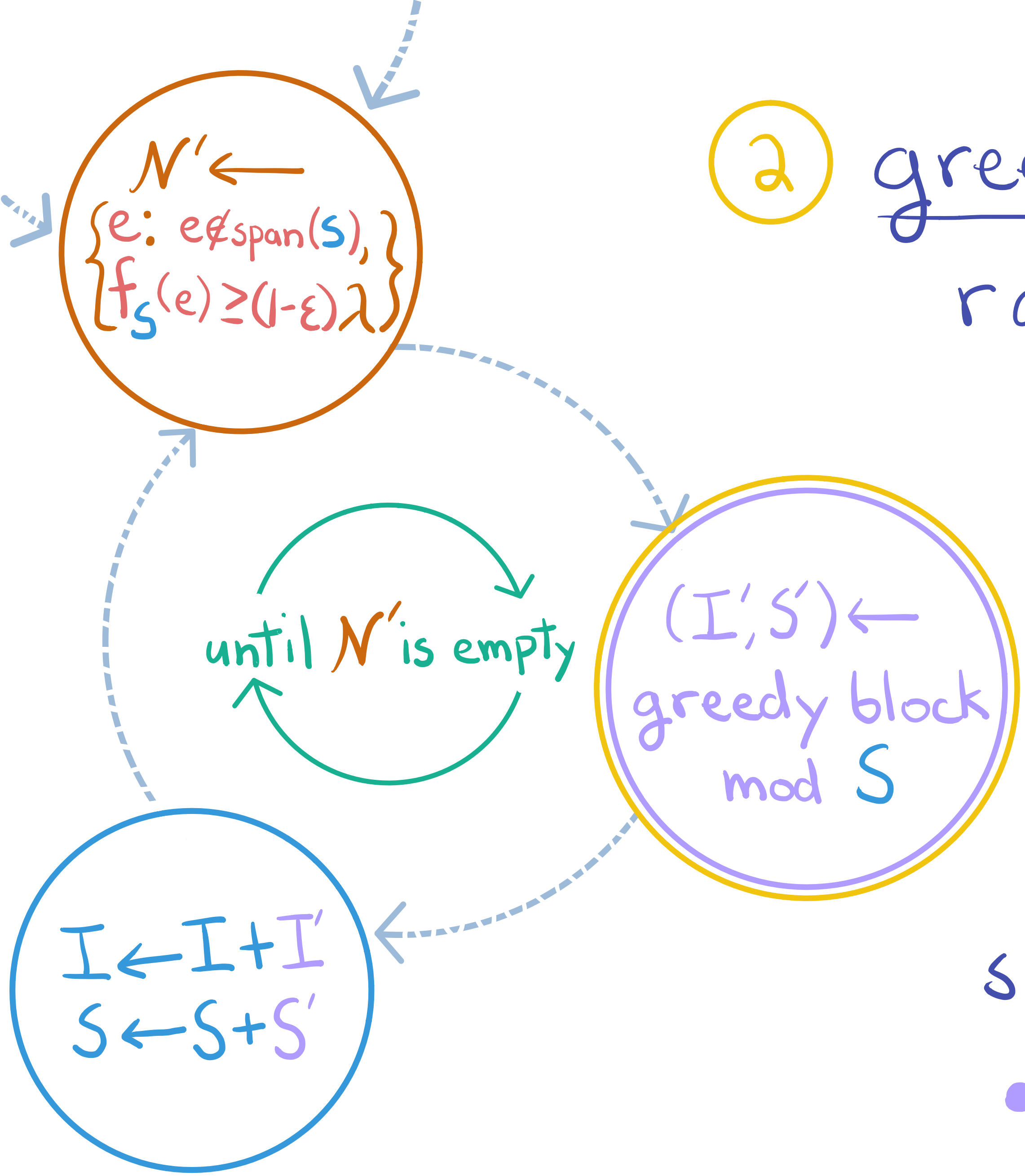
① replace I' with 2 sets

I' and S'

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

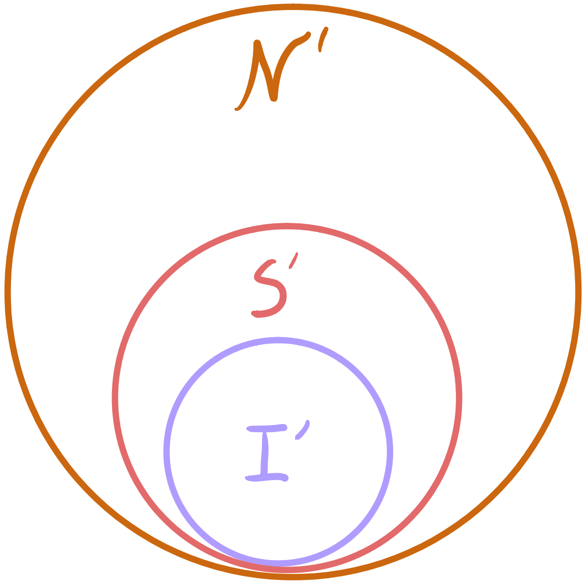
(1b) replace I w/ I and S

(for sake of induction)



② greedy block (I', S') mod S

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

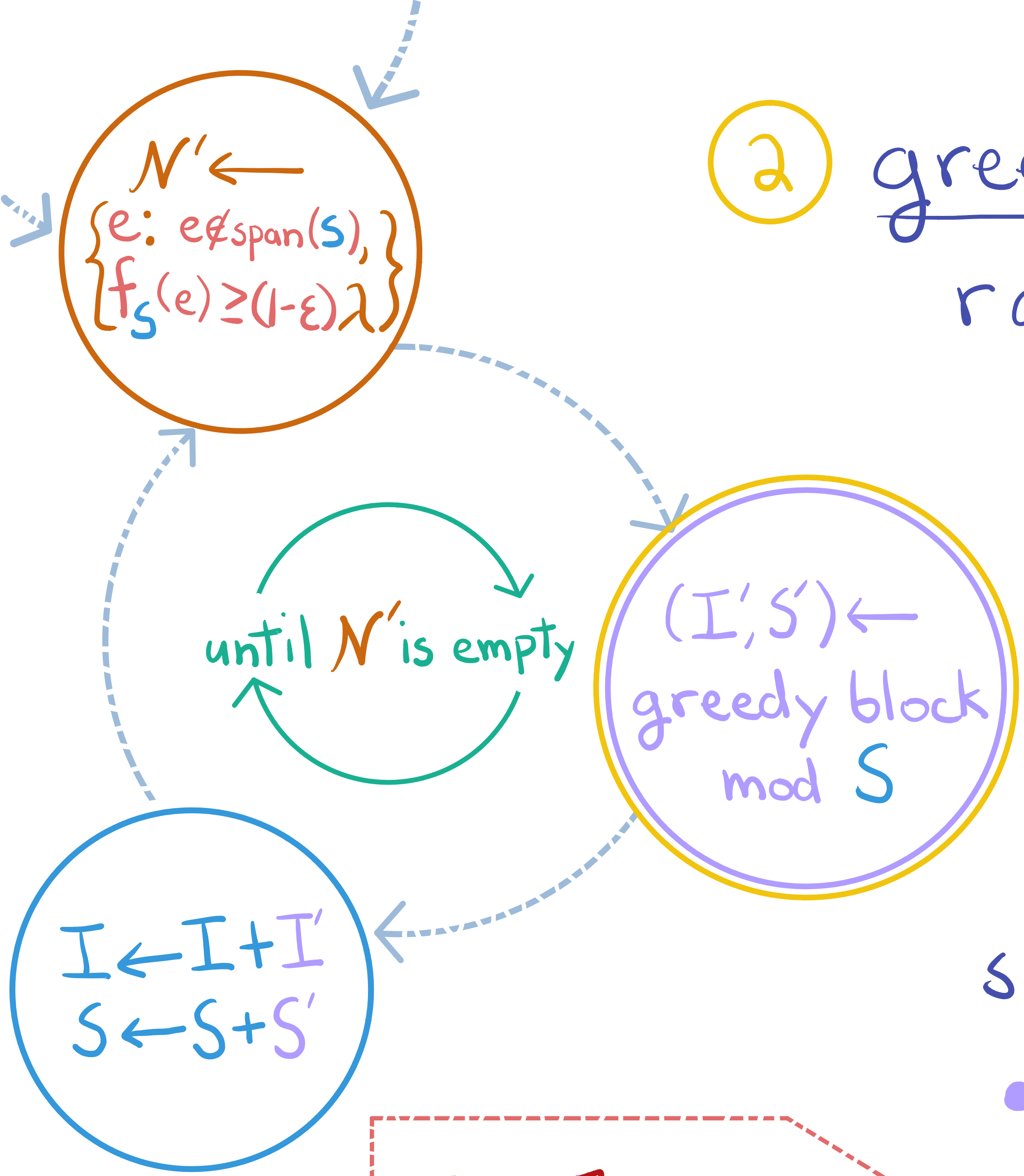


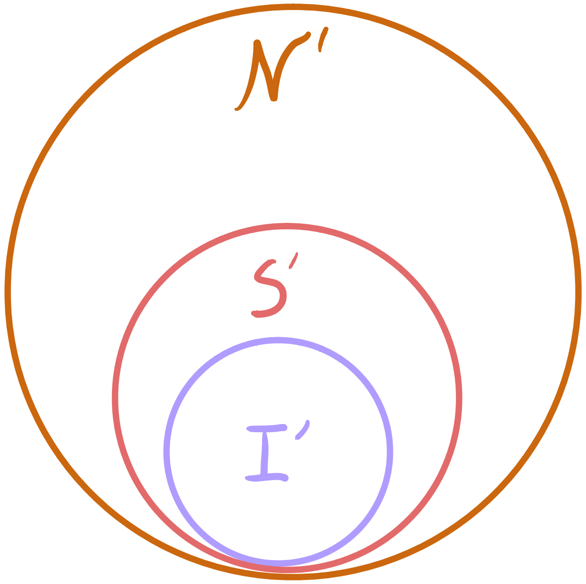
such that:

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

② greedy block (I', S') mod S

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

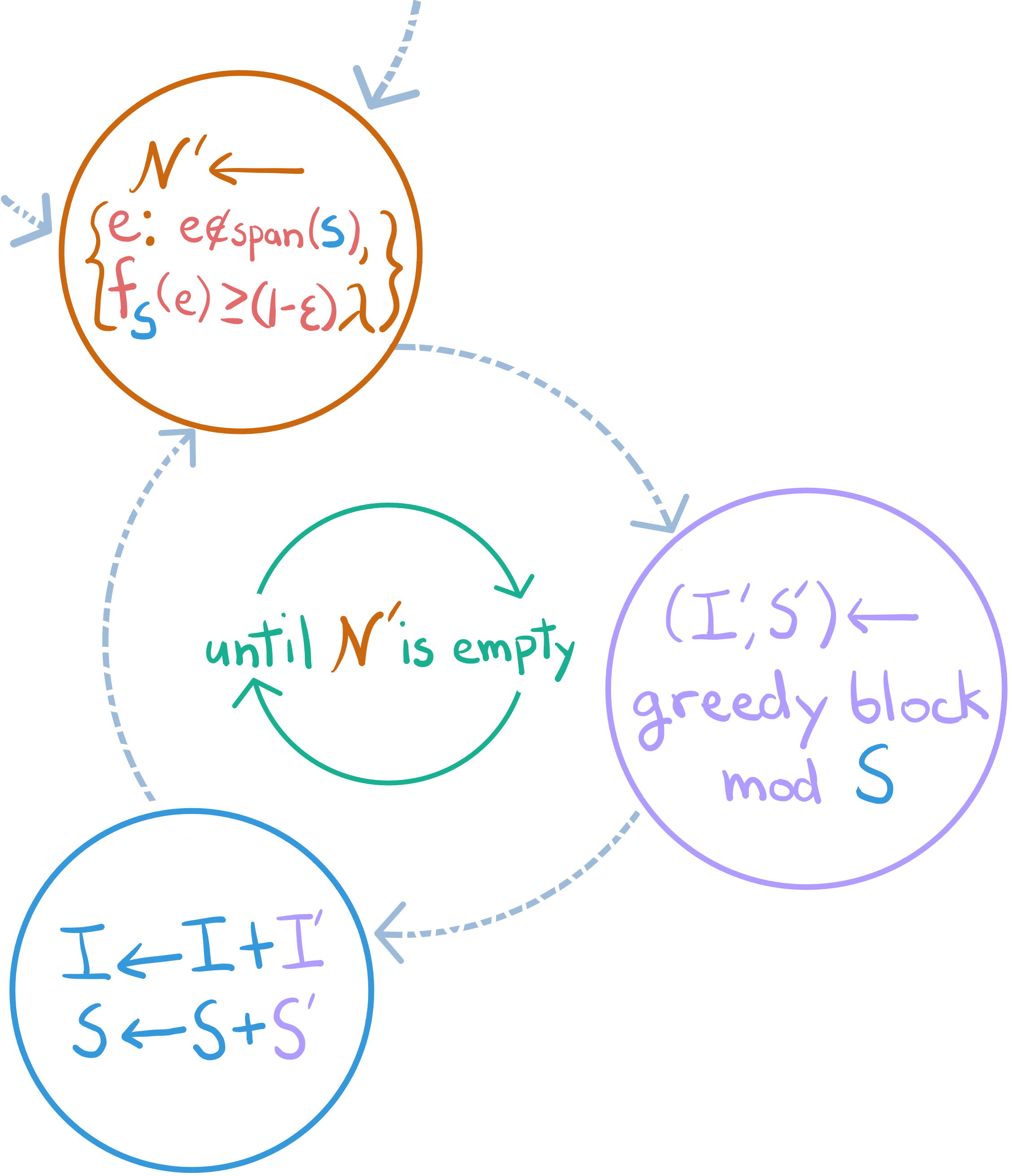




such that:

- $I + I'$ independent
- $$E[\underbrace{f_S(I')}_{\text{bang}}] \geq (1-\varepsilon) E[\underbrace{|S'|}_{\text{buck}}] \lambda$$

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



Need to show

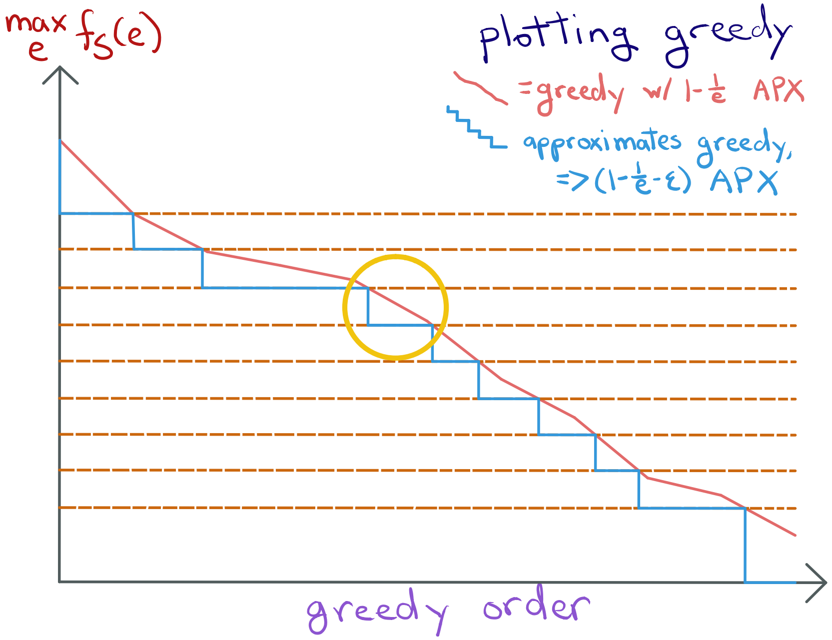
ⓐ APX factor

ⓑ 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 $= S$)

$\Rightarrow$  $\Rightarrow$

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

- for all $J \in \mathcal{I}$,

  $E[f(I)] \geq (1-\varepsilon)E[f_S(J)]$

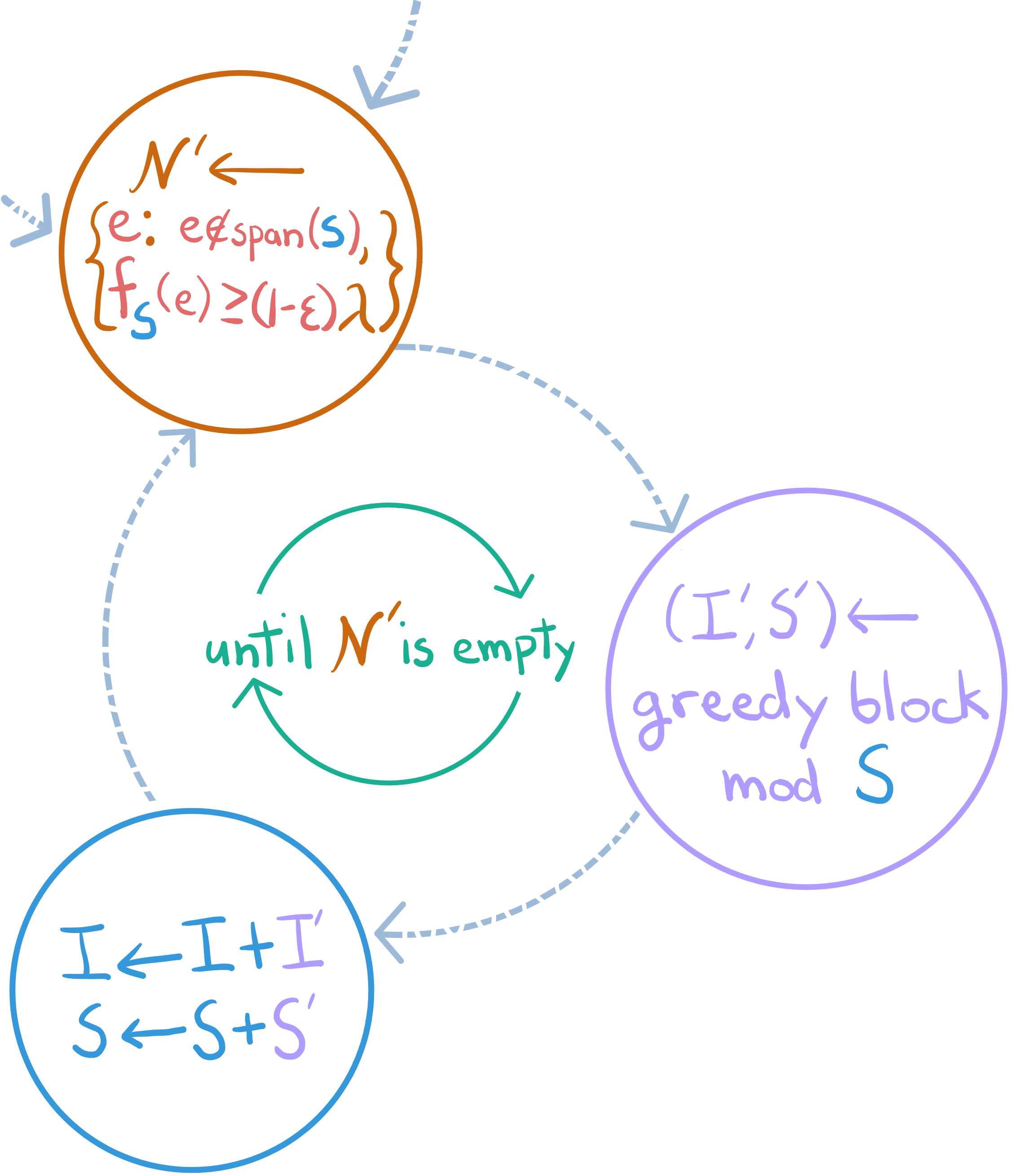
- $E[f(I)] \geq (1-\varepsilon)E[f(S)]$

[omit proof]

Corollary

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

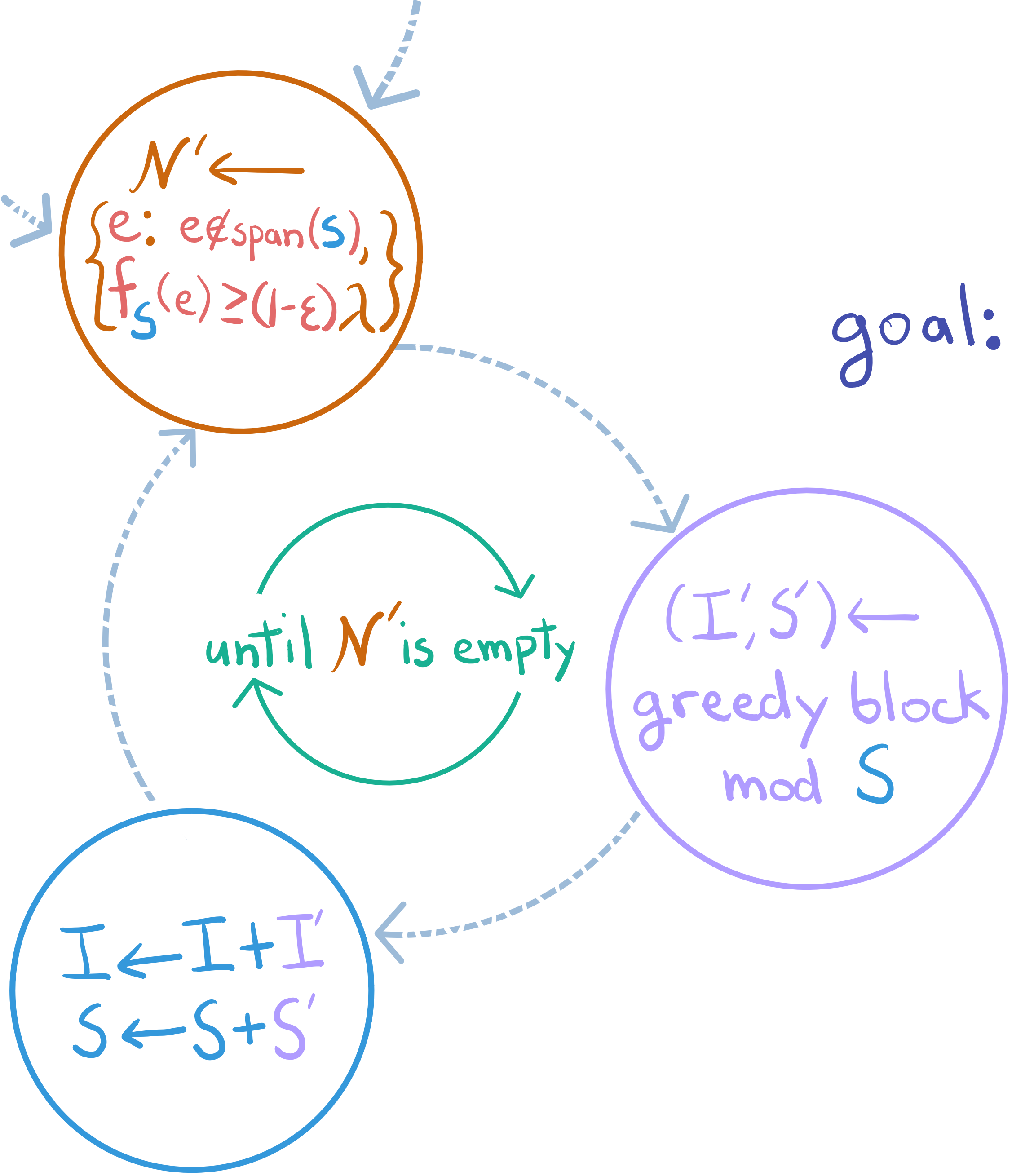
constant APX for nonnegative f



Need to show

ⓐ ~~APX factor~~

ⓑ how? and depth



"greedy sampling"

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

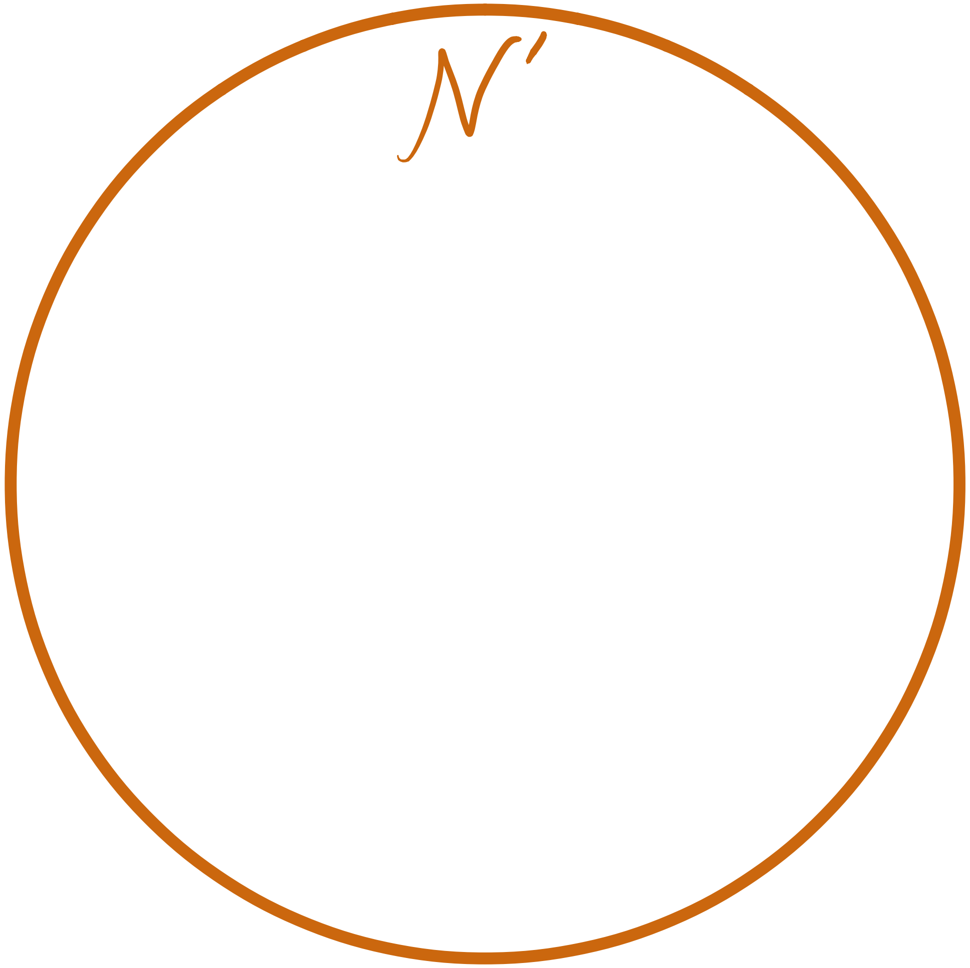
Need to show

ⓐ ~~APX factor~~

ⓑ how? and depth

"greedy sampling"

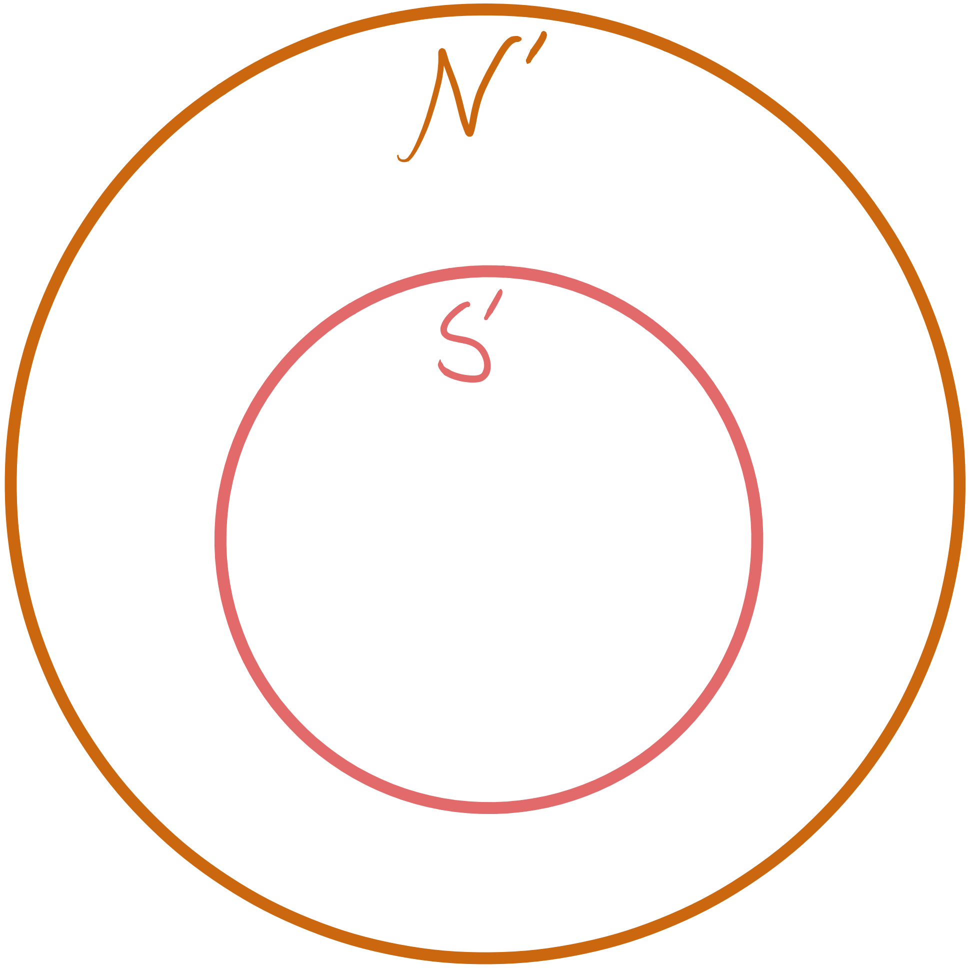
let $\delta > 0$ TBD



"greedy sampling"

let $\delta > 0$ TBD

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

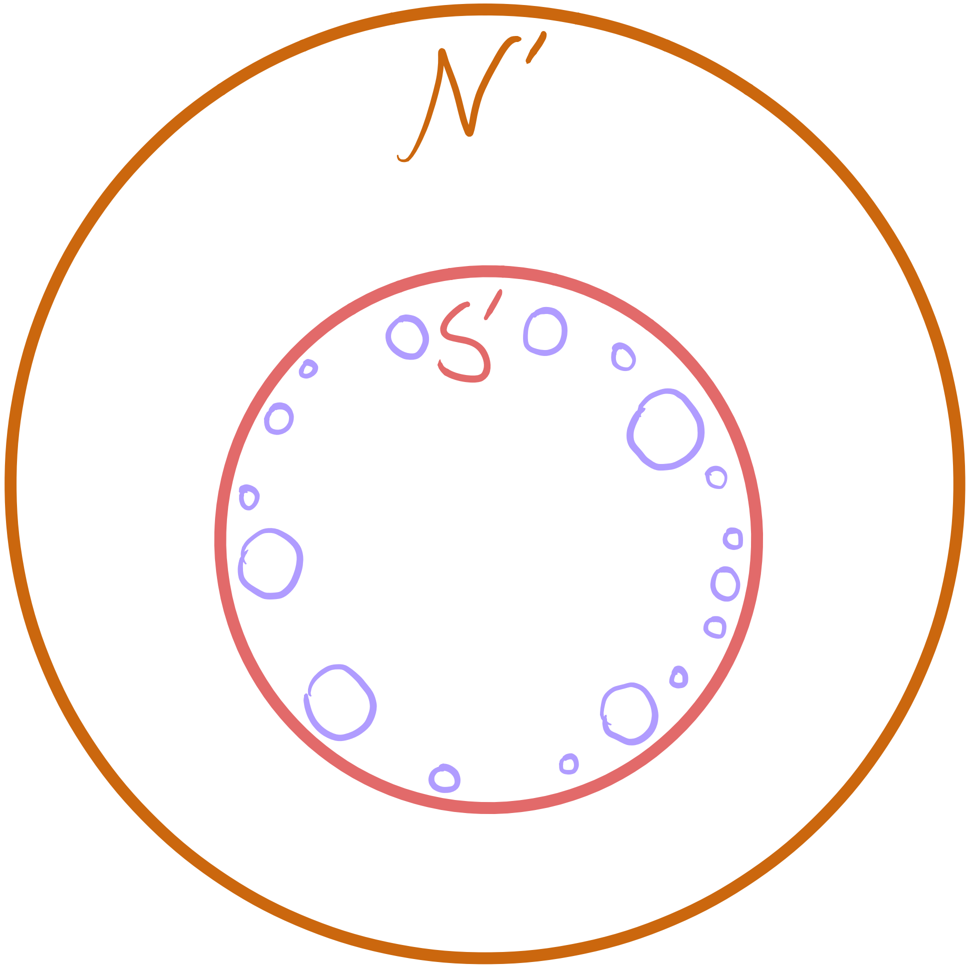


"greedy sampling"

let $\delta > 0$ TBD

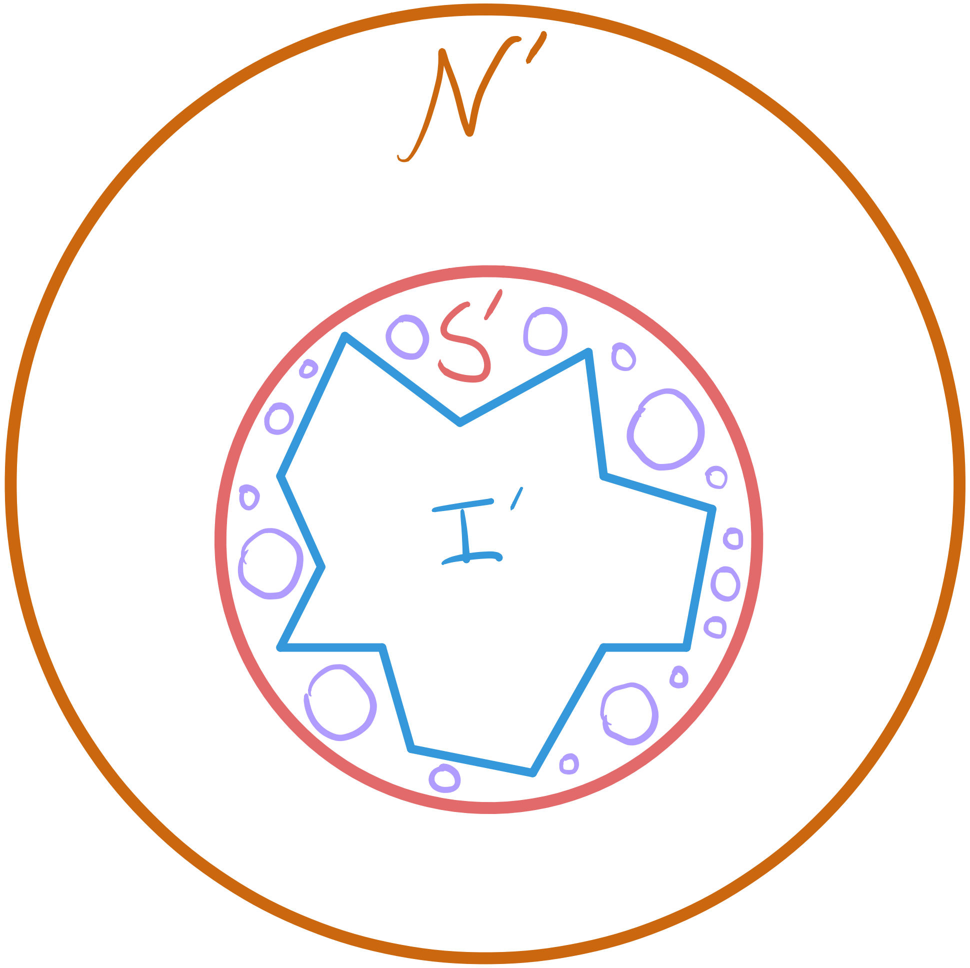
(a) sample $S' \sim \delta N'$

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



"greedy sampling"

let $\delta > 0$ TBD



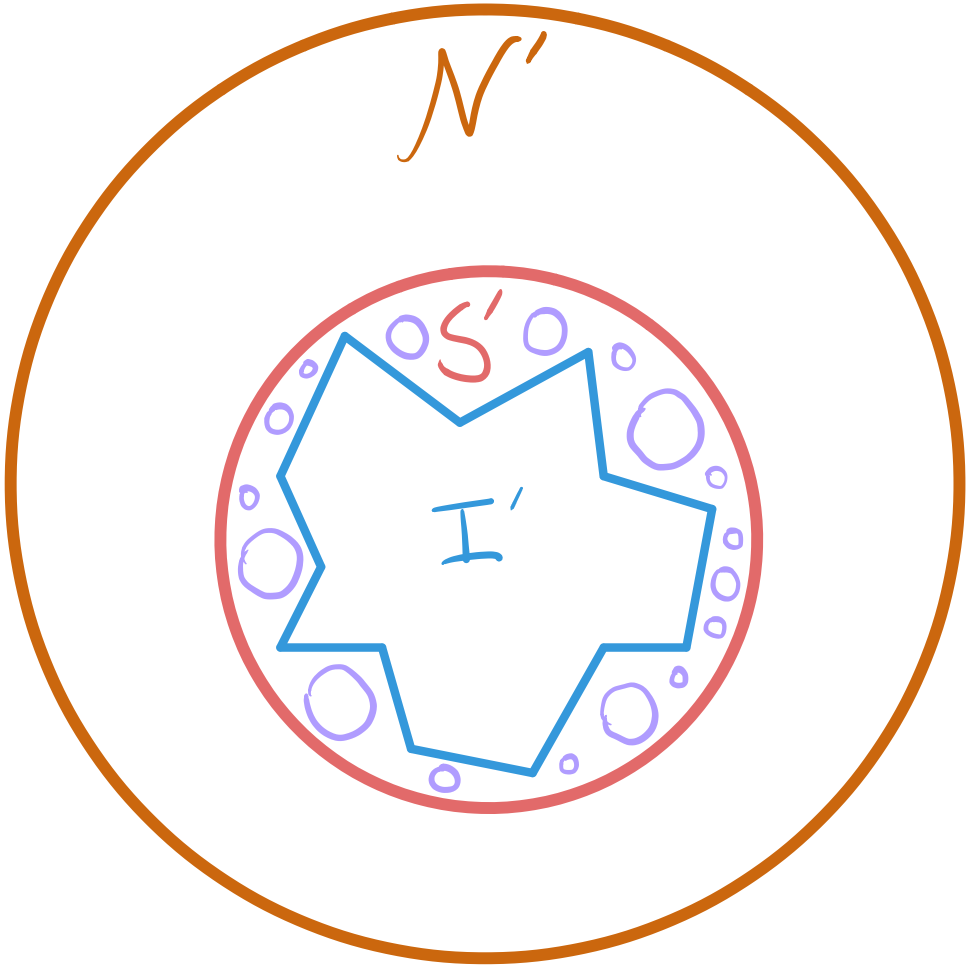
(a) sample $S' \sim_{\delta} \mathcal{N}'$

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

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

"greedy sampling"

let $\delta > 0$ TBD



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

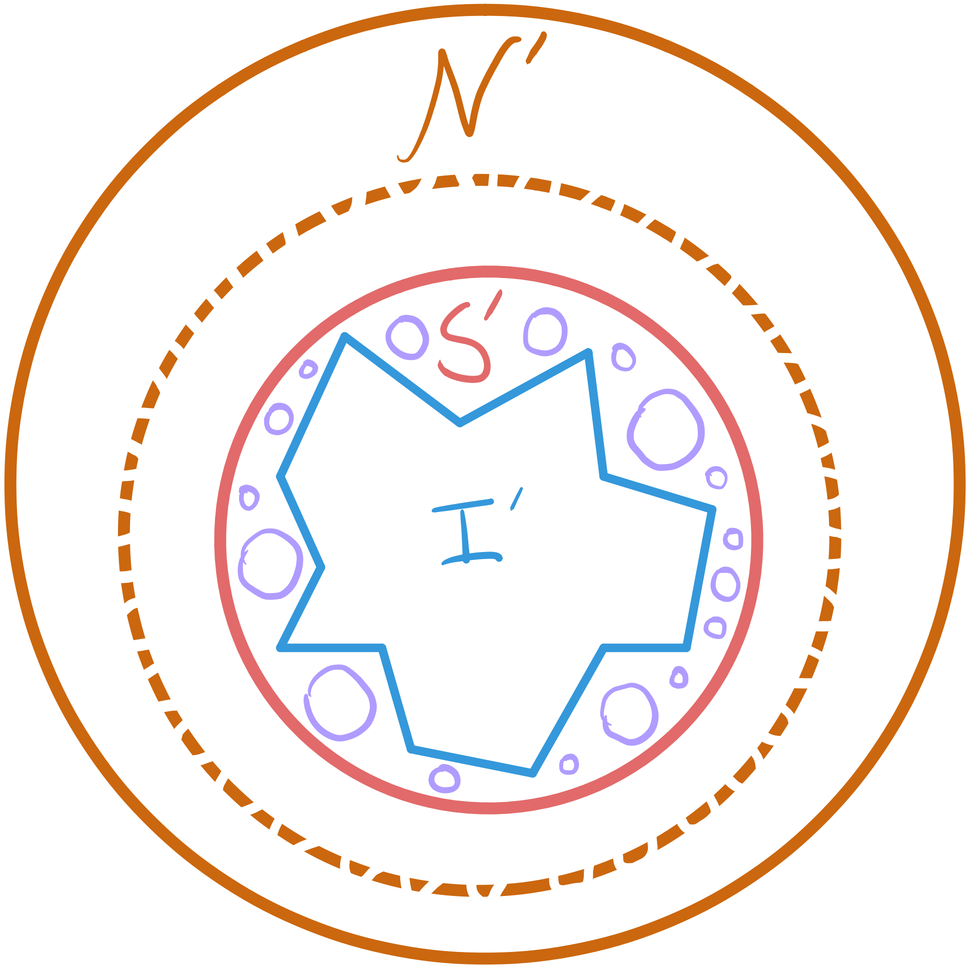
[S' may have dependencies, or bad marginal values]

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

← in parallel w/ independent queries to f and span oracles

"greedy sampling"

let $\delta > 0$ TBD

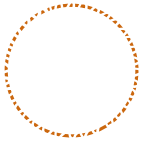


(a) sample $S' \sim_{\delta} N'$

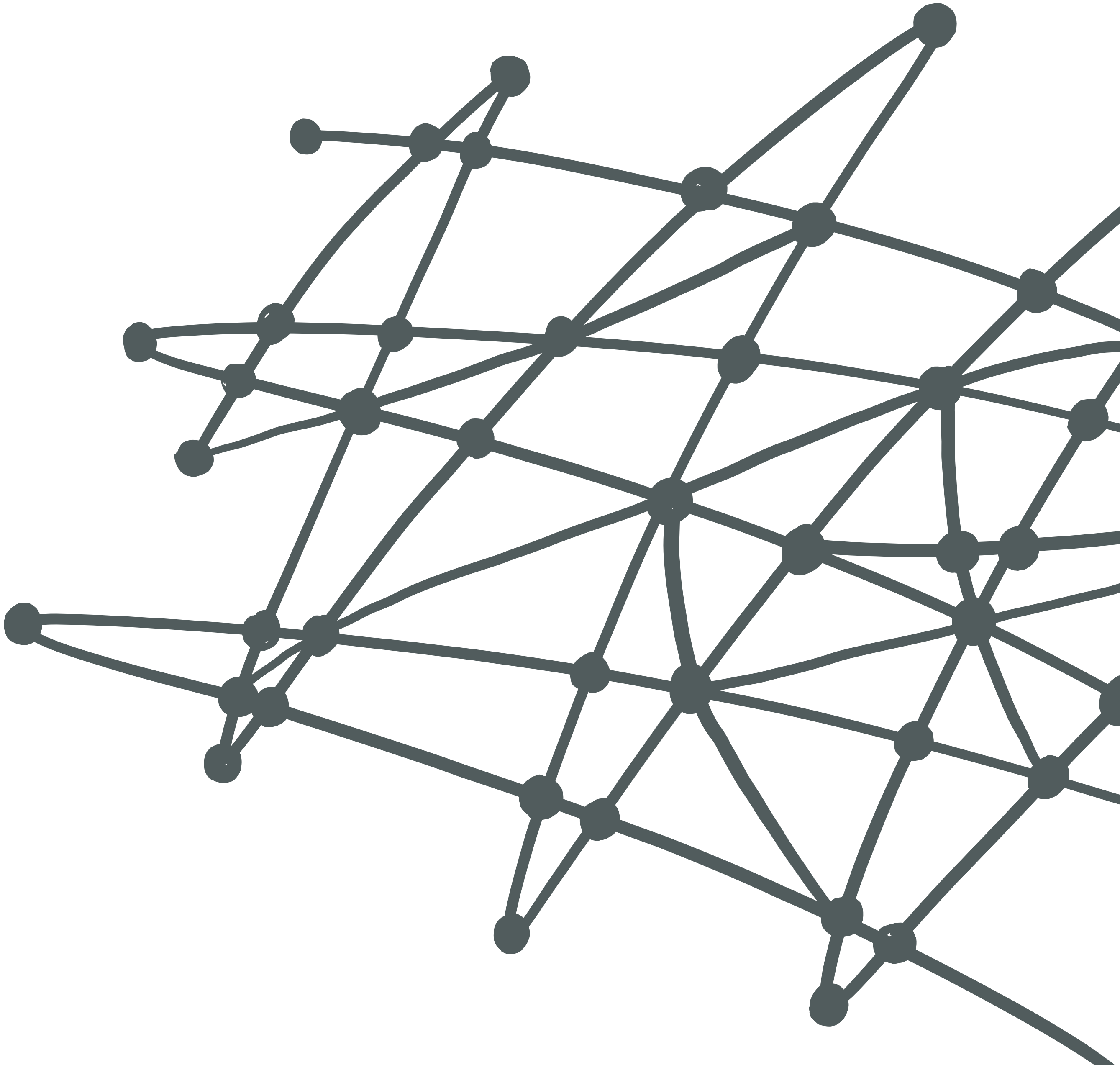
[S' may have dependencies, or bad marginal values]

(b) $I' \leftarrow \left\{ e \in S' : e \notin \mathrm{span}(S+S'-e) \text{ and } f_{S+S'-e}(e) \geq (1-\varepsilon)\lambda \right\}$ ← in parallel w/ independent queries to f and span oracles

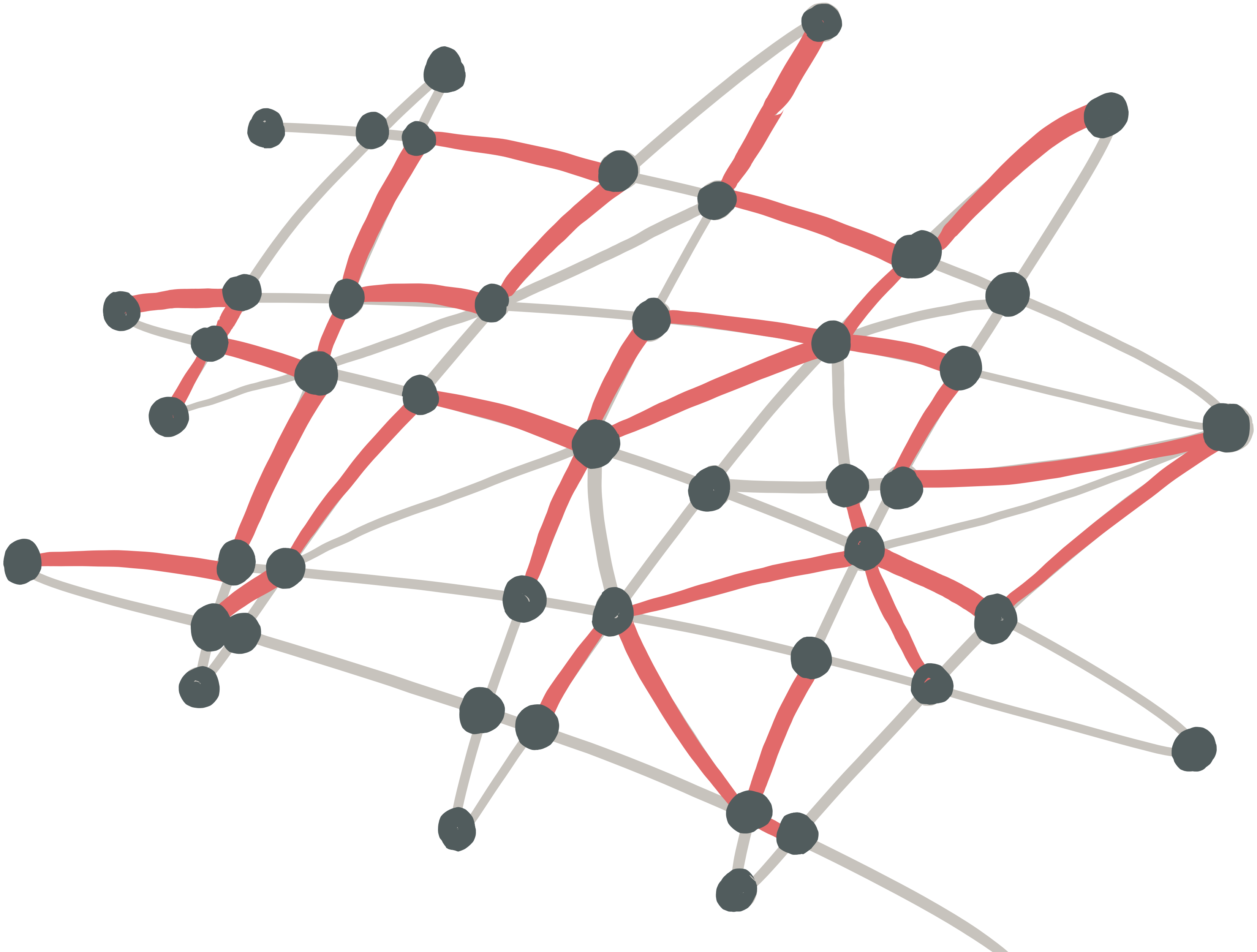
(and then N' shrinks)



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

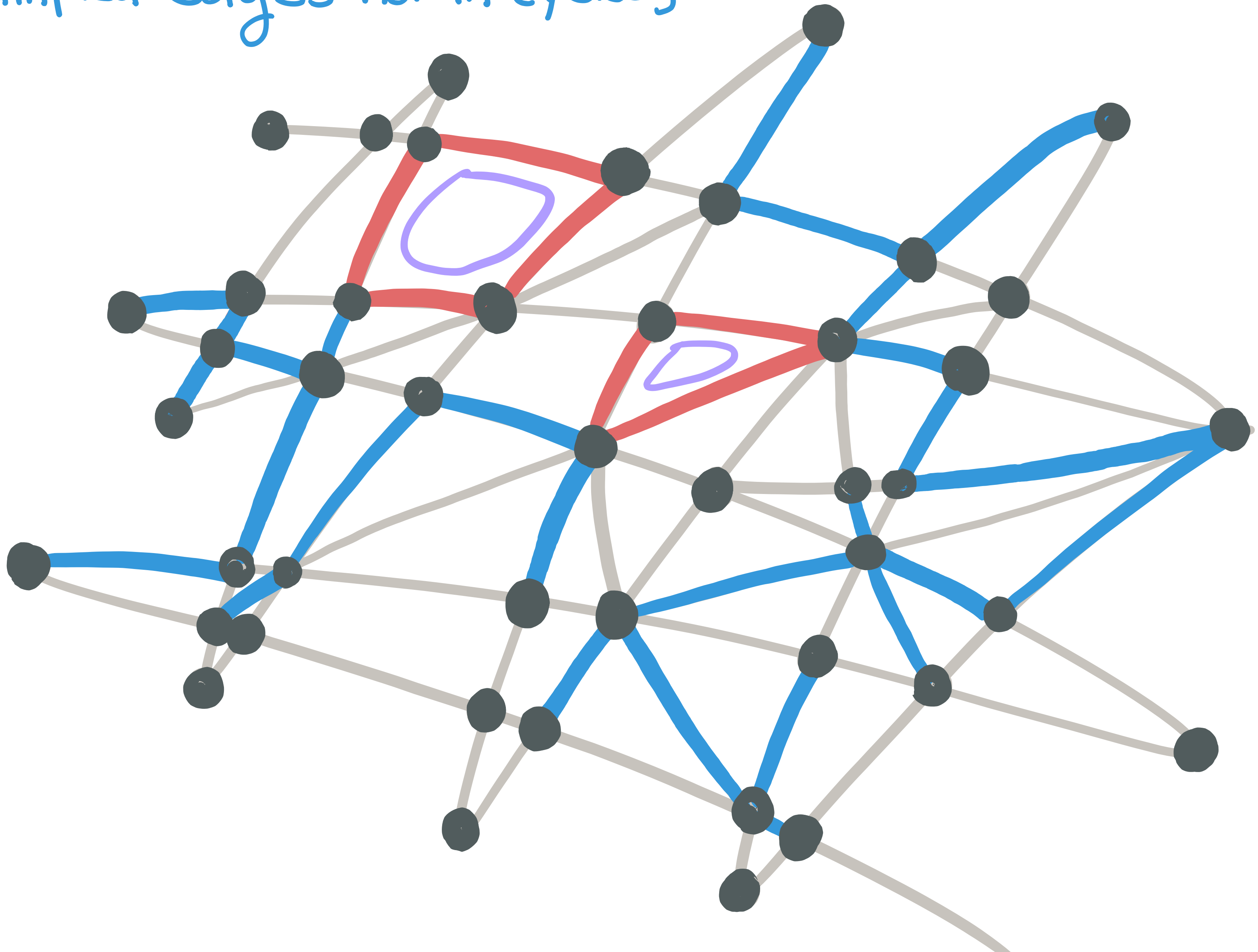


① let S′ sample each edge w/ prob δ



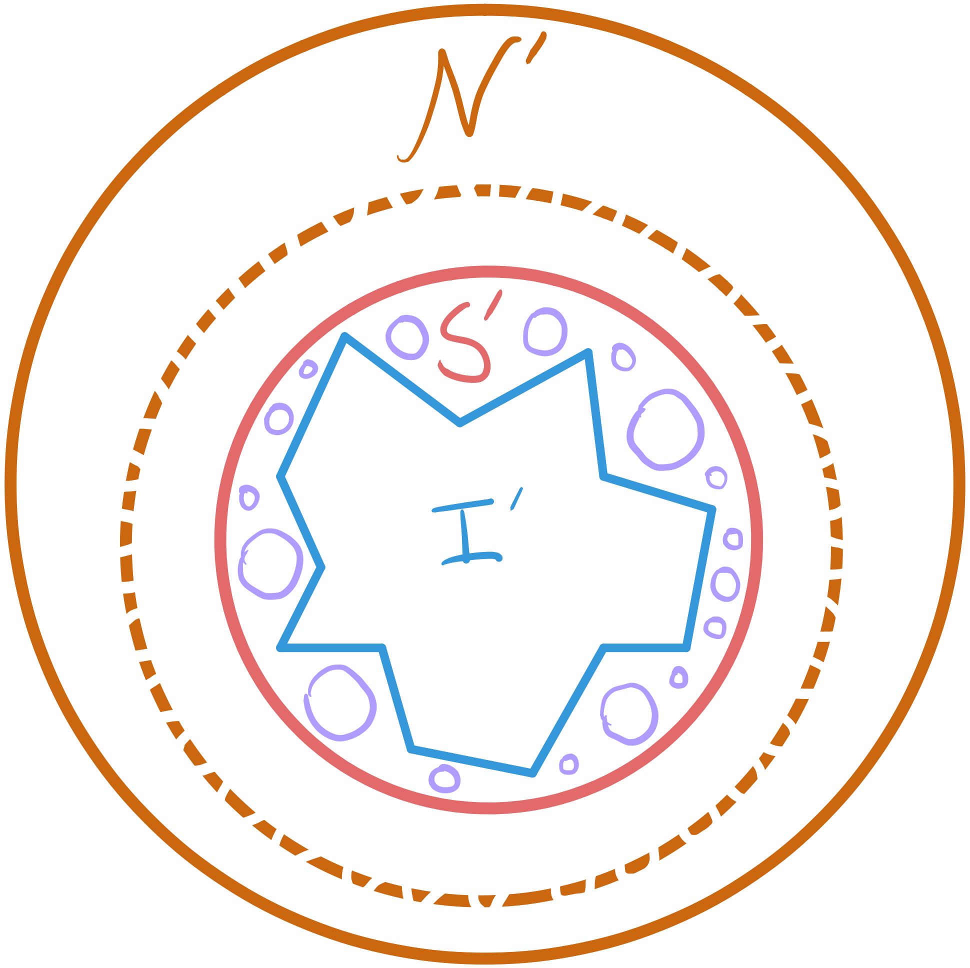
① let S' sample each edge w/ prob δ

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



"greedy sampling"

let $\delta > 0$ TBD

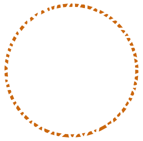


(a) sample $S' \sim_{\delta} N'$

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

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

(and then N' shrinks)



"greedy sampling"

let $\delta > 0$ TBD

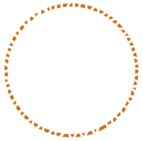
δ very very small

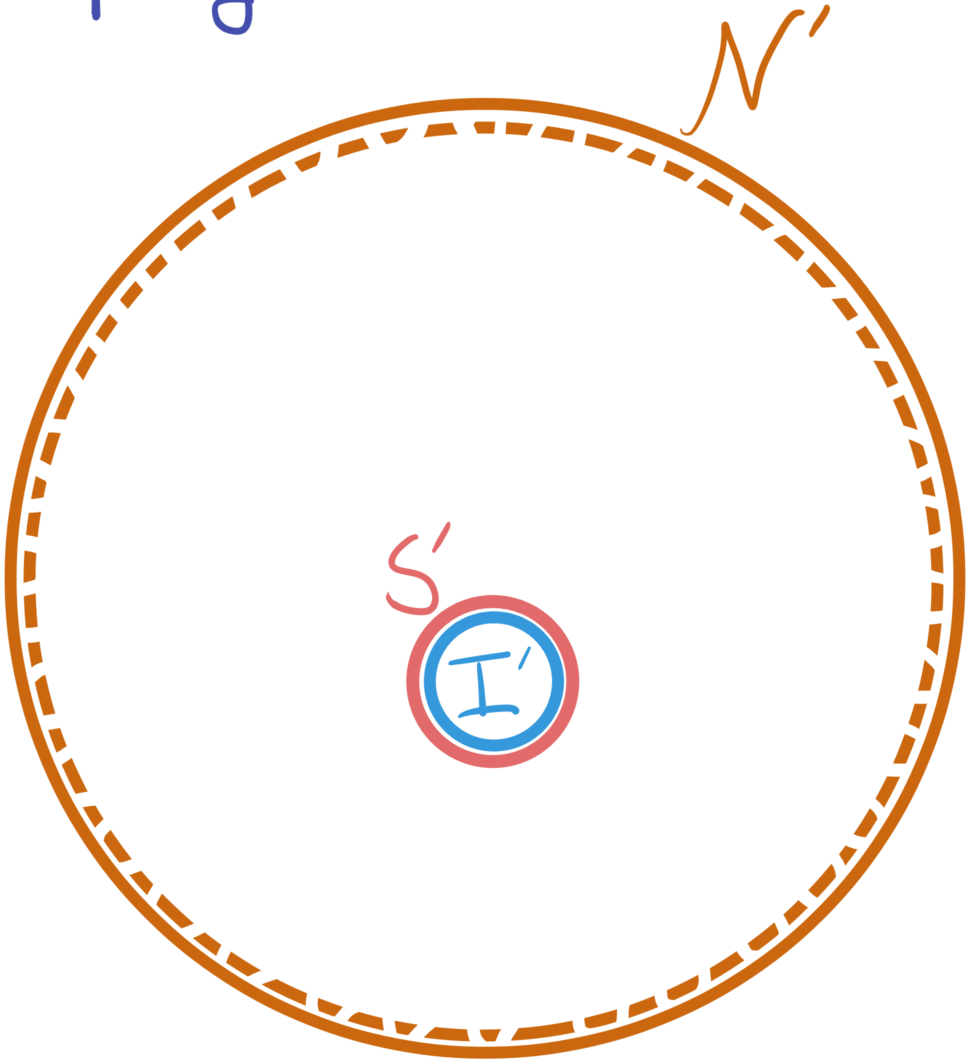
(a) sample $S' \sim_\delta N'$

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

(b) $I' \leftarrow \{ e \in S' : e \notin \mathrm{span}(S + S' - e)$ 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

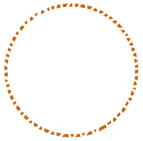
δ very very big

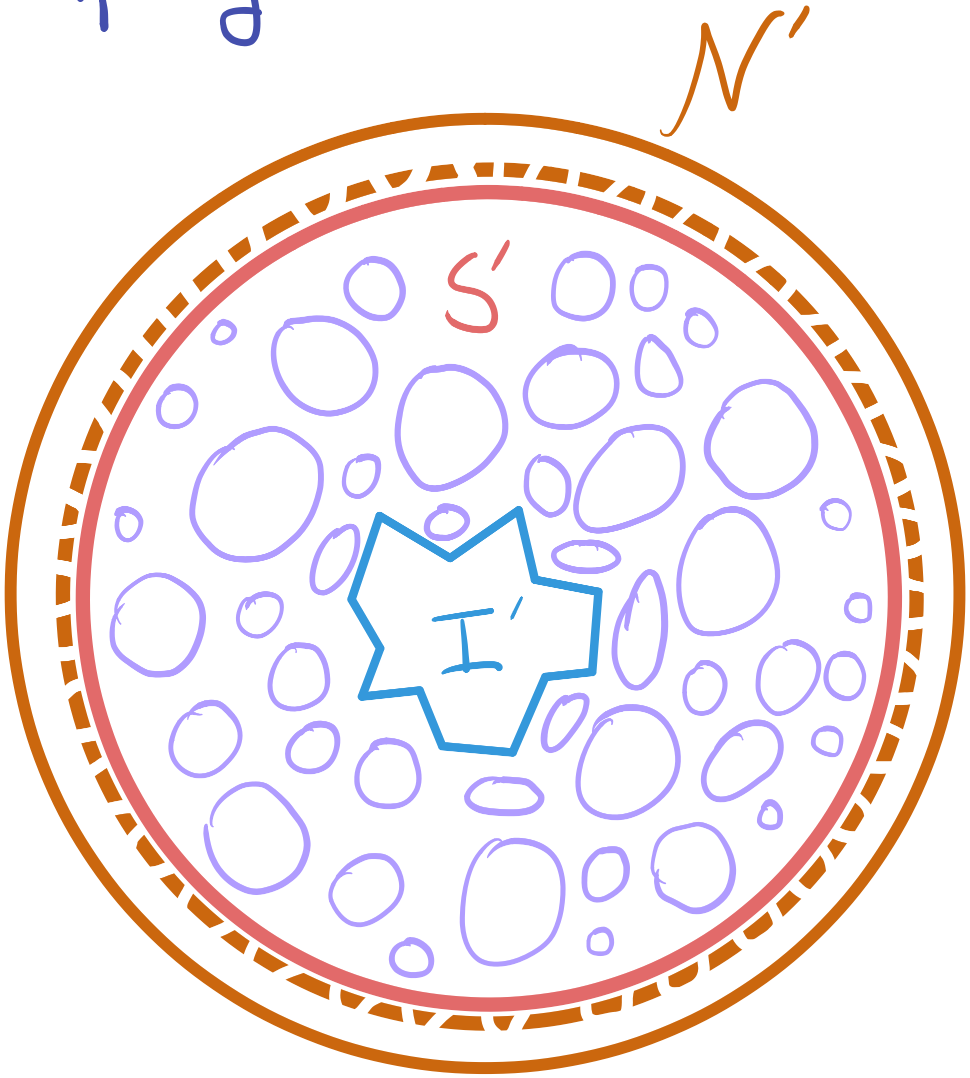
(a) sample $S' \sim \delta N'$

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

(b) $I' \leftarrow \{ e \in S' : e \notin \mathrm{span}(S + S' - e)$ 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"

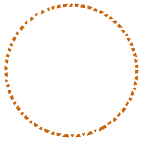
let $\delta > 0$ TBD

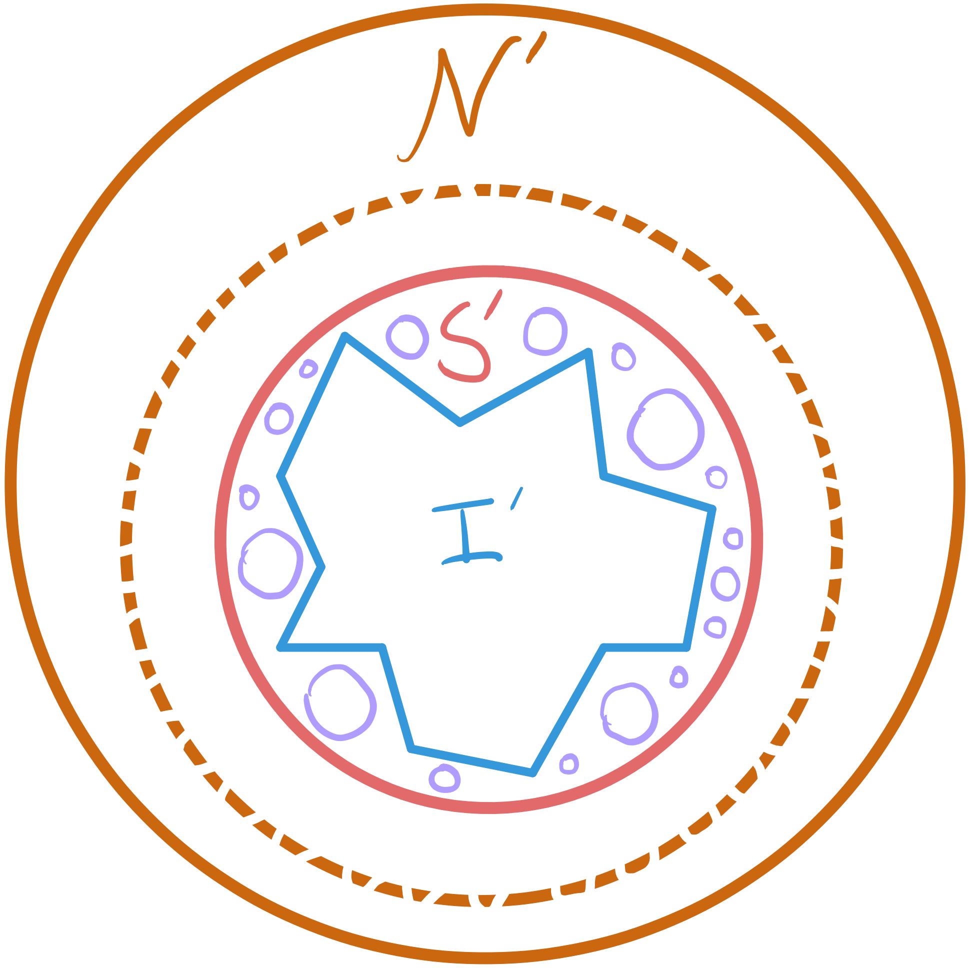
(a) sample $S' \sim_\delta \mathcal{N}'$

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

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

(and then $\mathcal{N}'$ shrinks)





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

---

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

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

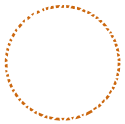
(b) $\mathbb{E}[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

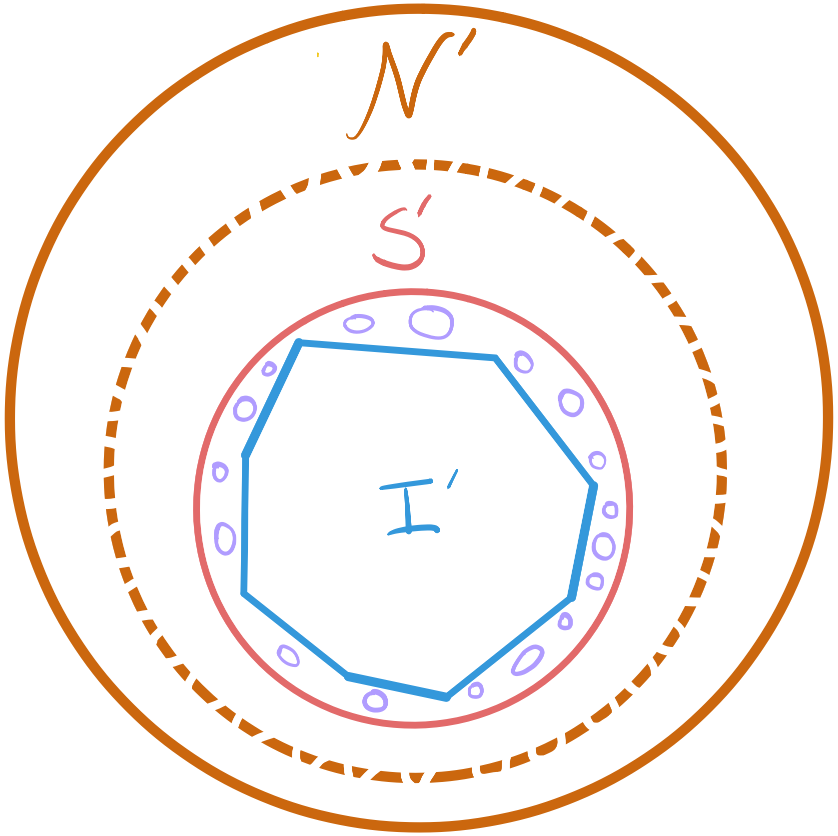
(a) sample $S' \sim \delta N'$

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

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

(and then N' shrinks)





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

---

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

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

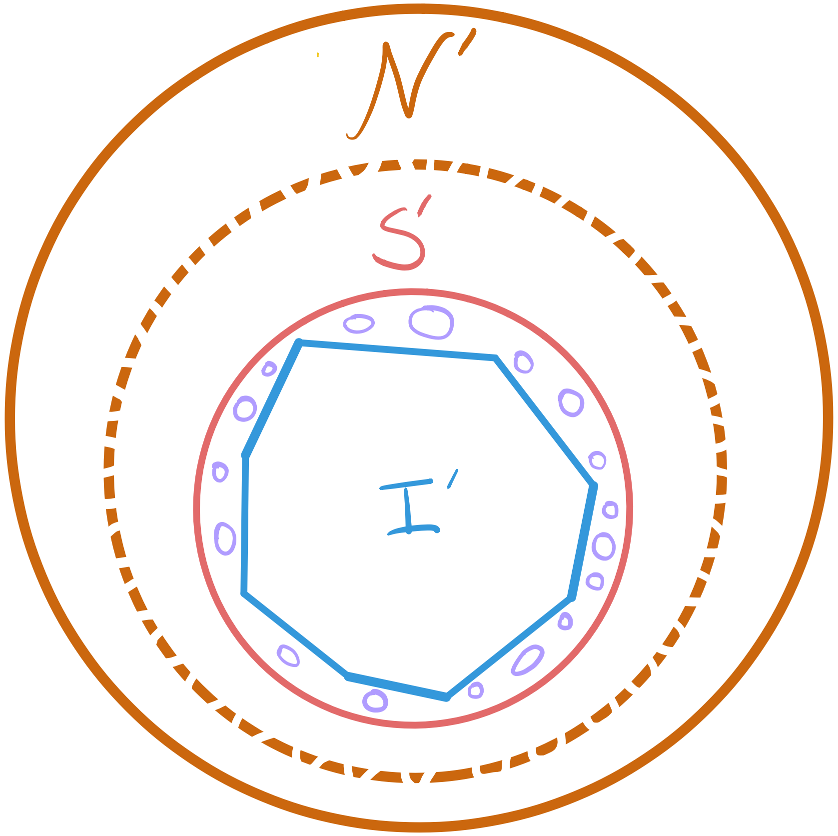
(b) $E[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

Key Claims

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

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

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



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

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

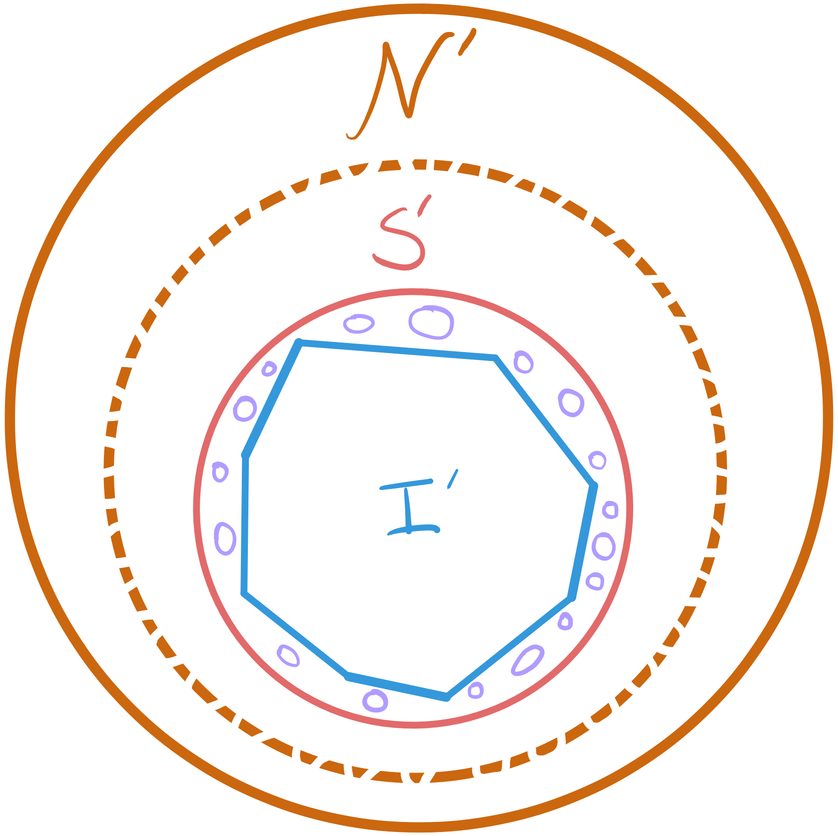
(b) $E[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \leq \varepsilon |N'|$

Key Claims

(1) ✓ (I', S') is greedy block mod S

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

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



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

---

for $S' \sim \delta S$:

both

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

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

---

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

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

---

for $S' \sim \delta S$:

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

~~both~~ either

(b) $E[|\{e \in N' : e \notin \operatorname{span}(S+S')\}|] \overset{\geq}{\cancel{\leq}} \varepsilon |N'|$

---

both $E[|\{\cdots\}|]$ terms ⬆ as δ ⬆

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

---

for $S' \sim \delta S$:

~~both~~ either

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

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

---

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

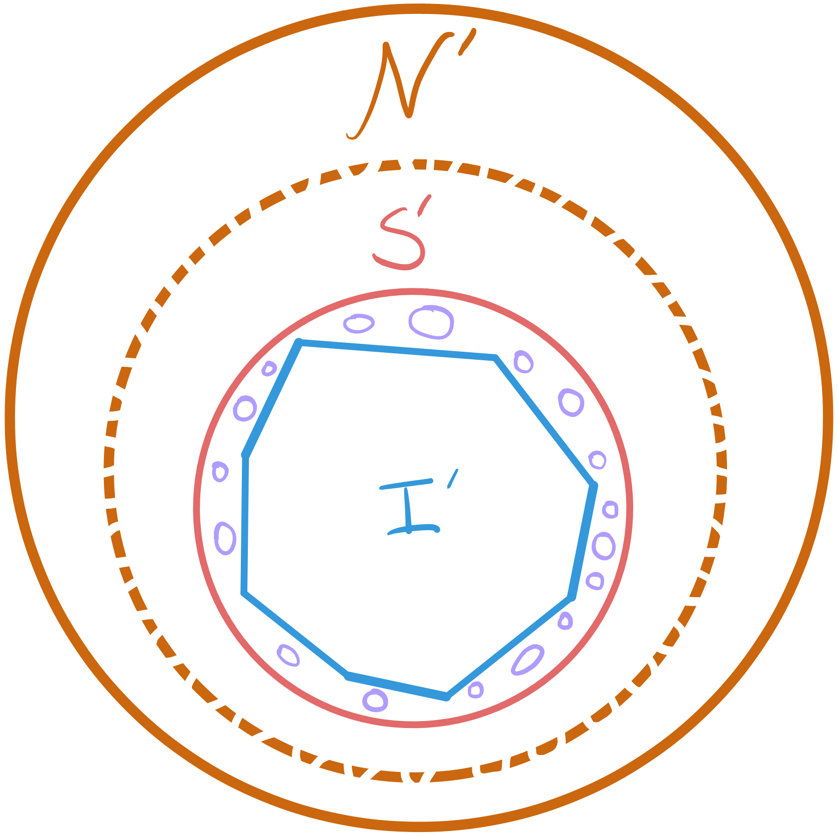
(b) $E[|\{e\in N': e\notin \mathrm{span}(S+S')\}|]\le\varepsilon|N'|$

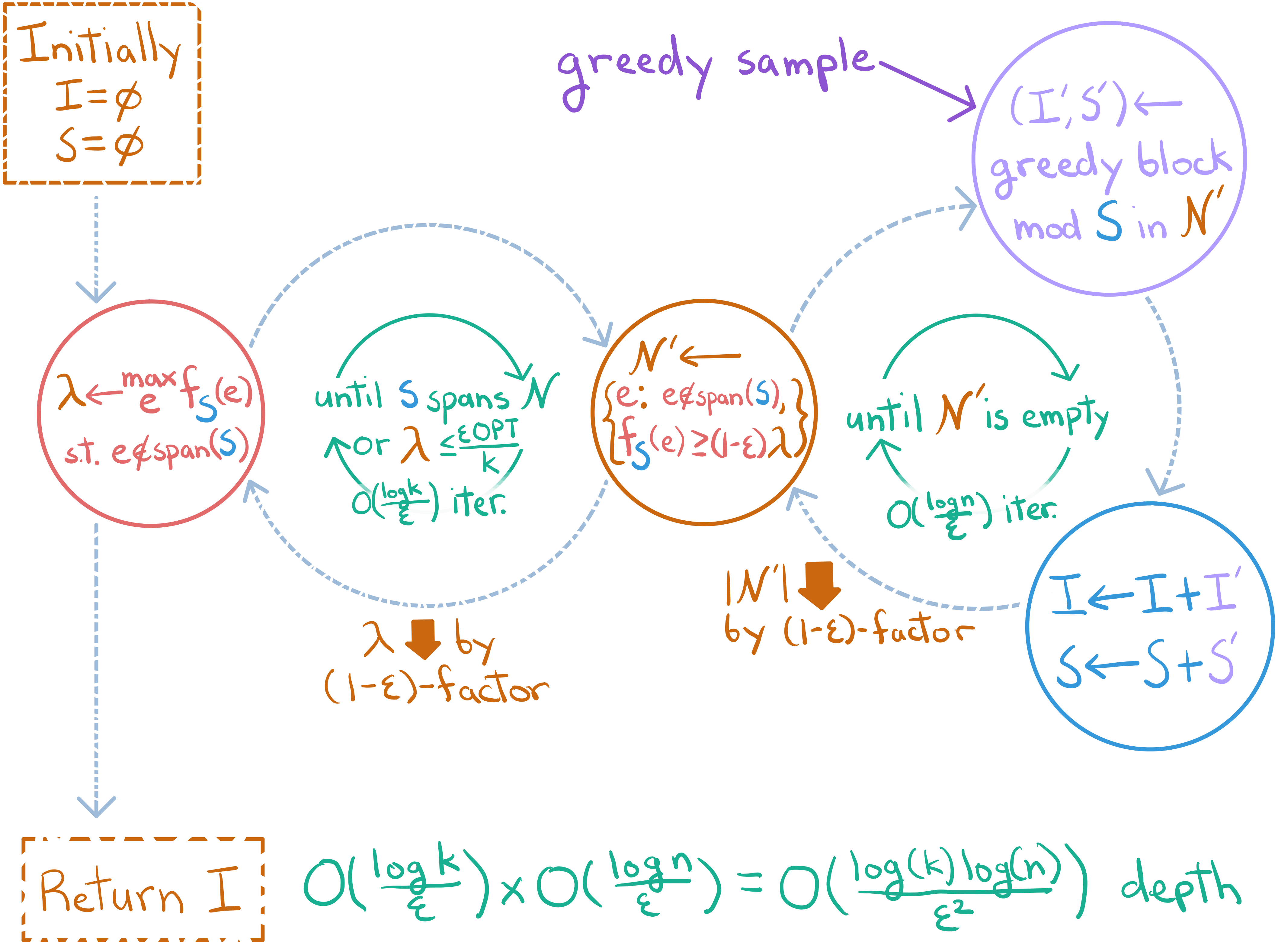
Key Claims

✓ (1) (I', S') is greedy block mod S

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

✓ (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. $e \notin span(S)$
until S spans N or $\lambda \leq \frac{\varepsilon OPT}{k}$
$O(\frac{\log k}{\varepsilon})$ iter.
$N' \leftarrow \{e : e \notin span(S), f_S(e) \geq (1-\varepsilon)\lambda\}$
until N' is empty
$O(\frac{\log n}{\varepsilon})$ iter.
$I \leftarrow I + I'$
$S \leftarrow S + S'$
λ ⬇ by $(1-\varepsilon)$-factor
$|N'|$ ⬇ 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

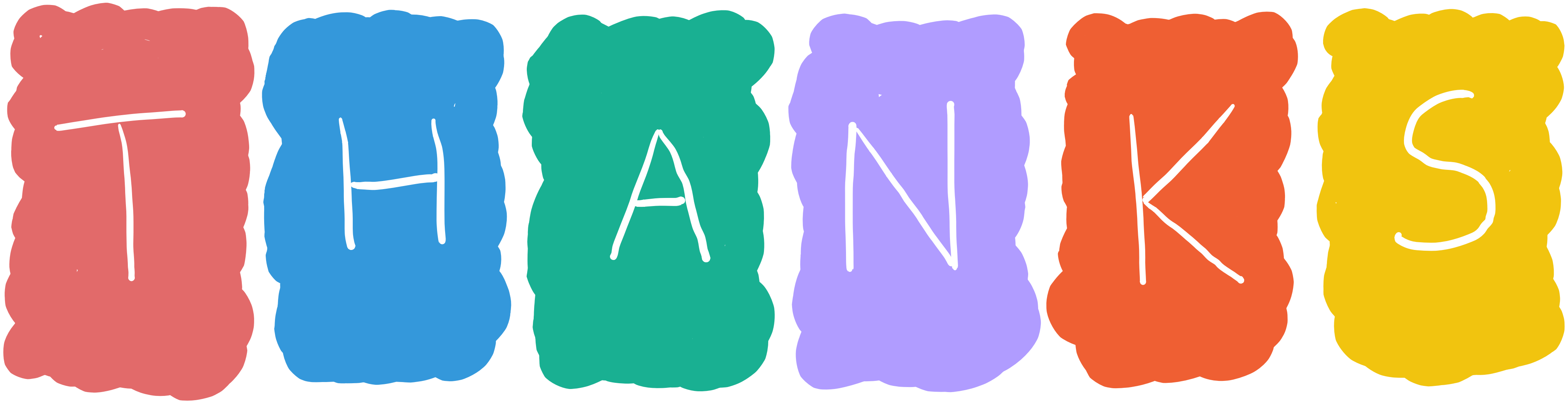
k-matchoid constraints in polylog depth

similar results to greedy

Open problems

rounding matroids in parallel

further (combinations) of constraints


THANKS

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

(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.

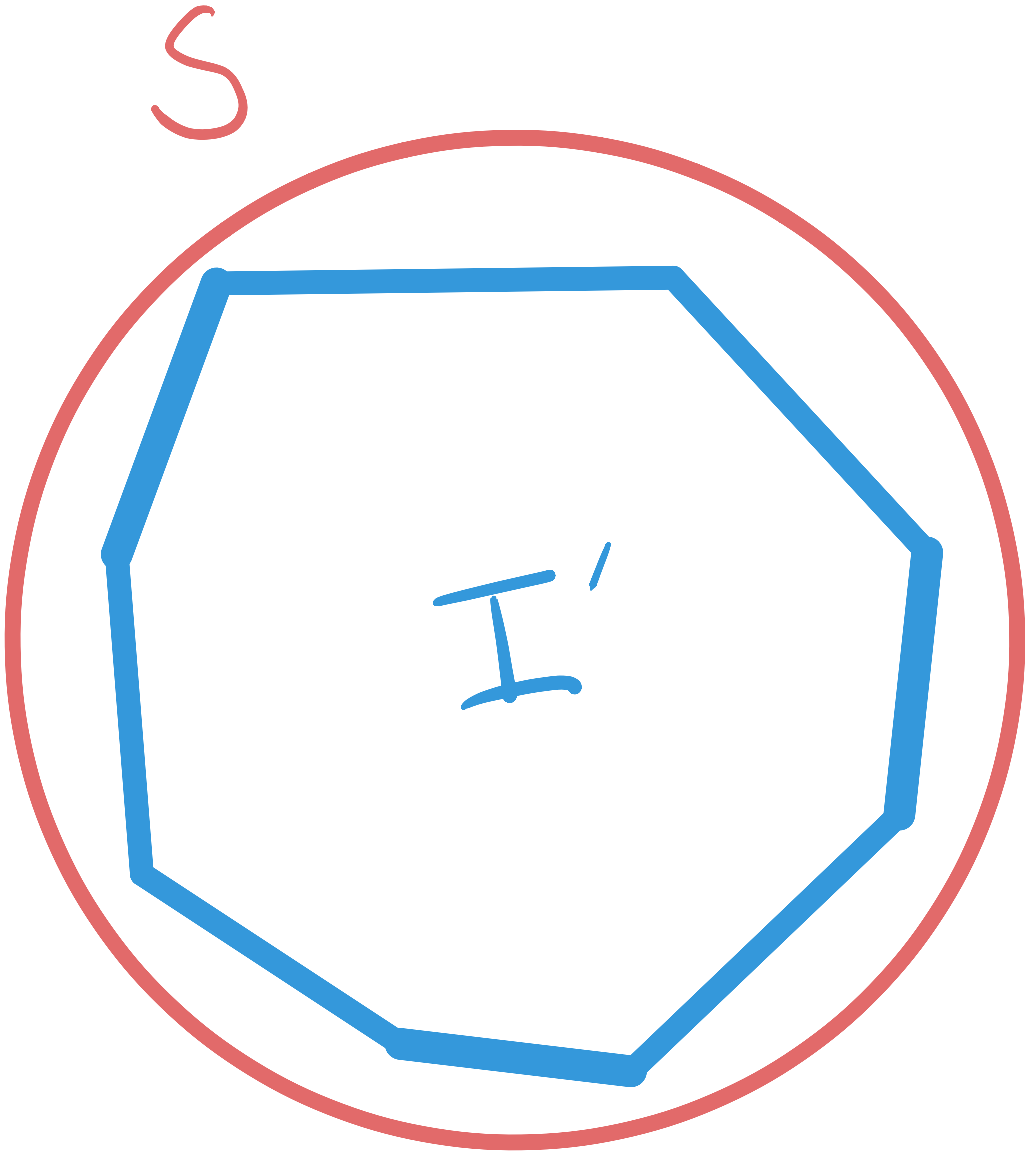
for $S' \sim_\delta S$:

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

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

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

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



"greedy sampling"

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

for $S' \sim \delta S$:

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

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

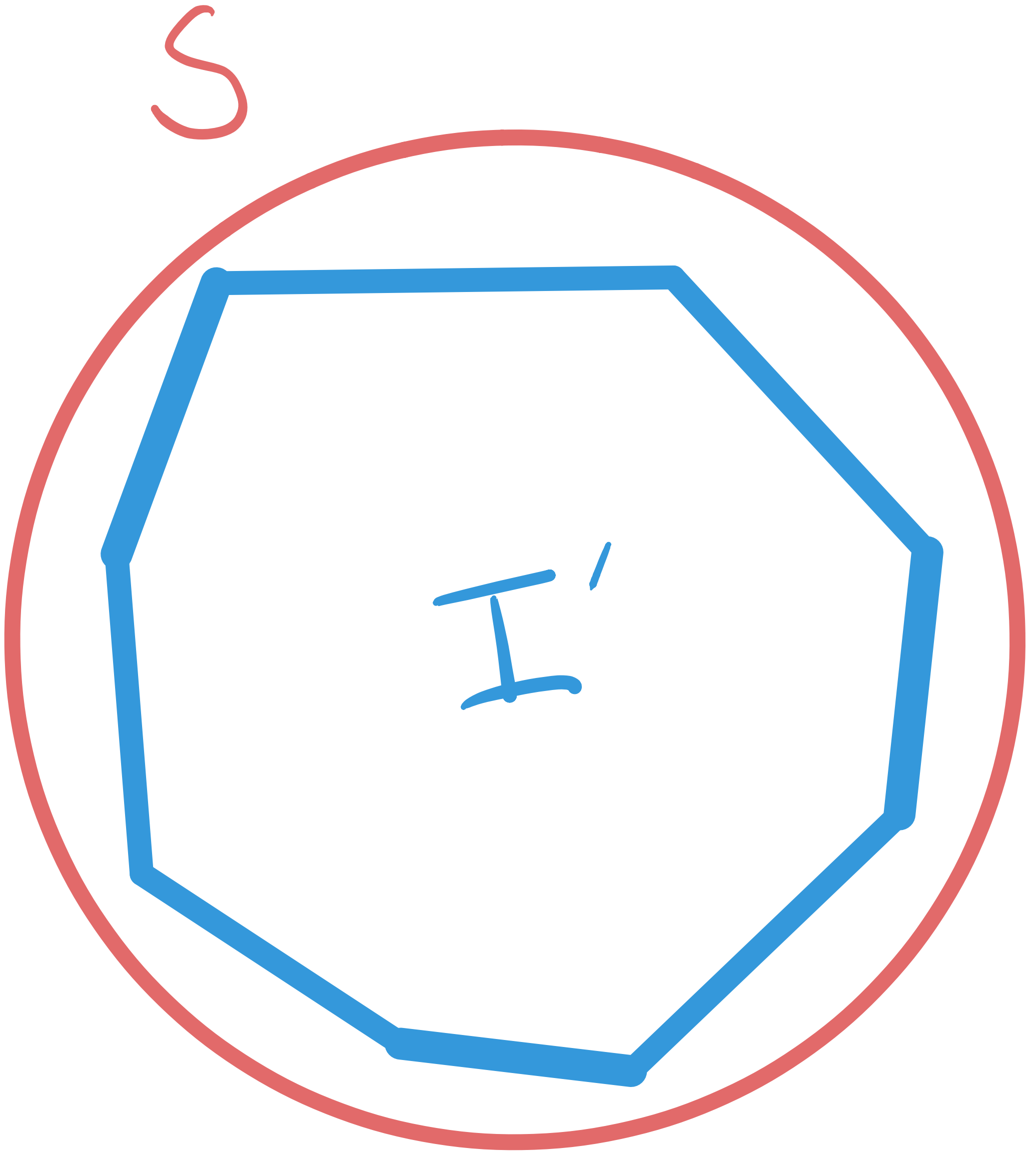
(a) sample $S' \sim \delta N'$

(b) $I' \leftarrow \left\{ \begin{array}{l} e \in S' : e \notin \operatorname{span}(S+S'-e) \\ \text{and } f_{S+S'-e}(e) \geq (1-\varepsilon)\lambda \end{array} \right\}$

Define

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

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



"greedy sampling"

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

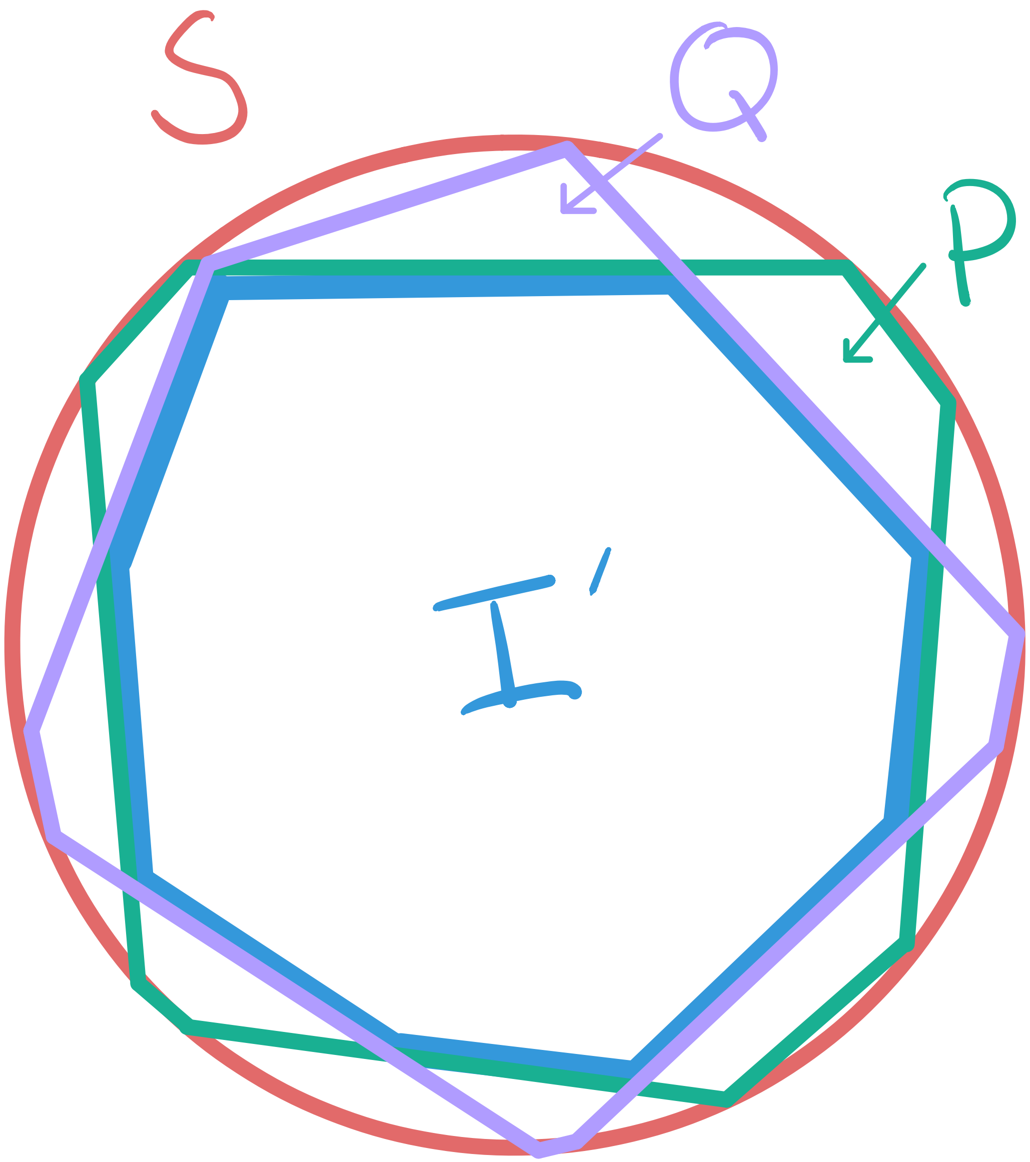
for $S' \sim_\delta S$:

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

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

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

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



Define

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

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

"greedy sampling"

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

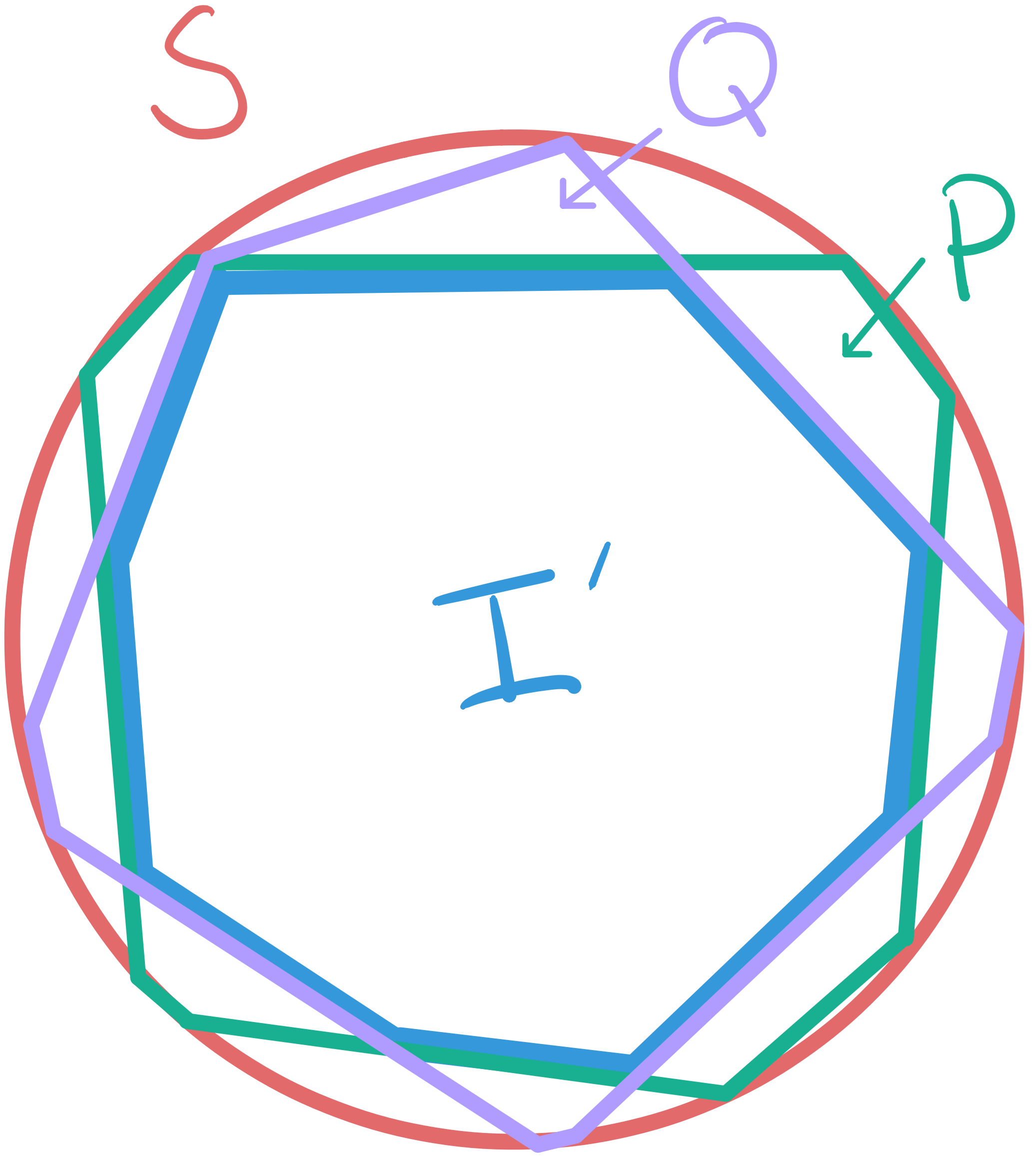
for $S' \sim_\delta S$:

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

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

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

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



Define

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

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

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

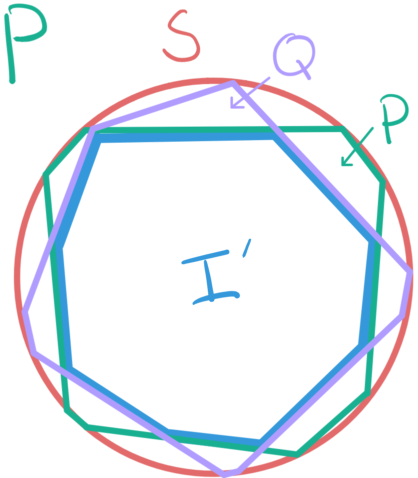
claim 3: $E[f_S(I')] \geq (1-2\varepsilon)\lambda E[|S'|]$

I' has good "bang-for-buck"

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

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

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



"greedy sampling"

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

for $S' \sim_{\delta} S$:

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

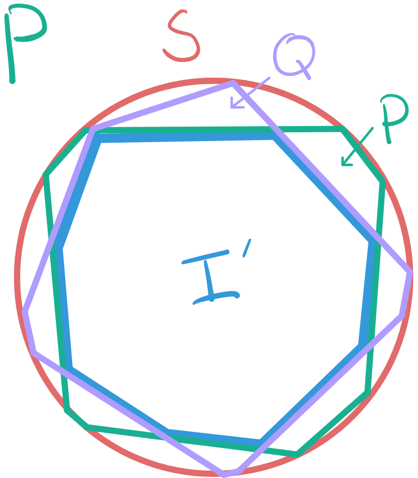
(b) $E[|\{e\in N': e\notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

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



3 claims

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

Q has good "bang-for-buck"

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

most sampled elements are not in cycles, retained by P

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

I' has good "bang-for-buck"

"greedy sampling"

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

for $S' \sim \delta S$:

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

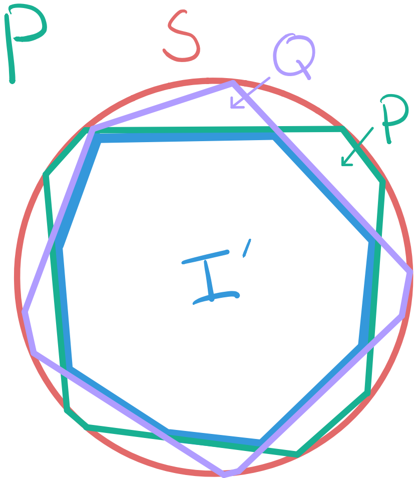
(b) $E[|\{e \in N' : e \notin \text{span}(S+S')\}|] \leq \varepsilon |N'|$

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

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

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

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



Proof ideas

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

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

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

δ small enough that P keeps most of S'

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

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

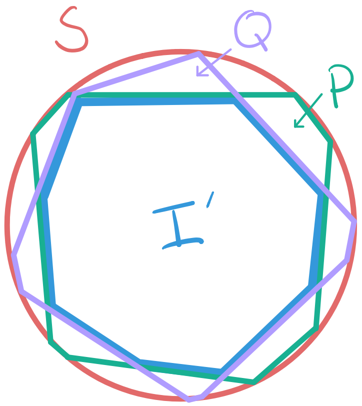
(b) $E[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \le \varepsilon |N'|$

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

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

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

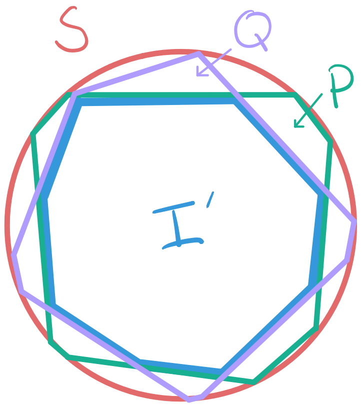
(b) $E[|\{e \in N' : e \notin \text{span}(S+S')\}|] \leq \varepsilon |N'|$

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

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

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

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



[skip]

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

Fix e.

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

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

by independence of &

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

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

"greedy sampling"

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

for $S' \sim \delta S$:

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

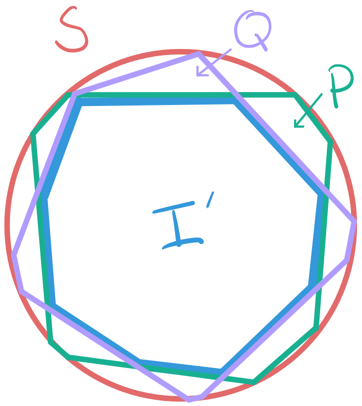
(b) $E[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \leq \varepsilon |N'|$

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

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

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

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



[skip]

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

Fix e.

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

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

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

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

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

"greedy sampling"

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

for $S' \sim \delta S$:

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

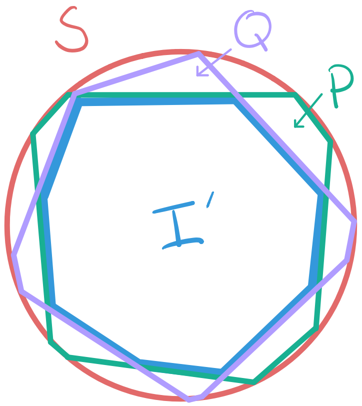
(b) $E[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

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



[skip]

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

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

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

(by submodularity)

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

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

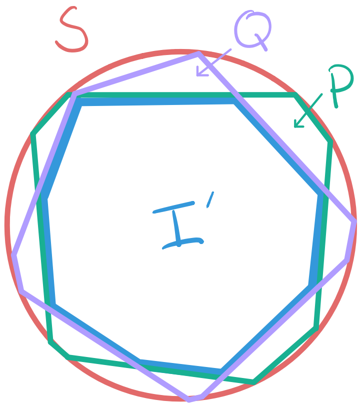
(b) $E[|\{e \in N' : e \notin span(S+S')\}|] \le \varepsilon |N'|$

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

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

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

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



[skip]

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

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

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

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

by the above (*)

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

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

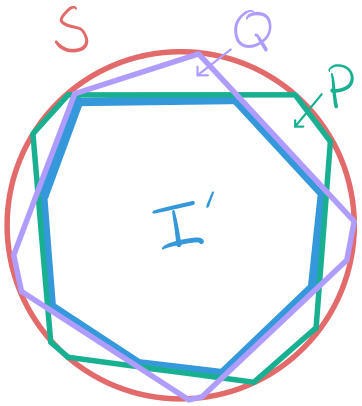
(b) $E[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \leq \varepsilon |N'|$

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

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

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

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



[skip]

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

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

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

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

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

by submodularity

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

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

"greedy sampling"

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

for $S'\sim\delta S$:

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

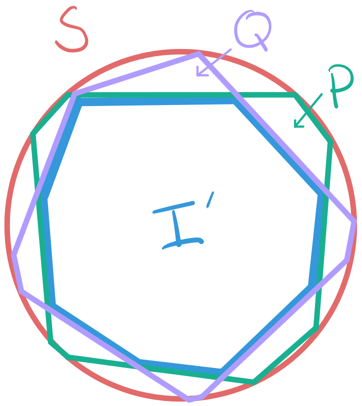
(b) $E[|\{e\in N': e\notin \mathrm{span}(S+S')\}|] \le \varepsilon|N'|$

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

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

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

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

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

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

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

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

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

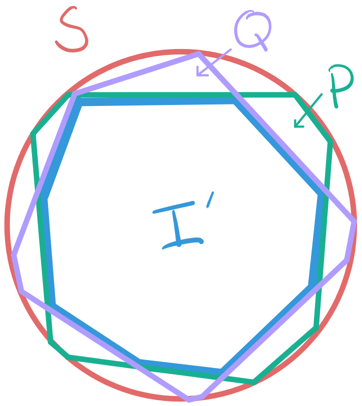
(b) $E[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \leq \varepsilon |N'|$

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

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

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

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



[skip]

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

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

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

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

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

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

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

by choice of δ per (a)

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

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

"greedy sampling"

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

for $S' \sim \delta S$:

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

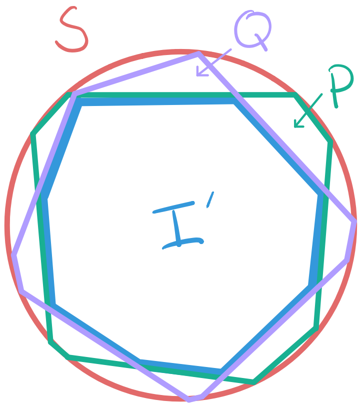
(b) $E[|\{e\in N': e\notin \mathrm{span}(S+S')\}|] \le \varepsilon|N'|$

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

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

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

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

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

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

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

since $S \sim \delta N'$ ∎

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

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

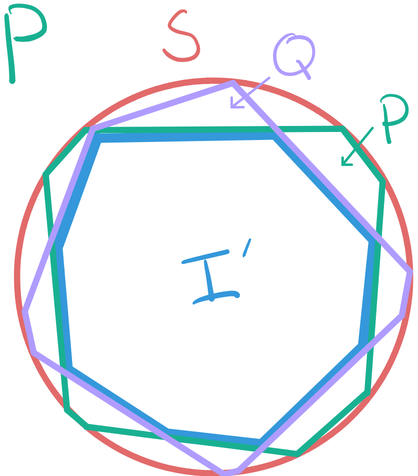
(b) $E[|\{e \in N' : e \notin span(S+S')\}|] \leq \varepsilon |N'|$

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

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

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

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



Proof ideas

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

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

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

δ small enough that P keeps most of S'

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

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

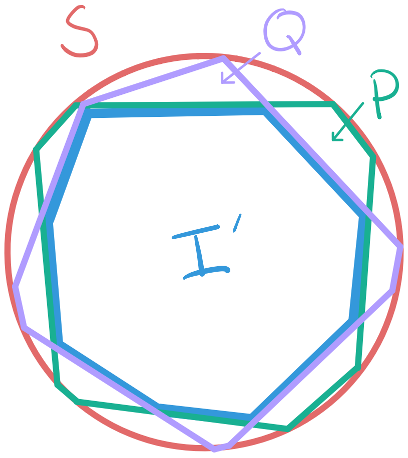
(b) $E[|\{e \in N' : e \notin \text{span}(S+S')\}|] \le \varepsilon |N'|$

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

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

fix e.

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

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

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

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

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

[skip]

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

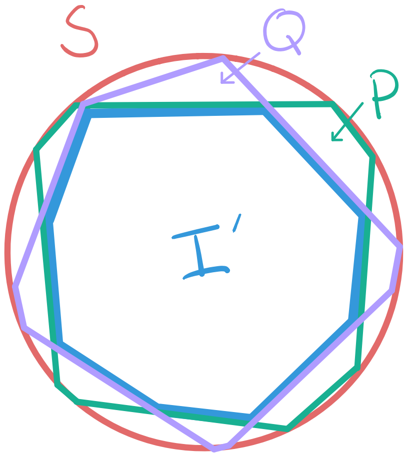
(b) $E[|\{e \in N' : e \notin \text{span}(S+S')\}|] \leq \varepsilon |N'|$

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

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

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

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



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

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

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

δ small enough that P keeps most of S'

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

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

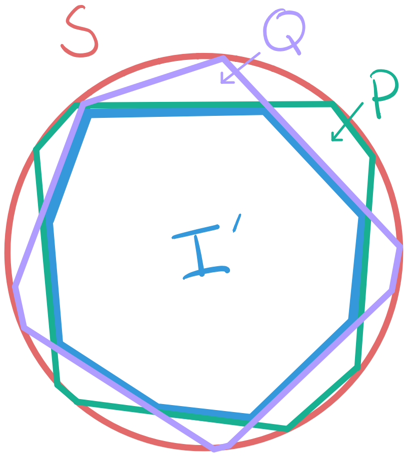
(b) $E[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \leq \varepsilon |N'|$

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

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

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

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



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

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

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

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

[since $I' \subseteq Q$]

"greedy sampling"

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

for $S' \sim_\delta S$:

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

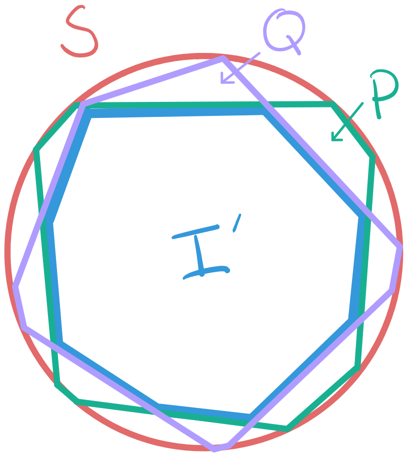
(b) $E[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \leq \varepsilon |N'|$

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

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

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

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



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

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

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

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

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

[by submodularity]

"greedy sampling"

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

for $S' \sim \delta S$:

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

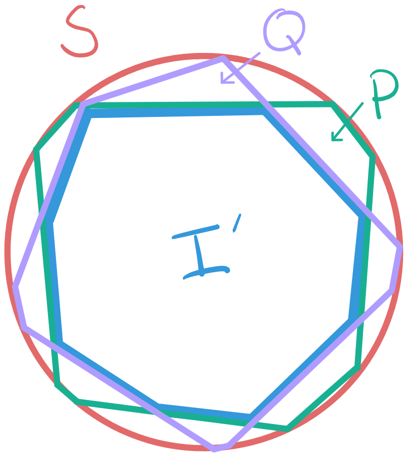
(b) $E[|\{e \in N' : e \notin \text{span}(S+S')\}|] \leq \varepsilon |N'|$

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

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

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

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



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

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

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

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

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

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

[by submodularity]

"greedy sampling"

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

for $S' \sim \delta S$:

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

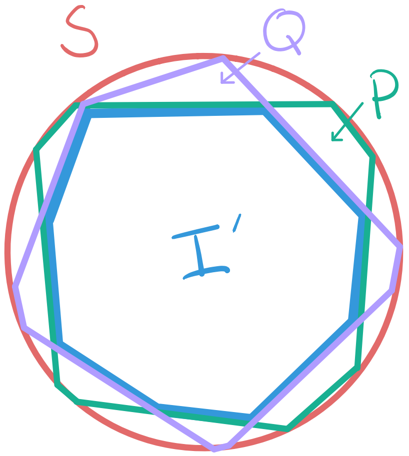
(b) $E[|\{e \in N' : e \notin \text{span}(S+S')\}|] \leq \varepsilon |N'|$

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

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

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

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



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

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

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

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

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

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

$$\geq E[f_S(Q)] - \lambda E[|Q \setminus I'|]$$

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

"greedy sampling"

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

for $S' \sim \delta S$:

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

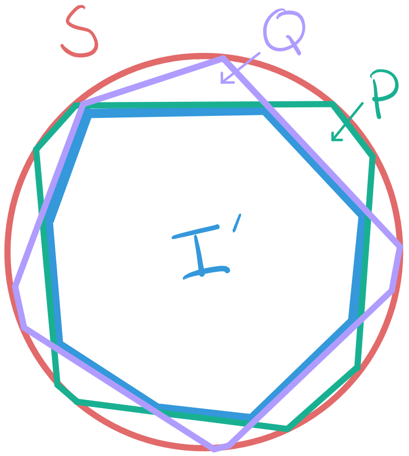
(b) $E[|\{e \in N' : e \notin \text{span}(S+S')\}|] \leq \varepsilon |N'|$

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

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

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

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



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

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

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

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

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

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

$$\geq E[f_S(Q)] - \lambda E[|Q \setminus I'|]$$

$$\geq E[f(Q)] - \lambda E[|S' \setminus P|]$$

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

"greedy sampling"

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

for $S' \sim \delta S$:

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

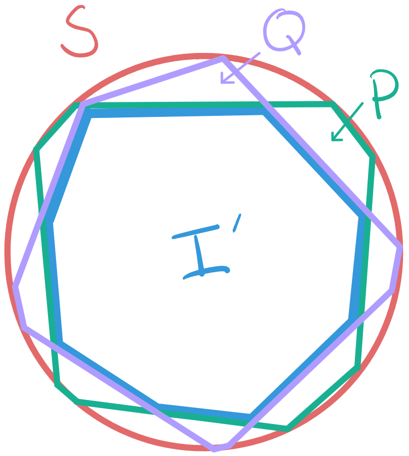
(b) $E[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \leq \varepsilon |N'|$

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

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

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

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



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

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

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

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

by claim 1 and claim 2

"greedy sampling"

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

for $S' \sim \delta S$:

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

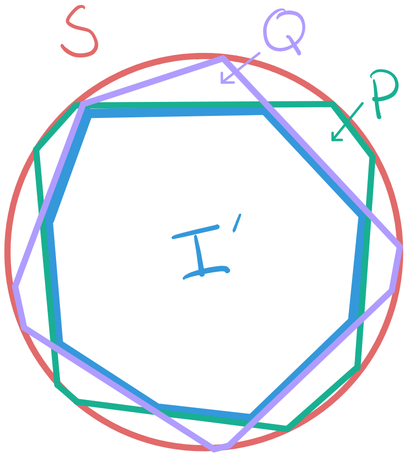
(b) $E[|\{e \in N' : e \notin \mathrm{span}(S+S')\}|] \le \varepsilon |N'|$

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

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

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

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 $I' = Q \cap P$