Pseudorandomness in Computer Science and in Additive Combinatorics

Luca Trevisan
University of California, Berkeley

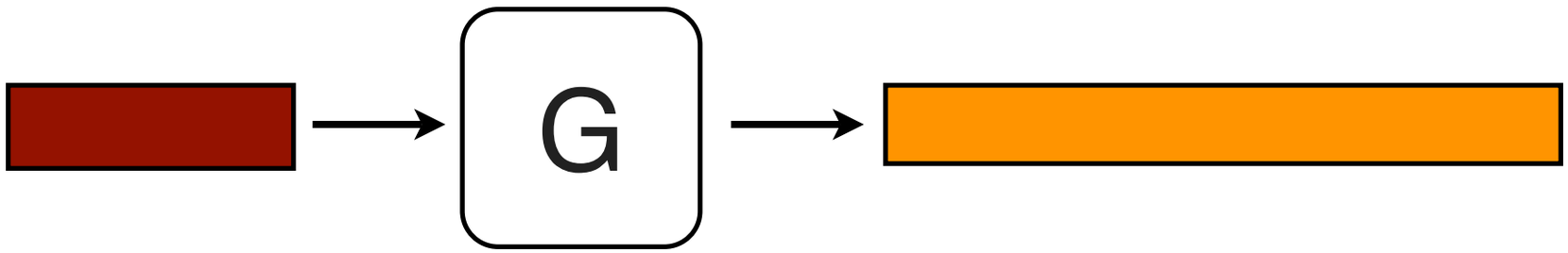
this talk

- explain the notions of pseudorandomness and indistinguishability from cryptography and complexity theory
- show their relation to notions of pseudorandomness and indistinguishability arise in additive combinatorics
- translate from language of norms, “decomposition” and “transference” theorems, etc.

this talk

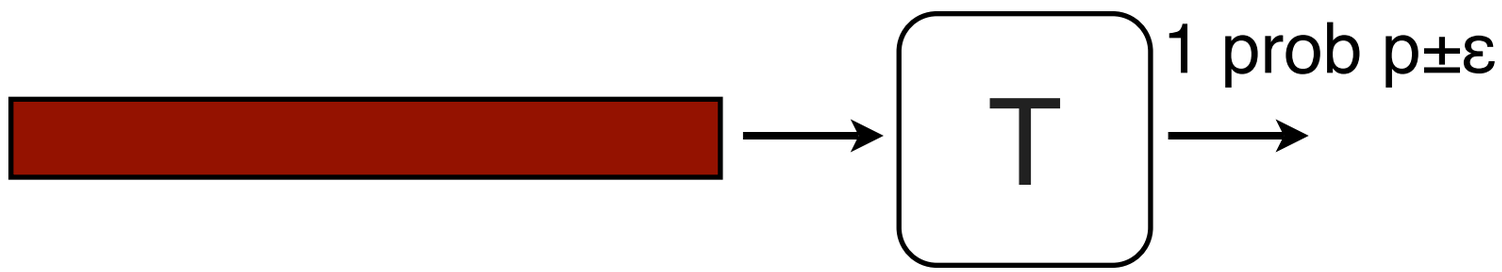
- quasirandom graphs, weak regularity lemma
- Gowers norm, decomposition thms
- Green-Tao transference thm

pseudorandom generator



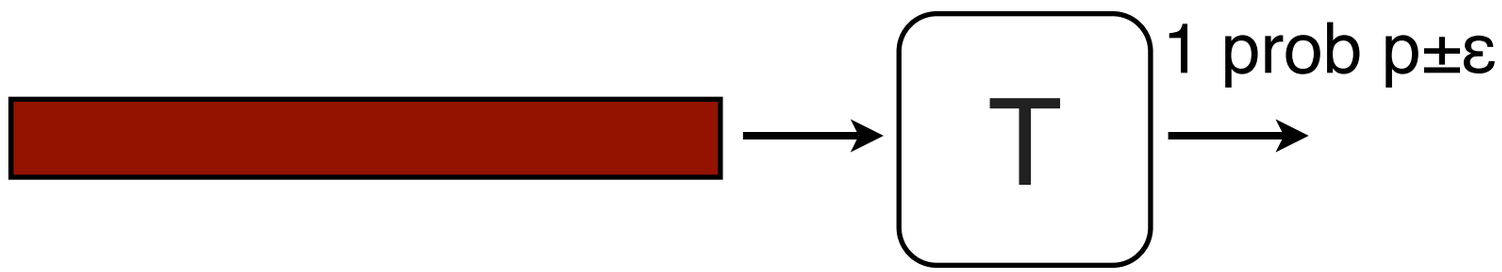
- deterministic procedure
- output longer than input
- when input is uniform, output “looks random”

pseudorandom generator



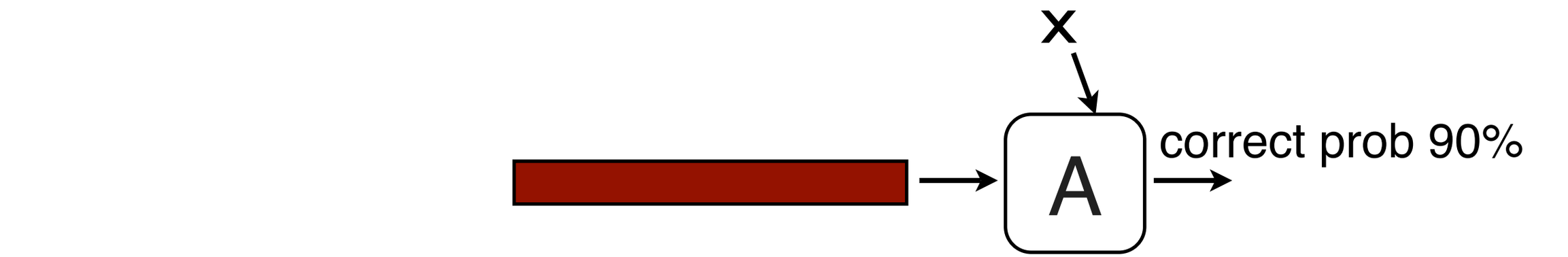
For every “efficient” test T

pseudorandom generator

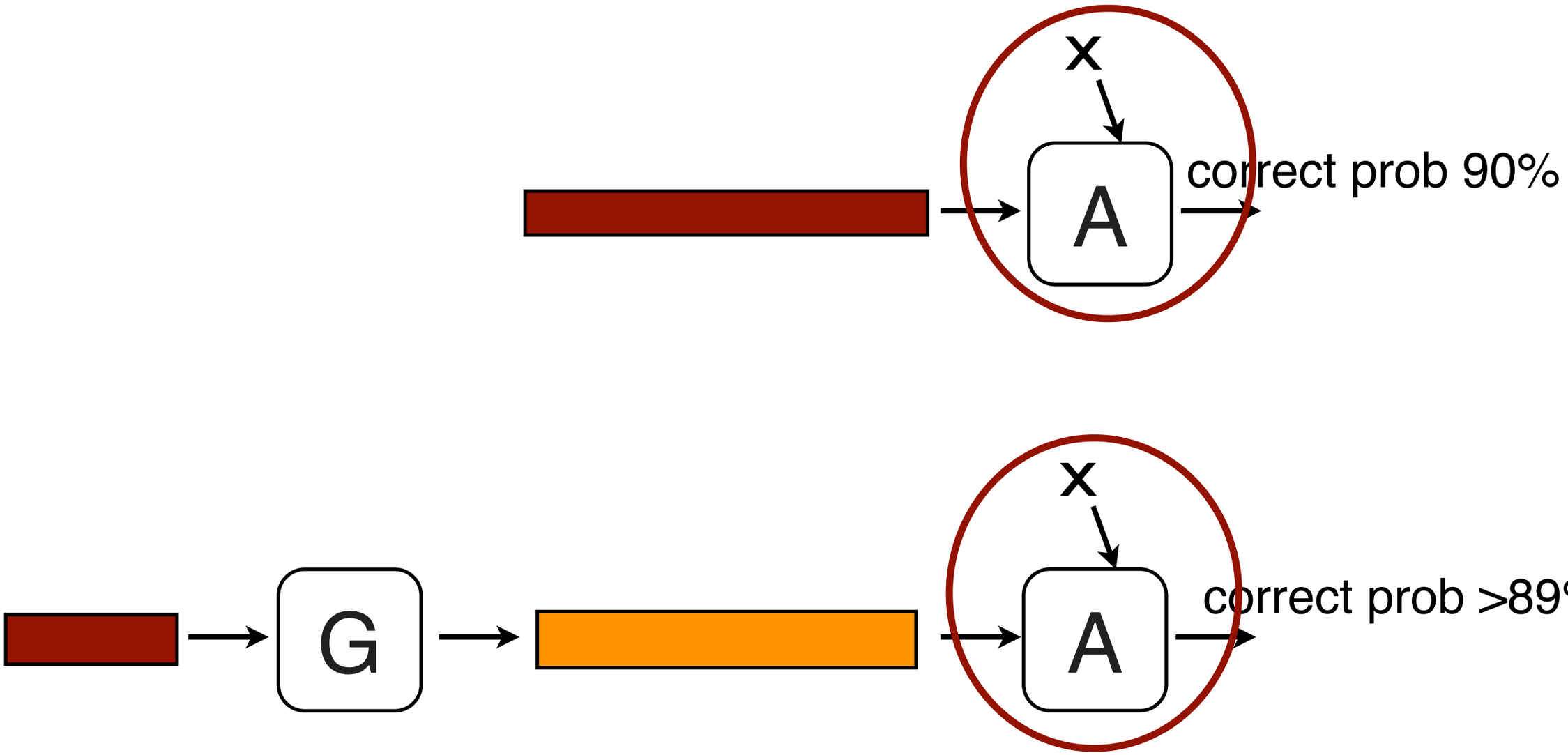


For every test T in a class ***C*** of functions
- Then we say G “ε-fools” ***C***

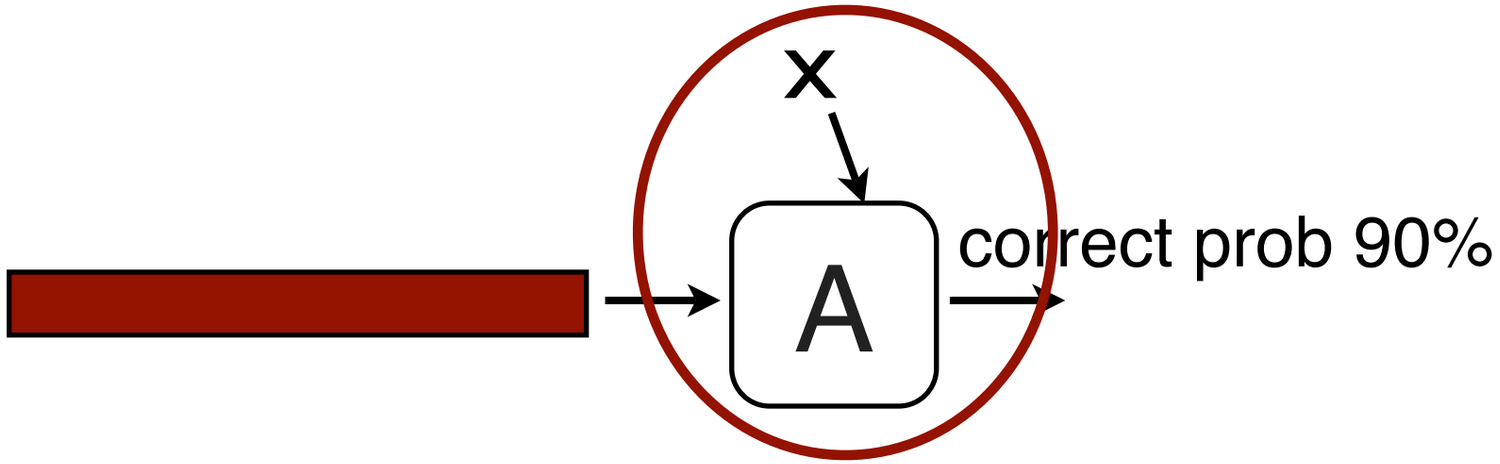
application

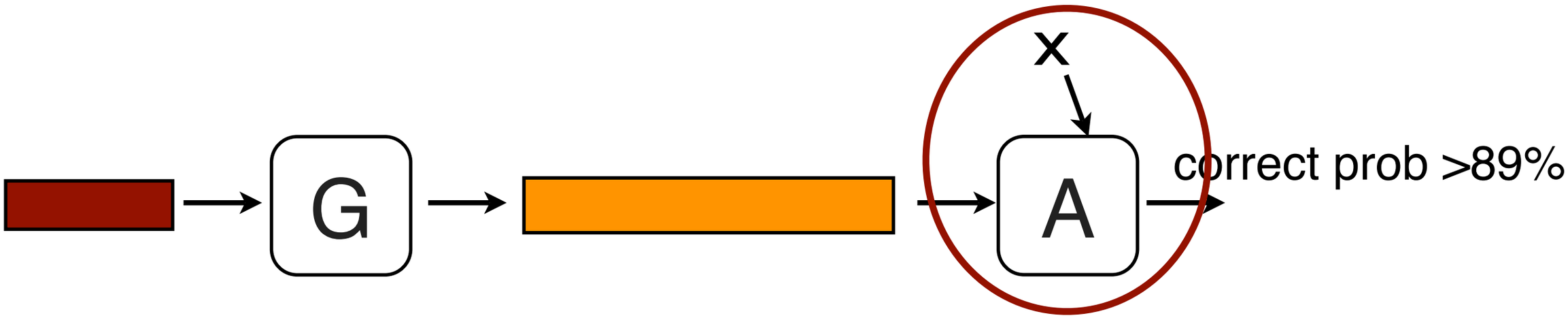


application



application





derandomization

pseudorandomness

random variable X taking values in $\{0,1\}^n$ is
ε-pseudorandom for class of algorithms ***C***
if for every T in ***C***:

$$| \Pr [T(X) = 1] - \Pr [T(U_n) = 1] | \leq \varepsilon$$

(U_n is uniform distribution over $\{0,1\}^n$)

indistinguishability

random variables X,Y taking values in $\{0,1\}^n$ are
ε-indistinguishable for class of algorithms ***C***
if for every T in ***C***:

$$| \Pr [T(X) = 1] - \Pr [T(Y) = 1] | \leq \varepsilon$$

pseudorandomness and graphs

quasirandom graph

[Thomason, Chung-Graham]

G=(V,E) is quasirandom if
for every sets A,B ⊆ V
# of edges between A,B is approximately

$$|E| \cdot |A| \cdot |B| \cdot 2 / |V|^2$$

quasirandom graph

[Thomason, Chung-Graham]

G=(V,E) is ε-quasirandom if
for every sets $A,B \subseteq V$
# of edges between A,B is

$$|E| \cdot |A| \cdot |B| \cdot 2 / |V|^2 \pm \varepsilon \cdot |V|^2$$

quasirandomness / indistinguishability

- Identify graph G=(V,E) with uniform distribution over E
- Define ***C*** to be class of functions
  $C_{A,B}$ (u,v) = 1 iff (u,v) crosses sets (A,B)
- Then G is ε-pseudorandom iff G and $K_{|V|}$ are ε-indistinguishable by ***C***

note: domain of functions in ***C*** is the set of all pairs of vertices

weak regularity lemma

[Frieze-Kannan]

Given G=(V,E) and

there is G' that

- is ε-indistinguishable from G

- has "complexity" dependent only on ε: it is a (edge-) disjoint union of $\exp(\varepsilon^{-O(1)})$ complete bipartite weighted graphs

Gowers norms

Szemeredi's thm

- For every k, every δ, every subset $A \subseteq \{1,\ldots,N\}$ with $|A| > \delta N$
  - *A* contains a length-k arithmetic progression provided $N > N(\delta,k)$

At least 4 different proofs; each proof uses notions of "pseudorandomness"

Roth's proof

Roth (1953) proved that if $A \subseteq \{1,...,N\}$ has size δN, and $N > \exp(\exp(1/\delta))$, then A must contain a length-3 progression.

Win-win argument:

- If A is "pseudorandom": done
  - then it has $\approx \delta^3 N^2$ progressions, like a random set of size |A|
- If A is not "pseudorandom": recurse
  - then enough to find progressions in $A' \subseteq \{1,...,N'\}$ of density $\delta+\delta^2$

Roth's proof

$$\mathbf{E}_{x,y} A(x)A(x+y)A(x+y+y) = \Sigma_s \hat{A}(-2s)\,\hat{A}^2(s)$$

- counts length-3 progressions
- It is δ^3 plus an expression that is, in absolute value, at most $\delta \cdot \max_{s\neq 0} |\hat{A}(s)|$

Roth's proof

of length 3 progressions in A is at least

$N^2 \cdot (\delta^3 - \delta \max_{s \neq 0} | \hat{A}(s) |)$

1. If all coefficients $<< \delta^2$ we are done (pseudorandom case)
2. If a coefficient $> \delta^2$ then recursion to a case with density $> \delta + \delta^2$

Gowers's proof

Progressions of length 4?

If A has small Fourier coefficients, it does not follow that A has $\approx \delta^4$ progressions of length 4

Gowers introduces stronger notion of pseudorandomness

Gowers uniformity norm

- f: Z_N -> **R**

- Def:

  $\| f \|_{U^k} := (\mathbf{E}_{x,y_1,\ldots,y_k} \prod_{S \subseteq \{1,\ldots,k\}} f(x + \sum_{i \in S} y_i))^{1/2^k}$

- Main point:

  if $\| f - g \|_{U^k}$ is small and f,g bounded, then

  $\mathbf{E}\, f(x)f(x+y) \cdots f(x+ky) \approx \mathbf{E}\, g(x)g(x+y) \cdots g(x+ky)$

Gowers uniformity

- Main point:
  if $\| f - g \|_{U^k}$ is small and f,g bounded, then
  $\mathbf{E}\, f(x)f(x+y) \cdots f(x+ky) \approx \mathbf{E}\, g(x)g(x+y) \cdots g(x+ky)$

- If $\| 1_A - 1_B \|_{U^k}$ is small, then A,B, have approximately same number of length-k progressions

- If A has density δ, and $\| 1_A - \delta \|_{U^k}$ is small,
  then A has $\approx \delta^{k+1}$ fraction of all length-(k+1) progressions

Gowers's proof

$A \subseteq Z_N$, $|A| = \delta N$

- If $\| A - \delta \|_{U^k}$ is small, done
  - then A has $\approx \delta^{k+1}N^2$ length-(k+1) progressions (pseudorandom case)

- If $\| A - \delta \|_{U^k}$ is not small, recursion
  - reduce to finding progressions in $A' \subseteq Z_{N'}$ of density $\delta+\delta^{O(1)}$ (100 of 128 pages in the paper)

Gowers norm as indistinguishability

- A,B (indicator functions of) sets
- $\| A - B \|_{U^k}$ small means A,B approximately same number of length-(k+1) progressions
- “Indistinguishable” by an “adversary” that counts progressions
- Does not match computer science notion of indistinguishability

Gowers inverse conjecture

ll f ll_{Uk} is small iff for every "polynomial" p of "degree" k-1, f and p are not "correlated"

- The current status of the inverse conjecture is complicated:
  - True in F_p, p>k, for polynomials
  - False in F_p, $p \leq k$, for polynomials
  - True in Z/NZ for "polynomial" := low-complexity (k-1)-step nilsequence, k=2,3. (Larger k in progress)
- [cf. Green-Tao, Samorodnitsky, Lovett-Meshulam-Samorodnitsky, Green-Tao, Bergelson-Tao-Ziegler,Tao-Ziegler, Green-Tao-Ziegler]

indistinguishability vs. correlation

- Let D1, D2 be two probability distributions
- D1,D2 are ε-indistinguishable by ***C*** iff for every function f in ***C***
- $| \mathbf{E}_{x \sim D1} f(x) - \mathbf{E}_{x \sim D2} f(x) | \leq \varepsilon$
- iff:

  $| \Sigma_x D1(x)f(x) - D2(x)f(x) | \leq \varepsilon$

indistinguishability vs. correlation

- Let D1, D2 be two probability distributions
- D1,D2 are ε-indistinguishable by ***C*** iff for every function f in ***C***
- $| \mathbf{E}_{x \sim D1} f(x) - \mathbf{E}_{x \sim D2} f(x) | \leq \varepsilon$
- iff:

  $| \Sigma_x (D1(x)-D2(x)) f(x) | \leq \varepsilon$

indistinguishability vs. correlation

- Let D1, D2 be two probability distributions
- D1,D2 are ε-indistinguishable by ***C*** iff for every function f in ***C***
- $| \mathbf{E}_{x \sim D1} f(x) - \mathbf{E}_{x \sim D2} f(x) | \leq \varepsilon$
- iff:
  $| < (D1\text{-}D2) , f > | \leq \varepsilon$

view as a norm

- ***C*** is a class of bounded functions f: X -> [0,1]
- for a function g: X -> **R**,
  ll g ll$_{\boldsymbol{C}}$:= max $_{f \in \boldsymbol{C}}$ l <g,f> l

- ls always a norm; it is L1 if ***C*** is all bounded functions
- ll D1 - D2 ll$_{\boldsymbol{C}} \leq \varepsilon$
  iff
  D1,D2 ε-indistinguishable by ***C***

inverse conjecture

- Let A,B be (dense) sets
- Then 1_A-1_B has small k-th Gowers norm
  iff
  U_A , U_B indistinguishable by degree (k-1) polynomials

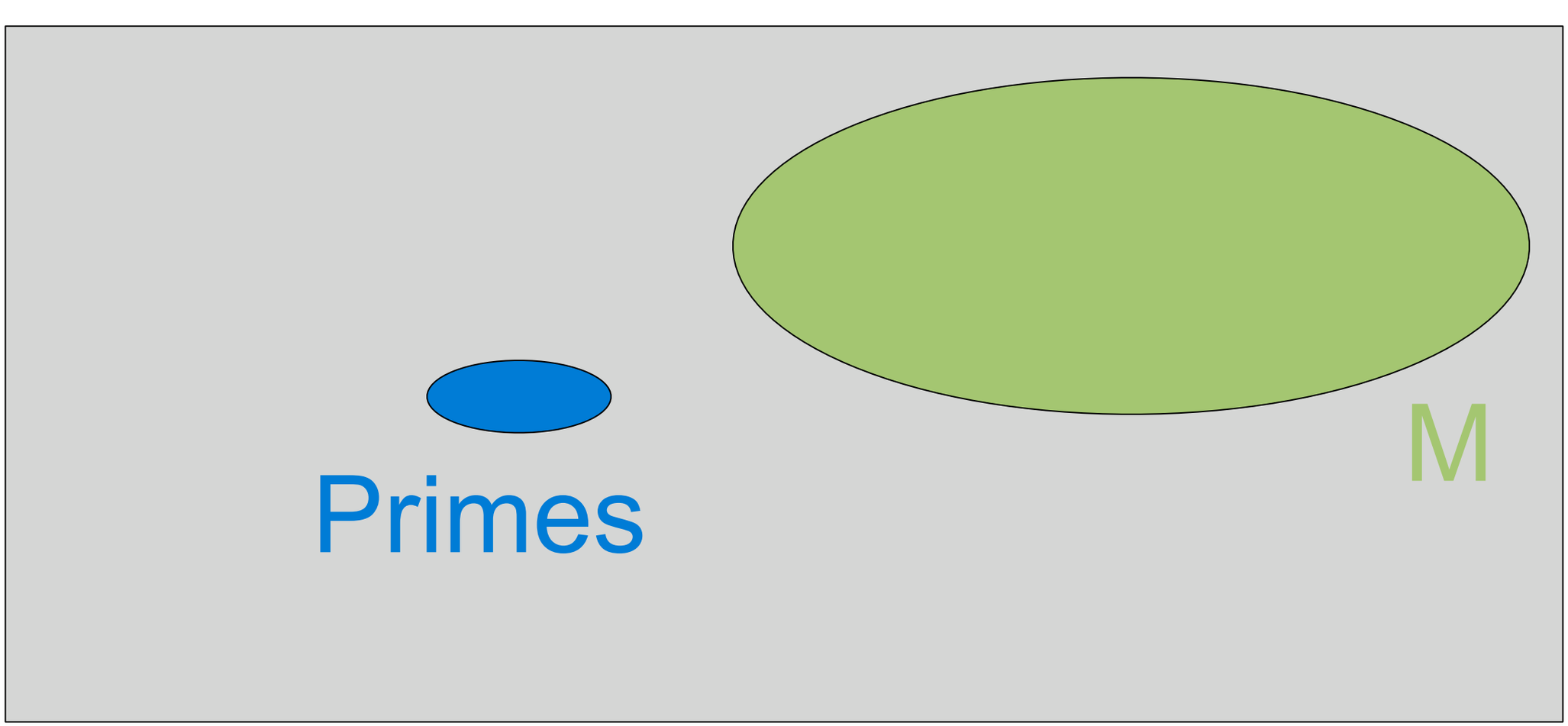
Transference thms

Green-Tao

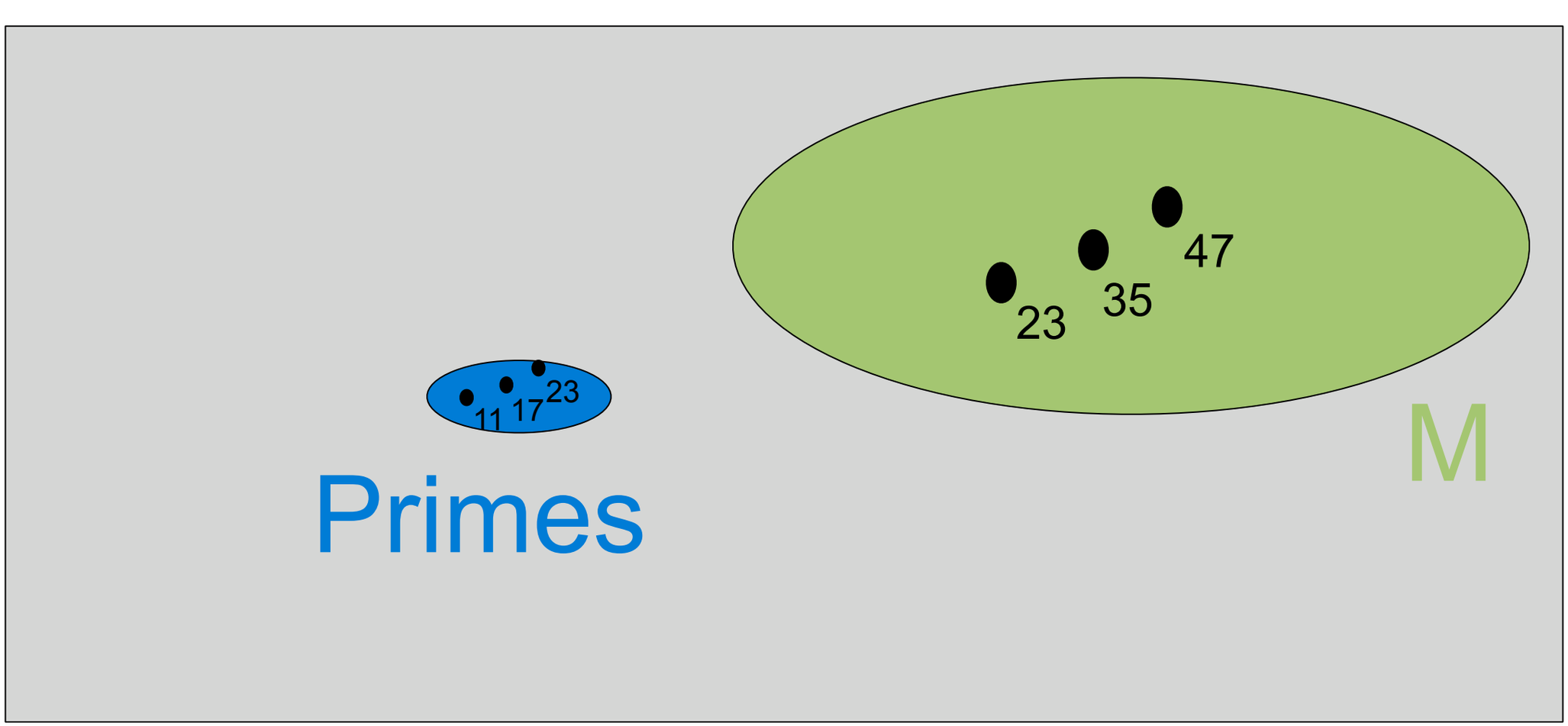
- The primes contain arbitrarily long arithmetic progressions

$\{1,\ldots,N\}$

$\{1,\ldots,N\}$



$\{1,\ldots,N\}$



model set M

Desired property of the model set:

$|M| > \Omega(N)$

$\| 1_M - 1_{Primes} \|_{Uk}$ small

model set M

Desired property of the model set M

$|M| > \Omega(N)$

$\| 1_M - 1_{Primes} \|_{Uk}$ small

impossible:

- functions of different averages are far in Uk norm

- Primes in {1,...,N} cannot have $\Omega(N^2)$ arithm. progr.

model set M

Desired property of the model set:

$\| 1_M - C*1_{Primes} \|_{U^k}$ small, $C=|M|/|Primes|$

- Problem:

  if $\| f - g \|_{U^k}$ is small and f,g bounded, then

  $\mathbf{E}\, f(x)f(x+y) \cdots f(x+ky) \approx \mathbf{E}\, g(x)g(x+y) \cdots g(x+ky)$

  but here $C*1_{Primes}$ is not bounded, $C \approx \log N$

- Can be overcome

model set M

Desired property of the model set:

$\| 1_M - C^* 1_{Primes} \|_{U^k}$ small

There is a class ***C*** of bounded functions f: [N] -> **R**

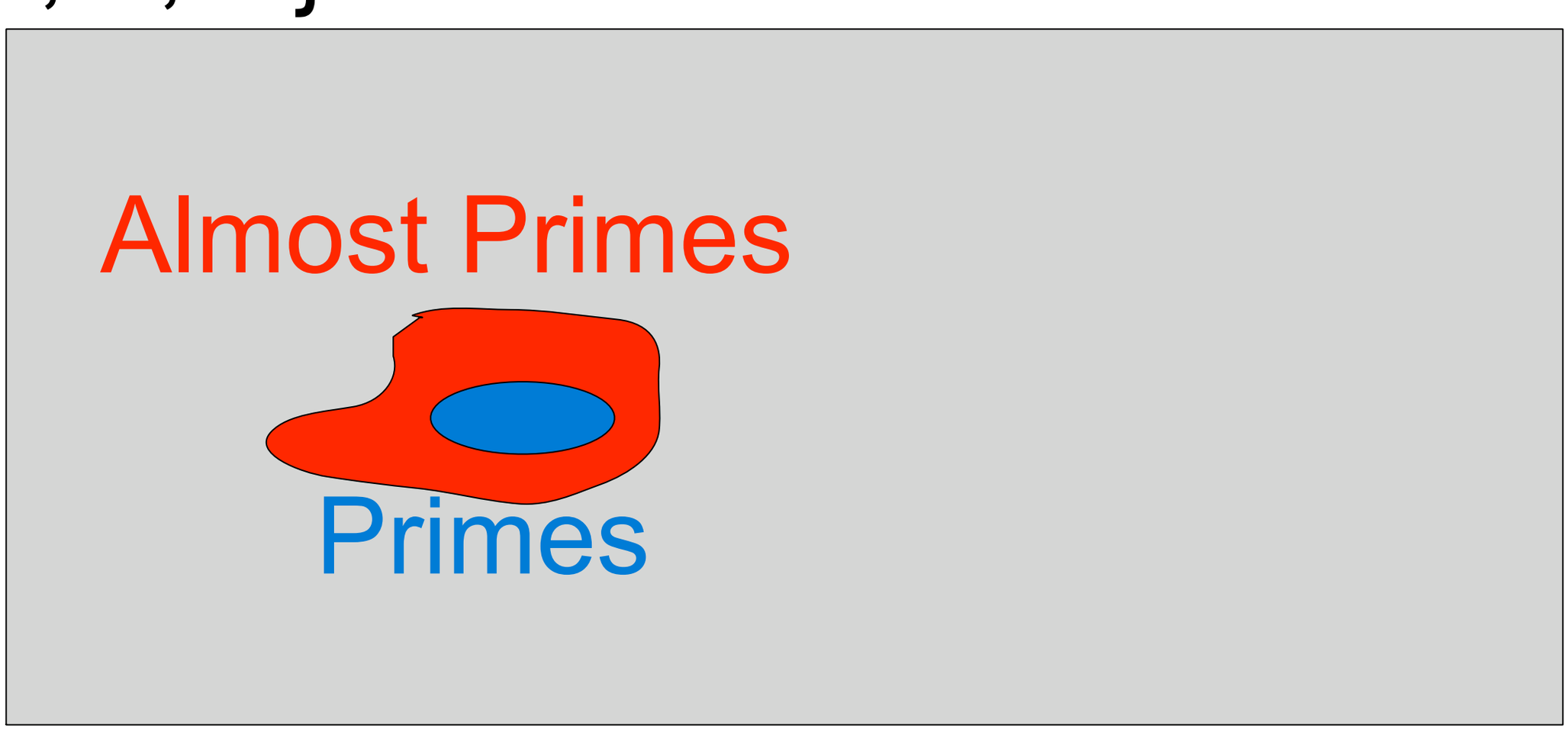
such that it is enough to prove

$< 1_M - C^* 1_{Primes} , f >$ small for every f in ***C***

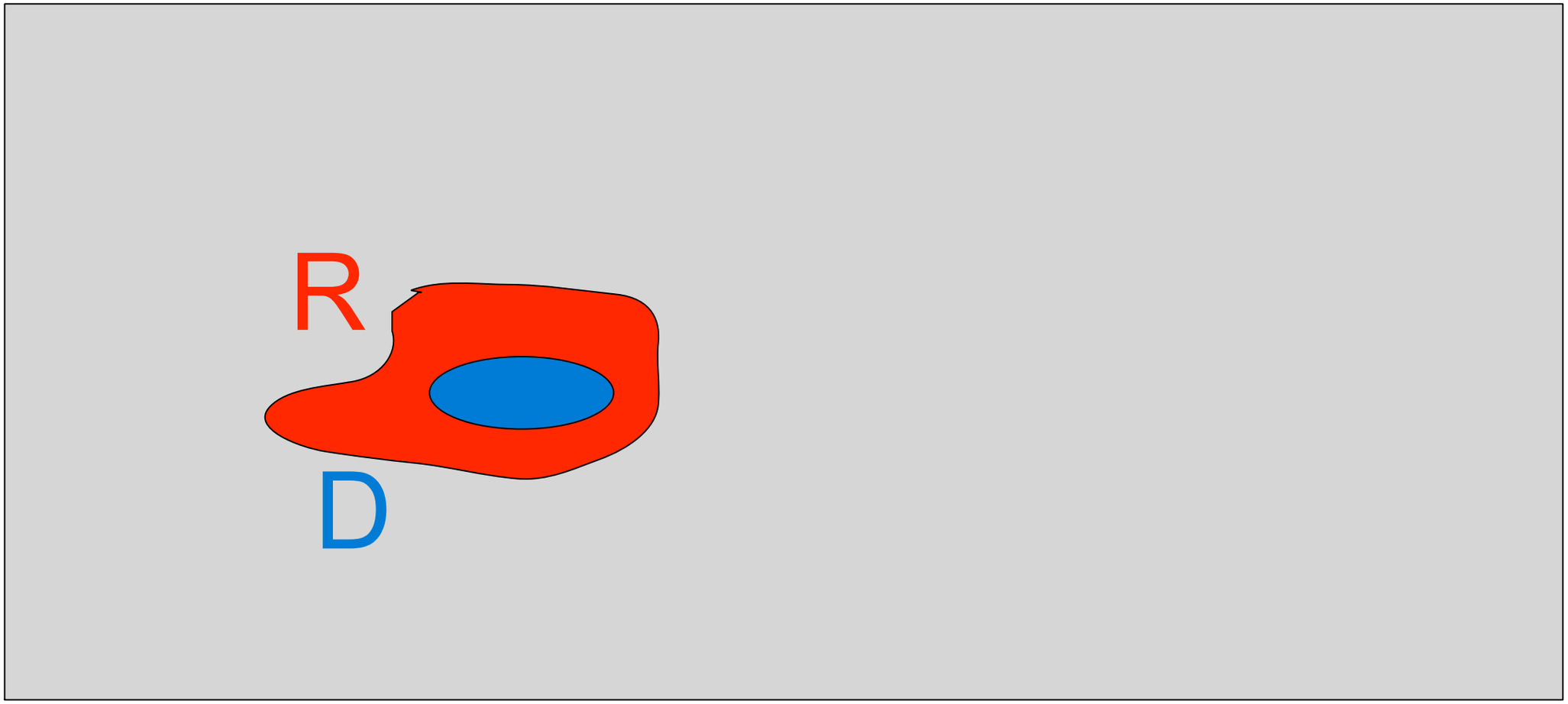
[***C*** could be (k+1)-step nilsequences given inverse conjecture, but a different ***C*** can be constructed otherwise]

$\{1,\ldots,N\}$

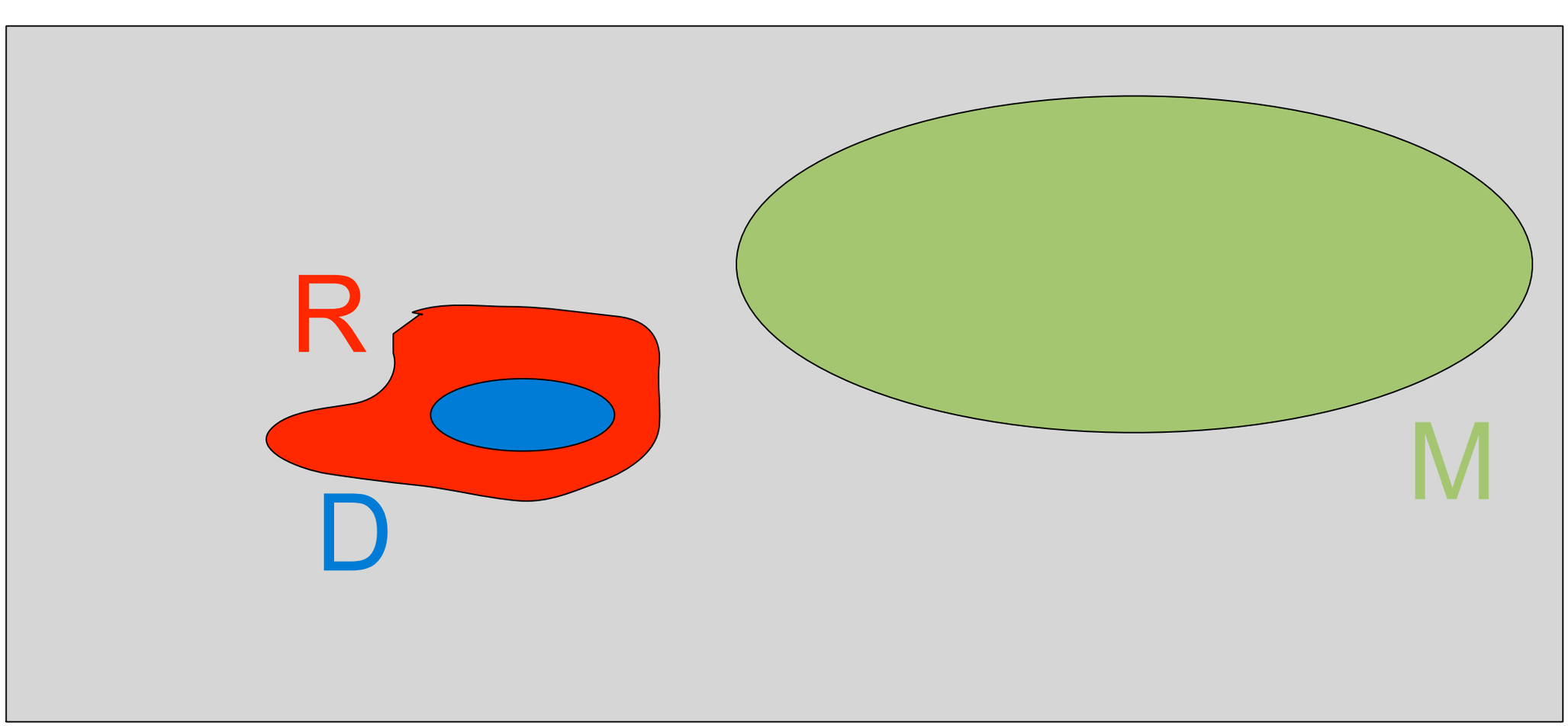
$\{1,\ldots,N\}$



X



X



dense model thm

[Green Tao] [Tao Ziegler] Given

r: X -> [0,C] **E**r=1 (indicator function of almost-primes)

g:X -> [0,C], $g \leq r$, **E**g = δ (primes)

C class of functions f: X -> [0,1], ε

Then either there is h: X -> [0,1], **E**h $\geq$ δ/2, s.t.

$\forall f \in \boldsymbol{C}$. $|\langle (h-g), f \rangle| \leq \varepsilon$

or $\exists\, d \in \boldsymbol{C'}$. $|\langle (r-1), d \rangle| \geq \varepsilon'$

(***C'*** contains simple combinations of functions from ***C***)

dense model theorem

- Can be proved in a computational setting:

- Computational setting: ***C'*** contains functions obtained by composing $(\varepsilon\delta)^{-O(1)}$ with operations of "complexity" $(\varepsilon\delta)^{-O(1)}$.

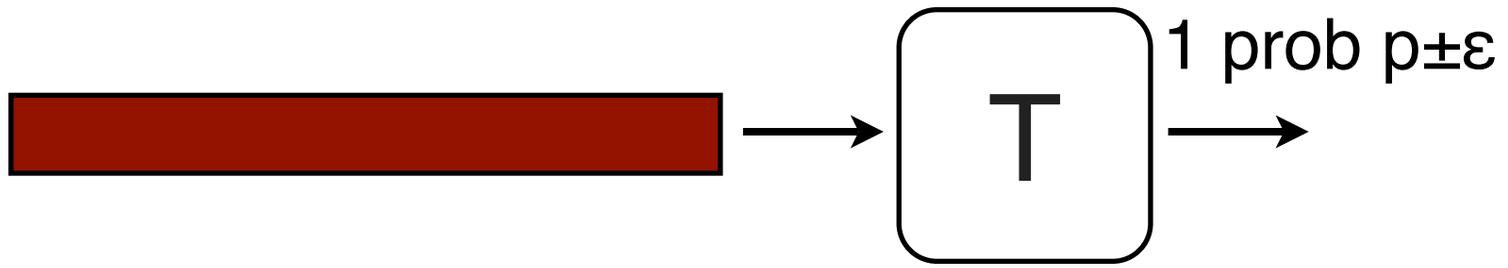
- In Green-Tao-Ziegler proofs: composition has $\exp((\varepsilon\delta)^{-O(1)})$ complexity

- [Reingold T Tulsiani Vadhan 2008,Impagliazzo 2008] Proof uses duality of linear programming
  [same idea in Gowers 2008, Gowers-Wolf 2009]

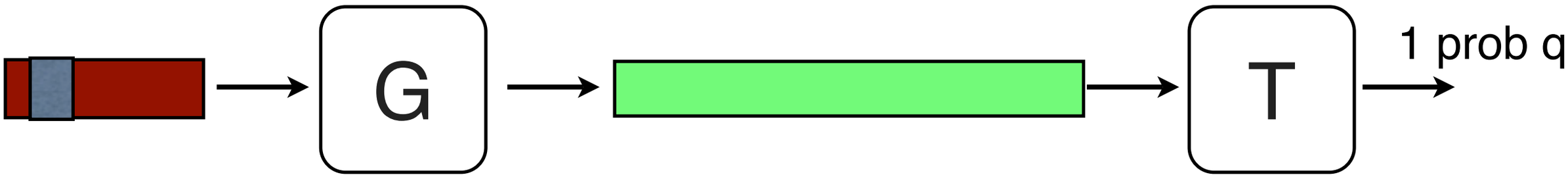
computational dense model theorem

- Application:
- Suppose G is a pseudorandom generator mapping t bits into n bits
- X is a distribution of entropy t-2
- There is distribution of M of entropy n-2 that is indistinguishable from G(X)
- Useful to secure against key leakage
  c.f. [Dziembowsky-Pietrzak]

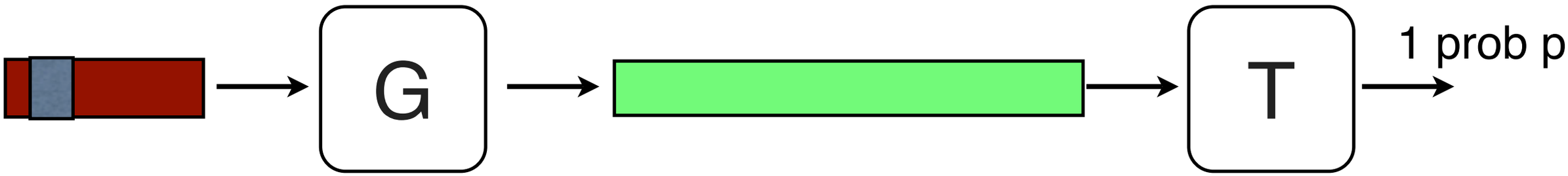
pseudorandom generator

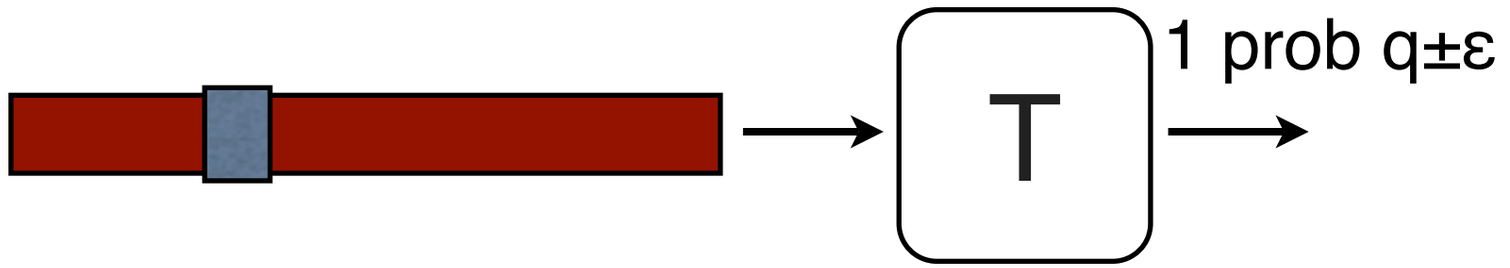


pseudorandom generator



pseudorandom generator





Decomposition thms

decomposition results in add comb

Many theorems have form:

given g arbitrary function, ***C*** class of functions

can write

$g = g_s + g_r$

where: g_s is “structured” (related to ***C***)

g_r has low correlation with ***C***

“efficient decomposition” result

- Given:
  C class of bounded function f: X-> [-1,1]
  g: X -> [-1,1]
  ε

- Can find $f_1,\ldots,f_k$, $k = O(\varepsilon^{-2})$ such that:
  - define $h(x) := \max\{ -1 , \min\{ 1, \Sigma_i \varepsilon f_i \} \}$
  - then $\langle g - h , f \rangle \leq \varepsilon$ for all f in ***C***

  [Tulsiani T Vadhan]

efficient decomposition

- Given ***C***, g, ε, there is a decomposition
  $g(x) = h_1(x) + h_2(x)$ where
  - $h_1(x)$ is "structured:" simple composition of ε^{-2} functions from ***C***
  - h_2 is "uniform:" $< h_2, f> \leq \varepsilon$ for all f in ***C***
- Implies Frieze-Kannan weak regularity lemma
- Implies (with a bit of work) dense model thm
- Every high-entropy distribution is indistinguishable from an efficiently computable distribution of same entropy

local testability

- A graph is pseudorandom iff it has approximately the same number of 4-cycles of a random graph with the same number of edges
- A function has low Gowers norm (hence pseudorandom w.r.t. low degree polynomials) if it is nearly unbiased in small dimensional "parallelograms"

- Is $\max_{f \in \boldsymbol{c}} |\langle g, f \rangle|$ small iff a "local" property of g holds?

local testability

- Is $\max_{f \in \boldsymbol{C}} |\langle g, f \rangle|$ small iff a "local" property of g holds?

- Not if $\boldsymbol{C}$ is the class of all efficiently computable functions, or even the class of functions obtained by a constant number of compositions of majority functions
[Razborov-Rudich, "Natural Proofs"]