Martingales, function computation in simple broadcast networks

Reading: See website.

Problems to be handed in:

1. Two martingales associated with a simple branching process
Let $Y = (Y_n : n \geq 0)$ denote a simple branching process. Thus, Y_n is the number of individuals in the n^{th} generation, $Y_0 = 1$, the numbers of offspring of different individuals are independent, and each has the same distribution as a random variable X.
(a) Identify a constant θ so that $G_n = \frac{Y_n}{\theta^n}$ is a martingale.
(b) Let $\mathcal{E}$ denote the event of eventual extinction, and let $\alpha = P\{\mathcal{E}\}$. Show that $P[\mathcal{E}|Y_0, \ldots, Y_n] = \alpha^{Y_n}$. Thus, $M_n = \alpha^{Y_n}$ is a martingale.
(c) Using the fact $E[M_1] = E[M_0]$, find an equation for α. (Note: a homework problem in the ECE534 notes shows that α is the smallest positive solution to the equation, and $\alpha < 1$ if and only if $E[X] > 1$.)

2. Function computation in simple broadcast networks (Lei Ying)
Consider the algorithm given in the last two slides of Lei's presentation. Show that there exists positive a_1 and a_2 such that $\lim_{N\to\infty} P\{\text{the fusion center knows all bits correctly}\} = 1$. (The result is in Gallager's paper, but try to prove it using other results in Lei's slides. A given node i belongs to $a_1 \log N$ subsets of size $a_1 \log N$ each. Each node is associated with one subset, and each subset has one node associated with it. The intersection of any two subsets is at most one node. Under the algorithm: (1) each node broadcasts its bit $a_2 \log \log n$ times. (2) Each node estimates the parity of the subset it is associated with, and transmits it ones. (3) The rest is up to the fusion center. Let $\widetilde{b_j}$ be the majority vote estimate of b_j formed at the fusion center after step (1), for $1 \leq j \leq N$. Fix a node i. Let $S_{i,k}$, $1 \leq k \leq a_1 \log N$ denote the subsets containing i, and let $P_{i,k}$ denote the estimate of the parity of the bits in $S_{i,k}$ obtained at the fusion center from step (2). Conditioned on the data bits, all the binary random variables in $(P_{i,k}, (\widetilde{b_j})_{j\in S_{i,k}})_{1\leq k\leq a_1 \log N}$ are mutually independent (why?). Let

$$\overline{b}_{i,k} = \left(P_{i,k} + \sum_{j\in S_{i,k}-\{i\}} \widetilde{b_j} \right) \mod 2$$

The final estimate of b_i is the majority vote of $\overline{b}_{i,k}$.)

3. A covering problem
Consider a linear array of n cells. Suppose that m base stations are randomly placed among the cells, such that the locations of the base stations are independent, and uniformly distributed among the n cell locations. Let r be a positive integer. Call a cell i covered if there is at least one base station at some cell j with $|i - j| \leq r - 1$. Thus, each base station (unless those near the edge of the array) covers $2r - 1$ cells. Note that there can be more than one base station at a given cell, and interference between base stations is ignored.
(a) Let F denote the number of cells covered. Apply the method of bounded differences based on the Azuma-Hoeffding inequality to find an upper bound on $P[|F - E[F]| \geq \gamma]$.
(b) (This part is related to the coupon collector problem and may not have anything to do with martingales.) Rather than fixing the number of base stations, m, let X denote the number of base stations needed until all cells are covered. In case $r = 1$ we have seen that $P[X \geq n \ln n + cn] \to \exp(-e^{-c})$ (the coupon collectors problem). For general $r \geq 1$, find $g_1(r)$ and $g_2(r)$ to so that for any $\epsilon > 0$, $P[X \geq (g_2(r) + \epsilon)n \ln n] \to 0$ and $P[X \leq (g_1(r) - \epsilon)n \ln n] \to 0$. (Ideally you can find $g_1 = g_2$, but if not, it'd be nice if they are close.)

4. Doob decomposition
Suppose $X = (X_k : k \geq 0)$ is an integrable (meaning $E[|X_k|] < \infty$ for each k) sequence adapted to a filtration $\boldsymbol{\mathcal{F}} = (\mathcal{F}_k : k \geq 1)$. (a) Show that there is sequence $B = (B_k : k \geq 0)$ which is predictable relative to $\boldsymbol{\mathcal{F}}$ (which means that B_0 is a constant and B_k is $\mathcal{F}_{k-1}$ measurable for $k \geq 1$) and a mean zero martingale $M = (M_k : k \geq 0)$, such that $X_k = B_k + M_k$ for all k. (b) Are the sequences B and M uniquely determined by X and $\boldsymbol{\mathcal{F}}$?

5. On uniform integrability
(a) Show that if $(X_i : i \in I)$ and $(Y_i : i \in I)$ are both uniformly integrable collections of random variables with the same index set I, then $(Z_i : i \in I)$, where $Z_i = X_i + Y_i$ for all i, is also a uniformly integrable collection. (b) Show that a collection of random variables $(X_i : i \in I)$ is uniformly integrable if and only if there exists a convex increasing function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ with $\lim_{c\to\infty} \frac{\varphi(c)}{c} = +\infty$ and a constant K, such that $E[\varphi(X_i)] \leq K$ for all $i \in I$.