1. We wish to prove that $E$ contains an information set, or that $\bar{E}$ is contained in a check set (an information set of $C^\perp$).

Let $s \neq 0$ be the syndrome that corresponds to $x$. For any check set $B$ there is a vector $e$ such that $He^T = s$, $\text{supp}(e) \subseteq B$.

Hence $|E| \leq n - k$. Moreover, since $x$ is a vector of minimum weight in its coset, there is no other vector $x'$ with the same syndrome such that $\text{supp}(x') \subset \bar{E}$. This implies that $\text{rk}(H(\bar{E})) = |\bar{E}|$.

Hence by Lemma 6a.3

$$0 = |\bar{E}| - \text{rk}(H(\bar{E})) = k - \text{rk}(G(E))$$

as required.

2. (a) is obvious. (b) follows from (a) and the equality

$$\sum_{i=1}^{n} \epsilon_{C}(t) \binom{n}{i} = 2^{n-k}.$$ 

Indeed,

$$2^{n-k} = \sum_{i=1}^{n} \epsilon_{C}(t) \binom{n}{i} \geq \sum_{i=0}^{t} \epsilon_{C}(t) \binom{n}{i} \geq \epsilon_{C}(t) \sum_{i=0}^{t} \binom{n}{i}.$$ 

(c) is implied by observing that for $h_2(p) > n-k$ the right-hand side of the inequality in (b) goes to $0$, so the fraction of correctable errors of relative weight $pm$ is vanishingly small.

3. Part (a) follows on observing that the outer code corrects $2e+x$ errors, so replacing one error by erasure will not increase its correcting capacity.

(b). Let $s$ be the number of different reliability values in the vector $W$. In particular, assume that the first $l_1$ coordinates are erased as a result of inner decoding (so their reliability is $w_1 = 1/2$), the next $l_2$ coordinates all have reliability $w_2$, and so on up to $l_s$. Consider $s$ vectors $W_i = ((1/2)\sum_{j=1}^{l_1} 0^{n_1} - \sum_{j=1}^{l_1} l_i)$. Let $\omega_0 = 1 - 2w_1, \omega_1 = 2(w_{i-1} - w_i), i = 1, \ldots, s-1, \omega_s = 2w_s$. The rest of the proof is the same.

4. The vector $y$ is linearly independent of the code $C$ and therefore is a unique codeword of minimum weight in the code $C' = \langle C, y \rangle$. Therefore running the algorithm on the code $C'$ finds $y$, which is the error vector. This solves the decoding task.

5. See the argument in Lecture 8, Remark 2 (on p.8). Essentially the only thing to be proved is that the error exponent $E_0(R, p)$ behaves as $O(\epsilon^2)$ as $\epsilon \to 0$. This is proved by computing the Taylor series (or just by showing that the first derivative $\frac{\partial}{\partial R} D(\delta_{GV}(R)||p)$ vanishes for $R = C$). Write $\delta$ for $\delta_{GV}(R)$.

$$\frac{\partial D(\parallel p)}{\partial R} = \frac{\log \frac{4(1-p)}{1-\delta p}}{\log \frac{1}{1-\delta}}$$

As $R \to C$, the value $\delta \to p < \frac{1}{2}$. Then the denominator tends to a constant $\log \frac{1-p}{p}$ and the numerator to 0.

6. Linear algebra (write out a parity check symbol computed in two ways and convince yourself that the result is the same).