## Using Chernoff's Bound to obtain effective bandwidth (discrete time case)

- let $S_{-k} \triangleq A_0 - A_{-k} - Ck$, for $k \geq 0$

  & let $X = \sup_{k \geq 0} (A_0 - A_{-k} - Ck)$

  $= \sup_{k \geq 0} S_{-k}$

  denote the random amount of data in the buffer in steady state

- We are interested in studying $Pr(X > x)$, the steady state probability that the amount of data in the buffer exceeds $x$

- Note that $Pr(\max_{0 \leq k \leq n} S_{-k} > x) \rightarrow Pr(X > x)$ as $n \rightarrow \infty$

- $Pr(\max_{0 \leq k \leq n} S_{-k} > x)$

$$= Pr\left( \bigcup_{0 \leq k \leq n} \{S_{-k} > x\} \right)$$

$$\leq \sum_{k=0}^{n} Pr(S_{-k} > x) \quad \text{by union bound}$$

$$\leq \sum_{k=0}^{n} E\left(e^{\theta(S_{-k} - x)}\right) \quad \text{by Chernoff Bound}$$

$$= e^{-\theta x} \sum_{k=0}^{n} E\left(e^{\theta S_{-k}}\right)$$

$$= e^{-\theta x} \sum_{k=0}^{n} e^{-\theta C k} E\left(e^{\theta(A_0 - A_{-k})}\right)$$

$$= e^{-\theta x} \sum_{k=0}^{n} e^{-\theta C k} E\left(e^{\theta(A_k - A_0)}\right) \quad — (1)$$

by stationarity

- let $\Gamma(\theta) \triangleq \lim_{k \rightarrow \infty} \frac{1}{k} \ln E\left(e^{\theta(A_k - A_0)}\right)$ (log moment generating function) if the limit exists

  $\rightarrow$ for large $k$,

  $$E\left(e^{\theta(A_k - A_0)}\right) \approx e^{k\Gamma(\theta)}$$

- If $\Gamma(\theta) < \theta C$, the terms on the RHS of (1) are geometrically decreasing in $k$ for large $k$

  $\rightarrow$ the sum is bounded for $n \rightarrow \infty$ by some finite value $G$

  $\rightarrow$ $Pr(X > x) \leq G e^{-\theta x}$

  ie if the link rate $C$ is greater than $\frac{\Gamma(\theta)}{\theta}$, the probability of exceeding $y$ in steady state has an upper bound that decays exponentially in $y$, with parameter $\theta$

- $e(\theta) = \frac{\Gamma(\theta)}{\theta}$ is called the <u>effective bandwidth</u> of the source for the QoS requirement of $\theta$ on the tail probability decay rate

- <u>Properties of $e(\theta)$</u>:

  a) If $R$ is the peak rate of the source, then $e(\theta) \leq R$

  $(\because \Gamma(\theta) \leq \lim_{k \rightarrow \infty} \frac{1}{k} \ln E(e^{\theta k R})$

  $= \theta R)$

  b) If $r$ is the average source rate, then $e(\theta) \geq r$

  $(\because \Gamma(\theta) \geq \lim_{k \rightarrow \infty} \frac{1}{k} \ln e^{\theta E(A_k - A_0)}$ by Jensen's Inequality

  $= \theta r)$

c) Effective bandwidths of independent sources add

- if $m$ independent sources $A_k^{(i)}$ have log moment generating functions $\Gamma_i(\theta)$, $1 \le i \le m$, the log moment generating function of their superposition $\sum_{i=1}^{m} A_k^{(i)}$ is

$$\Gamma(\theta) = \lim_{k\to\infty} \frac{1}{k} \ln E\left(e^{\theta \sum_{i=1}^{m}(A_k^{(i)} - A_0^{(i)})}\right)$$

$$= \lim_{k\to\infty} \frac{1}{k} \ln \prod_{i=1}^{m} E\left(e^{\theta(A_k^{(i)} - A_0^{(i)})}\right)$$

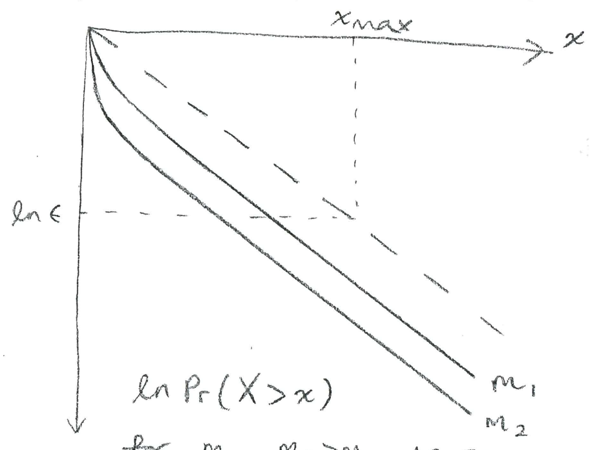
$$= \sum_{i=1}^{m} \lim_{k\to\infty} \frac{1}{k} \ln E\left(e^{\theta(A_k^{(i)} - A_0^{(i)})}\right)$$

$$= \sum_{i=1}^{m} \Gamma_i(\theta)$$

$\therefore$ the effective bandwidth of $\sum_{i=1}^{M} A_k^{(i)}$ is the sum of the effective bandwidths of $A_k^{(i)}$, $1 \le i \le m$

$\rightarrow$ straightforward use of the effective bandwidths can lead to overly conservative resource allocation

- Effective bandwidth requirement that $Pr(X>x)$ decays exponentially with parameter $\theta$ is excessive if we are concerned only with small values of $x$
  - for small $x$ (cf marginal buffering case), the required bandwidth scales sub-linearly in the number of iid sources
  - allocating capacity accord^g to effective bandwidth for an increasing no. of iid sources (linear scaling) leads to an increasingly overprovisioned system
    - typical plot for Markovian sources:




for $m_1$, $m_2 > m_1$ sources when capacity is allocated according to effective bandwith required for $Pr(X > x_{max}) \le \epsilon$

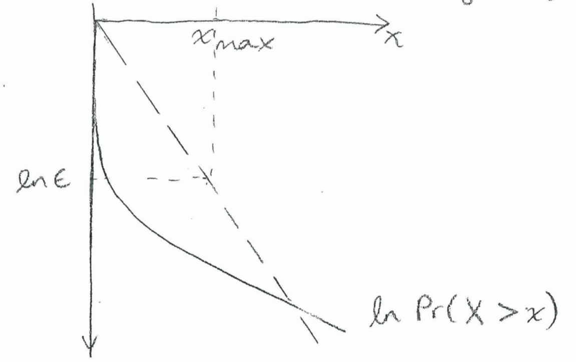
$$= e^{-\theta x_{max}}$$

where $\theta = \dfrac{-\ln \epsilon}{x_{max}}$

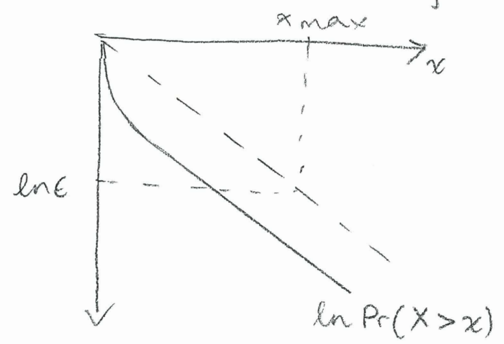
- Combining marginal & arbitrary buffering analysis for capacity allocation (Guérin et al 91)
  - $m$ independent sources with marginals $B_i$ & effective bandwidths $e_i(\theta)$, $1 \le i \le m$
  - QoS requirement $Pr(X > x_{max}) \le \epsilon$
  - allocated capacity $C = \min\{C_0, C_{ebw}\}$ where $C_0$ is obtained by marginal buffering analysis

& $C_{ebw}$ by effective bandwidth analysis

- $C_0$ is the min capacity required to achieve
  $$Pr\left(\sum_{i=1}^{m} B^{(i)} > C_0\right) \le \epsilon$$
- $C_{ebw} = \sum_{i=1}^{m} e^{(i)}(\theta)$ with $\theta = \frac{-\ln \epsilon}{x_{max}}$
- Case 1: $C = C_0$, $C < C_{ebw}$
  - marginal buffering sufficient
  - statistical multiplexing significant

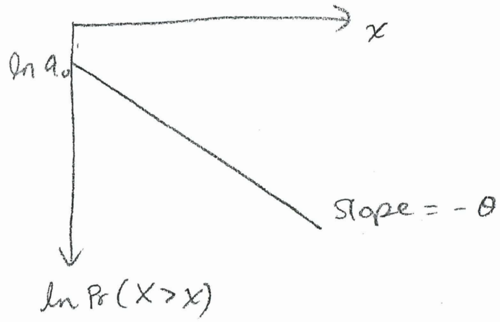



- Case 2: $C = C_{ebw}$
  - statistical multiplexing less significant




- Affine approximation for Markovian sources
  - for a given capacity $C$ between the average & peak arrival rate,
    $$Pr(X > x) \approx a_0 e^{-\theta x}$$
    where $\theta$ satisfies $C = \frac{\Gamma(\theta)}{\theta}$
    & $a_0 \doteq Pr\left(\sum_{i=1}^{m} B^{(i)} > C\right)$ (loss probability w marginal buffering)

$\Rightarrow \ln Pr(X > x) \approx \ln a_0 - \theta x$



Example: random on-off source

- source transmits

$$B_k = \begin{cases} R & \text{with prob } p \\ 0 & \text{with prob } 1-p \end{cases}$$

iid in each timeslot $k$

- marginal buffering:

$$M(\theta) = E(e^{\theta B})$$

$$= pe^{\theta R} + 1-p$$

$$\ell(\gamma) = \sup_{\theta \geq 0} (\theta\gamma - \ln(pe^{\theta R} + 1-p))$$

- arbitrary buffering:

$$\Gamma(\theta) = \lim_{k\to\infty} \frac{1}{k} \ln E(e^{\theta(A_k - A_0)})$$

$$= \lim_{k\to\infty} \frac{1}{k} \ln E(e^{\theta \sum_{i=1}^{k} B_k})$$

$$= \lim_{k\to\infty} \frac{1}{k} \ln \prod_{i=1}^{k} E(e^{\theta B_k})$$

$$= \lim_{k\to\infty} \frac{1}{k} \ln \prod_{i=1}^{k} (pe^{\theta R} + 1-p)$$

$$= \ln(pe^{\theta R} + 1-p)$$