PROJECT: THE BLAHUT-ARIMOTO ALGORITHM

I. THE ALGORITHM

Let $p(y|x)$ be a discrete memoryless channel with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$. Recall that the capacity of $p(y|x)$ is given by

$$C = \max_{p(x)} I(X;Y). \tag{1}$$

For any distribution $p(x)$ on $\mathcal{X}$ and any conditional distribution $q(x|y)$ on $\mathcal{X} \times \mathcal{Y}$, define

$$f(p(x), q(x|y)) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x)p(y|x) \log \frac{q(x|y)}{p(x)}.$$

Prove the following facts:

1) The conditional distribution $q(x|y)$ that maximizes $f(p(x), q(x|y))$ is given by

$$q^*(x|y) = \frac{p(x)p(y|x)}{\sum_{x' \in \mathcal{X}} p(x')p(y|x')}.$$

2) The distribution $p(x)$ that maximizes $f(p(x), q(x|y))$ is given by

$$p^*(x) = \frac{\prod_{y \in \mathcal{Y}} q(x|y)^{p(y|x)}}{\sum_{x' \in \mathcal{X}} \prod_{y \in \mathcal{Y}} q(x'|y)^{p(y|x')}}.$$

3) The capacity of $p(y|x)$ can be expressed as

$$C = \max_{p(x)} \max_{q(x|y)} f(p(x), q(x|y)).$$

The above facts naturally suggest the following iterative algorithm (known as the Blahut-Arimoto algorithm) for computing the channel capacity as well as the capacity-achieving input distribution. Let $p^{(0)}(x) = \frac{1}{|\mathcal{X}|}$ for all $x \in \mathcal{X}$. For $k \geq 0$, let

$$q^{(k)}(x|y) = \frac{p^{(k)}(x)p(y|x)}{\sum_{x' \in \mathcal{X}} p^{(k)}(x')p(y|x')},$$

$$p^{(k+1)}(x) = \frac{\prod_{y \in \mathcal{Y}} q^{(k)}(x|y)^{p(y|x)}}{\sum_{x' \in \mathcal{X}} \prod_{y \in \mathcal{Y}} q^{(k)}(x'|y)^{p(y|x')}}.$$

It can be shown that $p^{(k)}(x)$ converges to a capacity-achieving distribution $p^{\dagger}(x)$ (i.e., $p^{\dagger}(x)$ is an optimal solution to the maximization problem in (1)) as $k \to \infty$.

Remark: You are encouraged to consult the relevant literature on the Blahut-Arimoto algorithm.

II. ADDITIVE GAUSSIAN NOISE CHANNEL WITH PEAK POWER CONSTRAINT

Consider an additive Gaussian noise channel $Y = X + Z$, where $Z \sim \mathcal{N}(0,1)$ is independent of X. Use the Blahut-Arimoto algorithm to compute the capacity-achieving distribution of this channel subject to the peak power constraint $|X| \leq A$ for $A = 0.1$, $A = 1$, and $A = 10$. Record and explain your findings.

Remark: You may need to discretize the input alphabet and the output alphabet in order to apply the Blahut-Arimoto algorithm.