Solution 1.

(a) The density evolution function for the \((l, r)\)-regular ensemble is

\[
f(x, \epsilon) = \epsilon (1 - (1 - x)^{r-1})^{l-1}
\]

As we have seen in the class, we can plot the fixed points using the parametric expression

\[
\epsilon(x) = \frac{x}{f(x)} = \frac{x}{(1 - (1 - x)^{r-1})^{l-1}}
\]

For the \((2, 4)\)-regular ensemble and \((4, 8)\)-regular ensemble,

\[
\epsilon_{BP}^* = \frac{1}{3}, \quad \epsilon_{BP}^* \approx 0.3834.
\]

For the \((3, 6)\)-regular ensemble and \((5, 10)\)-regular ensemble,

\[
\epsilon_{BP}^* \approx 0.42944, \quad \epsilon_{BP}^* \approx 0.3416.
\]

(b)

(i) The design rate of the code is given by

\[
R = 1 - \frac{\int_0^1 \rho(x) \, dx}{\int_0^1 \lambda(x) \, dx} = \frac{177}{352} \approx 0.503
\]
(ii) Here is the plot of fixed points obtained by plotting
\[ \epsilon(x) = \frac{x}{\lambda (1 - \rho (1 - x))} \]
which shows the threshold is
\[ \epsilon_{\text{BP}}^* \approx 0.47524 \]
We can see that this is indeed higher than the best threshold for the previous regular ensembles (which was \( \approx 0.429 \) for the (3,6)-regular ensemble).

Solution 2.

(a) Recall that if \( W \) is a BEC with erasure probability \( \epsilon \), \( W_s \), \( s \in \{-, +\} \) is BEC with erasure probability \( T_s(\epsilon) \) where
\begin{align*}
T_- (\epsilon) &= 2\epsilon - \epsilon^2 \\
T_+ (\epsilon) &= \epsilon^2
\end{align*}
(1) (2)

Since both of the above functions are increasing in \( \epsilon \), \( \epsilon_1 \leq \epsilon_2 \) implies
\[ T_s(\epsilon_1) \leq T_s(\epsilon_2) \quad \forall s \in \{-, +\} \]
which proves the claim.

(b) The claim follows by induction. Fix a sign sequence \((s_1, \ldots, s_n)\) and define two sequences \( a_\ell \) and \( b_\ell \) as
\begin{align*}
a_0 &= \epsilon_1 \\
a_\ell &= T_{s_\ell}(a_{\ell-1}) \\
b_0 &= \epsilon_2 \\
b_\ell &= T_{s_\ell}(b_{\ell-1})
\end{align*}

Obviously \( a_n \) (respectively \( b_n \)) is the erasure probability of \( W^{s_1, \ldots, s_n}_1 \) (resp. \( W^{s_1, \ldots, s_n}_2 \)). Moreover \( a_0 \leq b_0 \) by assumption and \( a_{\ell-1} \leq b_{\ell-1} \) implies \( a_\ell \leq b_\ell \) due to (a). Thus, \( a_n \leq b_n \) which means \( W^{s_1, \ldots, s_n}_1 \succeq W^{s_1, \ldots, s_n}_2 \).

(c) While the polar code that is being used for communication over the better channel (with erasure probability \( \epsilon^0 < \epsilon \)) is not the optimal one, we know from our analysis that it performs at least as well as what we would have expected if the channel indeed had erasure probability \( \epsilon \) (and perhaps even better because we have strict inequality and \( T_- \) and \( T_+ \) are strictly increasing).
SOLUTION 3. Let us focus on the worst channel. Fix any sign sequence \( s_1, \ldots, s_n \) with \( m \) minuses and \( n - m \) pluses and assume this sequence contains at least one \((+,-)\) pair—i.e., a plus followed by a minus. We claim that by swapping those two signs we get a worst channel. (Hence, the worst channel is indeed the one indexed by a sequence that contains no \((+,-)\) pair — which is exactly the one claimed to index the worst channel.) Suppose \( t_1, \ldots, t_n \) is the sequence obtained by changing one of \((+,-)\) pairs in \( s_1, \ldots, s_n \) to \((-,+).\) More precisely, assume \( s_j = + \) and \( s_{j+1} = - \) and we let

\[
\begin{align*}
t_\ell &= s_\ell, \quad \ell = 1, 2, \ldots, j - 1, j + 2, \ldots, n \\
t_j &= s_{j+1} = - \\
t_{j+1} &= s_j = +.
\end{align*}
\]

(It is obvious that the \( t \) sequence also has \( m \) minuses and \( n - m \) pluses.)

Define two sequences \( a_\ell, b_\ell \) as

\[
\begin{align*}
a_0 &= b_0 = \epsilon \\
a_\ell &= T_{s_\ell}(a_{\ell-1}) \\
b_\ell &= T_{t_\ell}(b_{\ell-1})
\end{align*}
\]

(where \( T_- \) and \( T_+ \) are defined in Equations (1) and (2) previously). Obviously, \( a_n \) is the erasure probability of \( W^{s_1\ldots s_n} \) and \( b_n \) is the erasure probability of \( W^{t_1\ldots t_n}.\)

By construction, we, moreover, have that

\[
a_\ell = b_\ell, \quad \ell = 1, \ldots, \ell - 1
\]

Now,

\[
\begin{align*}
a_j &= T_+(a_{j-1}) = a_{j-1}^2, \\
a_{j+1} &= T_-(a_j) = 2a_j - a_j^2 = 2a_{j-1}^2 - a_{j-1}^4
\end{align*}
\]

while

\[
\begin{align*}
b_j &= T_-(b_{j-1}) = 2b_{j-1} - b_{j-1}^2, \\
b_{j+1} &= T_+(b_j) = b_j^2 = (2b_{j-1} - b_{j-1}^2)^2 = (2a_{j-1}^2 - a_{j-1}^4)^2
\end{align*}
\]

It is easy to check that

\[
(2\alpha^2 - \alpha^4) \leq (2\alpha - \alpha^2)^2, \quad \forall \alpha \in [0, 1]
\]

thus,

\[
a_{j+1} \leq b_{j+1}
\]

Moreover since \( s_\ell = t_\ell \) for \( \ell = j + 2, \ldots, n, \) the monotonicity of \( T_- \) and \( T_+ \) implies

\[
a_\ell \leq b_\ell, \quad \ell = j + 2, \ldots, n.
\]

which, in particular, shows \( a_n \leq b_n. \) Thus, \( W^{s_1\ldots s_n} \geq W^{t_1\ldots t_n}.\)

The claim on the best channels follows exactly in the same way.

SOLUTION 4.

(a) If any of the synthetic channels erase, a block-error event happens. Therefore, the block-error probability of the polar code which uses channels with indices in \( \mathcal{A}_n \) is lower-bounded by the erasure probability of any of those channels — in particular the one with the highest erasure probability.

(b)
(i) Let
\[
L_-(x) := x \\
L_+(x) := x^2.
\]
Then,
\[
a_\ell := T_{s\ell}(a_{\ell-1}) \\
b_\ell := L_{s\ell}(b_{\ell-1})
\]
where \(T_-\) and \(T_+\) are defined in (1) and (2) respectively. Similarly to \(T_\pm\), both transformations \(L_\pm\) are monotone. Moreover, \(L_+(x) = T_+(x)\) while \(L_-(x) \leq T_-(x)\). Therefore we can prove the claim by induction as follows: Clearly \(a_0 \geq b_0\) (in fact here we have equality). Furthermore if \(a_\ell \geq b_\ell\), then,
\[
a_{\ell+1} = T_{s_{\ell+1}}(a_\ell) \geq L_{s_{\ell+1}}(a_\ell) \geq L_{s_{\ell+1}}(b_\ell) = b_{\ell+1}.
\]
(ii) This is obvious by construction:
\[
\log(b_\ell) = 2^1 1^{s_\ell=+} \log(b_{\ell-1}) = 2\sum_{n=1}^{\ell} 1^{s_\ell=+} \log(b_0)
\]
which shows
\[
\log(b_\ell) = 2^p \log(\epsilon) \iff b_\ell = \epsilon^{(2^p)}
\]
(iii) Since \(b_n \leq a_n\), if \(a_n \leq 2^{-2}\beta n\), then we must have
\[
\epsilon^{(2^p)} \leq 2^{-2}\beta n \iff 2^p \log(\epsilon) \leq -2\beta n
\]
Since \(\log(\epsilon) < 0\) (as \(\epsilon < 1\)) the above implies
\[
2^p \geq \frac{2\beta n}{\log(1/\epsilon)} \iff p \geq \beta n - \log(\log(1/\epsilon))
\]
(c)
(i) The expansion of \([z + (1 - z)]^n\) is
\[
1 = [z + (1 - z)]^n = \sum_{i=0}^{n} \binom{n}{i} z^i (1 - z)^{n-i}
\]
Since all the terms in the summation are positive it can be lower-bounded by keeping any of them, say the \(j\)th term:
\[
1 \geq \binom{n}{j} z^j (1 - z)^{n-j}
\]
which proves the claim.
(ii) Taking \(z = j/n\) in (i) we have
\[
\binom{n}{j} \leq \left(\frac{j}{n}\right)^{-j} \left(\frac{n-j}{n}\right)^{(n-j)} = 2^{-n\left[\frac{1}{n} \log\left(\frac{j}{n}\right) + \frac{n-j}{n} \log\left(\frac{n-j}{n}\right)\right]} = 2^{nh_2(j/n)}
\]
(iii) Since $p \geq n/2$,

$$\sum_{j=p}^{n} \binom{n}{j} \leq \sum_{j=p}^{n} \binom{n}{p} \leq (n-p+1)\binom{n}{p} \leq n\binom{n}{p} \leq n2^{nh_2(p/n)}$$

(iv) Using (iii), the fraction of such sequences is upper-bounded by

$$\frac{n2^{nh_2(p/n)}}{2^n} \leq n2^{-n[1-h_2(\beta-c(\epsilon)/n)]}$$

where the inequality follows by lower-bounding $p/n \geq \beta - c(\epsilon)/n$. Since $\beta > \frac{1}{2}$ for large enough $n$, $\beta - c(\epsilon)/n \geq \frac{1}{2}$, which shows the exponent of the right-hand-side of the above is strictly positive (since $h_2(\beta - c(\epsilon)/n) < 1$). Therefore, the right-hand-side of the above goes to 0 exponentially fast as $n \to \infty$.

(d) In (b) we showed that for a synthetic channel to have an erasure probability decaying faster than $2^{-N^{\beta}}$, its ‘index’ must have at least $\beta n$ pluses. In particular if $\beta > \frac{1}{2}$, its index must more than $\frac{1}{2} n$ pluses. As we showed in (c), the fraction of such sign sequences vanishes as $n \to \infty$.

(e) If the block-error probability were to decay faster than $2^{-\sqrt{N}}$, that is $2^{-N^{\beta}}$, $\beta > \frac{1}{2}$, then because of (a) the erasure probability of all synthetic channels used for the transmission of information bits would have to decay as fast as $2^{-N^{\beta}}$, $\beta > \frac{1}{2}$. This means $A_n$ must be a subset of a vanishing fraction of channels because of (d). This requires the code rate to decay to 0.