Chernoff Bound for Sum of $\{-1, +1\}$ Random Variables

Theorem 1. *Let* $X_1, ..., X_n$ *be independent random variables with*

$$Pr(X_i = 1) = Pr(X_i = -1) = \frac{1}{2}.$$

Let $X = \sum_1^n X_i$*. For any* $a > 0$*,*

$$Pr(X \geq a) \leq e^{-a^2/2n}$$

Proof. For any $t > 0$,

$$E[e^{tX_i}] = \frac{1}{2}e^t + \frac{1}{2}e^{-t}.$$

$$e^t = 1 + t + \frac{t^2}{2!} + \ldots + \frac{t^i}{i!} + \ldots$$

and

$$e^{-t} = 1 - t + \frac{t^2}{2!} + \ldots + (-1)^i\frac{t^i}{i!} + \ldots$$

Thus,

$$\begin{aligned} E[e^{tX_i}] &= \frac{1}{2}e^t + \frac{1}{2}e^{-t} = \sum_{i\geq 0}\frac{t^{2i}}{2i!} \\ &\leq \sum_{i\geq 0}\frac{\left(\frac{t^2}{2}\right)^i}{i!} = e^{t^2/2} \end{aligned}$$

$$E[e^{tX}] = \prod_{i=1}^{n} E[e^{tX_i}] \leq e^{nt^2/2},$$

$$Pr(X \geq a) = Pr(e^{tX} > e^{ta}) \leq \frac{E[e^{tX}]}{e^{ta}} \leq e^{t^2 n/2 - ta}.$$

Setting $t = a/n$ yields

$$Pr(X \geq a) \leq e^{-a^2/2n}.$$

$\square$

By symmetry we also have

Corollary 1. *Let* $X_1, ..., X_n$ *be independent random variables with*

$$Pr(X_i = 1) = Pr(X_i = -1) = \frac{1}{2}.$$

Let $X = \sum_{i=1}^{n} X_i$*. Then for any* $a > 0$*,*

$$Pr(|X| > a) \leq 2e^{-a^2/2n}.$$

Application: Set Balancing Revisited

Theorem 2. *For a random vector* $\bar{b}$*, with entries chosen independently and with equal probability from the set* $\{-1, 1\}$*,*

$$Pr(||\mathcal{A}\bar{b}||_\infty \geq \sqrt{4n \ln n}) \leq \frac{2}{n} \quad (1)$$

Proof. Consider the i-th row $\bar{a}_i = a_{i,1}, ..., a_{i,n}$.

Let k be the number of 1's in that row.

$Z_i = \sum_{j=1}^{k} a_{i,i_j} b_{i_j}$.

If $k \leq \sqrt{4n \ln n}$ then clearly Z_i satisfies the bound.

If $k > \sqrt{4n \log n}$, the k non-zero terms in the sum Z_i are independent random variables, each with probability $1/2$ of being either $+1$ or -1.

Using the Chernoff bound:

$$Pr\left\{|Z_i| > \sqrt{4n \log n}\right\} \leq 2e^{-4n \log n/2k} \leq \frac{2}{n^2},$$

where we use the fact that $n \geq k$.

$\square$

Packet Routing on Parallel Computer

Communication network:

- Nodes - processors, switching nodes.
- edges - communication links.

The n-cube:

$N = 2^n$ nodes.

Let $\bar{x} = (x_1, ..., x_n)$ be the number of node x in binary.

Nodes x and y are connected by an edge iff their binary representations differ in exactly one bit.

Bit-wise routing: correct bit i in the i-th transition - route has length n.

A permutation communication request: each node is the source and destination of exactly one packet.

Up to one packet can cross an edge per step, each packet can cross up to one edge per step.

What is the time to route an arbitrary permutation on the n-cube?

Oblivious Routing

Definition 1. *A routing algorithm is* **oblivious** *if the path taken by one packet is independent of the source and destinations of any other packets in the system.*

Theorem 3. *Given an* n*-node network with maximum degree* d *the routing time of any* $deterministic\ oblivious$ *routing scheme is*

$$\Omega(\sqrt{\frac{n}{d^3}}).$$

Two phase routing algorithm:

1. Send packet to a randomly chosen destination.
2. Send packet from random place to real destination.

Path: Correct the bits, starting at x_0 to x_{n-1}.

Any greedy queuing method - if some packet can traverse an edge one does.

Theorem 4. *The two phase routing algorithm routes an arbitrary permutation on the* n*-cube in* $O(\log N) = O(n)$ *parallel steps with high probability.*

We focus first on phase 1. We bound the routing time of a given packet M.

Let $e_1, ..., e_m$ be the $m \leq n$ edges traversed by a given packet M is phase 1.

Let $X(e)$ be the total number of packets that traverse edge e at that phase.

Let $T(M)$ be the number of steps till M finished phase 1.

Lemma 1.

$$T(M) \leq \sum_{i=1}^{m} X(e_i).$$

$$E[T(M)] \leq \sum_{i=1}^{m} E[X(e_i)].$$

Let $P = (e_1, \ldots, e_m)$, $(m \leq n)$ be any path followed by the bit fixing algorithm. Nodes are $v_0, \ldots, v_m$.

For any path P let $T(P) = \sum_{i=1}^{m} X(e_i)$.

We bound the probability that $T(P)$ is large for any P.

A packet is **active** at a node v_{i-1} if it reaches v_{i-1} and has the possibility of traversing e_i.

A packet is active if it is active at some node i.e., reaches, some node of path P.

Since traversing e_i "fixes" the j-th bit, a packet can cross that edge only in its j-th transition.

Assume that e_i connects $v_{i-1} = (a_1, ..., a_{j-1}, a_j, ..., a_n)$ to $v_i = (a_1, .., \bar{a_j}, ..., a_n)$.

Only packets that started in address

$$(*, ..., *, a_j, ..., a_n)$$

can reach vertex v_{i-1}, before the jth bit is fixed and only if their destination addresses are

$$(a_1, ..., a_{j-1}, *, *, ..., *)$$

.

There are 2^{j-1} possible packets, each has probability $2^{-(j-1)}$ to reach v_{i-1}.

Thus expected number of active packets per vertex is 1.