Research Results

November 18, 2009

Contents

List of Figures

Acronyms Used In This Document

MC Monte Carlo

AV antithetic variables

PAV parameterized antithetic variables

CRLB Cramer Rao Lower Bound

PDF probability density function

MVU minimum variance unbiased

1 Single Link

We consider the case where a single source is transmitting a sequence of packets to a single destination (Figure 1) and the source records the outcome ($X_1^{(n)}$ of the packet. Examples of outcomes are loss information or delay information; we consider loss based tomography. In this case, the outomes are Bernoulli random variables, with $X_1^{(n)} = 1$ with probability α denoting successful reception of the packet. We are interested in developing an unbiased estimator for the link success probability α with a low variance.

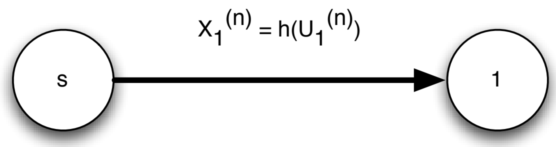


Figure 1: Single Link

1.1 Loss Base Tomography

1.1.1 Monte Carlo Estimation

The simplest estimation technique is Monte Carlo (MC). We generate a sequence of unifrom random variables $U_1^{(n)}$ and transforming them to Bernoulli random variables $X_1^{(n)} = h(U_1^{(n)}) = \mathbb{I}_{U_1^{(n)} < \alpha}$. The MC estimator is then

$$\hat{\alpha}_{MC} = \frac{1}{n} \sum_{i=1}^{n} X_1^{(i)} \tag{1}$$

This is clearly unbiased and has a normalized variance of

$$n \operatorname{Var}(\hat{\alpha}_{MC}) = \alpha(1-\alpha) \tag{2}$$

The variance as a function of α is shown in Figure 2. The variance (or mean squared error since the estimator is ubiased) is maximum at $\alpha = \frac{1}{2}$.

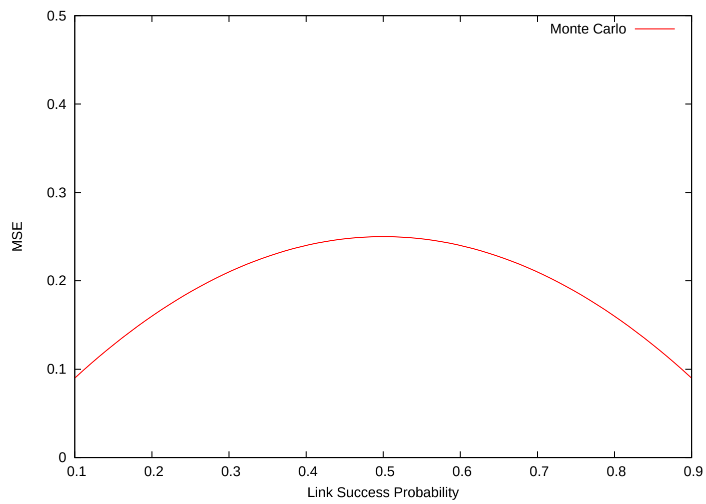


Figure 2: Normalized variance for Monte Carlo (MC) as a function of link success probability, α

1.1.2 Antithetic Variables

Our first step towards improving over the simple MC estimation is the use of antithetic variables (AV)[3]. The idea here is that instead of generating a sequence with n random variables, we generate a sequence of length $\frac{n}{2}$ and form the estimator

$$\hat{\alpha}_{AV} = \frac{1}{n} \sum_{i=1}^{n/2} h(U_1^{(i)}) + h(1 - U_1^{(i)}) \tag{3}$$

Because expectation is linear and because $h(U_1^{(n)})$ and $h(1-U_1^{(i)})$ are identically distributed, it is easy enough to verify that the estimator is unbiased. The normalized variance is given as

$$
\begin{aligned}
n \operatorname{Var}(\hat{\alpha}_{AV}) &= \frac{1}{n} \sum_{i=1}^{n/2} \operatorname{Var}\left(h(U_1^{(i)}) + h(1-U_1^{(i)})\right) \\
&= \frac{\operatorname{Var}\left(h(U_1^{(i)})\right) + \operatorname{Var}\left(h(1-U_1^{(i)})\right) + 2\operatorname{Cov}\left(h(U_1^{(i)}), h(1-U_1^{(i)})\right)}{2} \\
&= \operatorname{Var}\left(h(U_1^{(i)})\right) + \operatorname{Cov}\left(h(U_1^{(i)}), h(1-U_1^{(i)})\right) \\
&= \mathbb{E}\left[h(U_1^{(i)})^2\right] + \mathbb{E}\left[h(U_1^{(i)})h(1-U_1^{(i)})\right] - 2\mathbb{E}\left[h(U_1^{(i)})\right]^2
\end{aligned}
$$

$$
n \operatorname{Var}(\hat{\alpha}_{AV}) = \alpha + 2\mathbb{I}_{\alpha > \frac{1}{2}}(\alpha - \frac{1}{2}) - 2\alpha^2 \tag{4}
$$

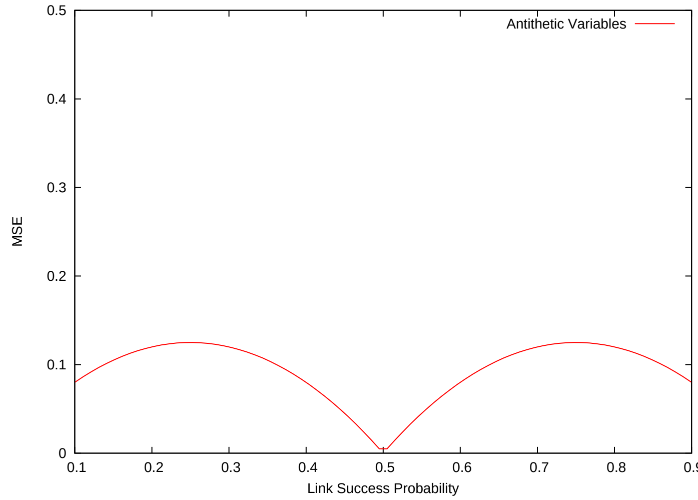


Figure 3: Normalized variance for antithetic variables (AV) as a function of link success probability, α

The normalized variance as a function of α is shown in Figure 3. Note that the normalized variance is maximum at $\alpha = \frac{1}{4}$ and $\alpha = \frac{3}{4}$. From this figure we observe:

Proposition 1.

$$
\operatorname{Var}(\hat{\alpha}_{AV}) < \operatorname{Var}(\hat{\alpha}_{MC}), \forall \alpha \in (0,1)
$$

Proof. First, consider when $\alpha < \frac{1}{2}$

$$
\alpha - 2\alpha^2 < \alpha(1-\alpha)
$$

$$
\alpha - 2\alpha^2 < \alpha - \alpha^2
$$

$$
0 < \alpha^2
$$

Next consider when $\alpha > \frac{1}{2}$

$$
3\alpha - 1 - 2\alpha^2 < \alpha - \alpha^2
$$

$$
0 < \alpha^2 - 2\alpha + 1
$$

$$
0 < (\alpha - 1)^2
$$

Finally, when $\alpha = \frac{1}{2}$ we have $\operatorname{Var}(\hat{\alpha}_{AV}) = 0$ and $\operatorname{Var}(\hat{\alpha}_{MC}) = \frac{1}{4}$ $\square$

Using AV, we have the greatest variance reduction when the variance of the MC estimator is at its maximum. The closer α is to either 0 or 1, there is less variance reduction.

1.1.3 Parameterized Antithetic Variables

We wish to develop an estimator with a tunable parameter β such that the variance reduction is maximized when $\beta = \alpha$, but which is an unbiased estimator for all α and all β.

Consider α fixed and unknown and to be estimated, and $\beta \in [0,1]$ chosen based on some prior assumptions about α. Generate $U_1^{(n)} \sim \text{Uni}(0,\beta)$ and from there set

$$\bar{U}_1^{(n)} = \frac{1-\beta}{\beta} U_1 + \beta. \tag{5}$$

Note that $\bar{U}_1^{(n)} \sim \text{Uni}(\beta, 1)$. We consider an estimator of the form

$$\hat{\alpha}_{PAV} = \frac{1}{n} \sum_{i=1}^{n/2} \left(\beta \left(\mathbf{1}_{U_1^{(i)} \le \alpha} + \mathbf{1}_{\beta - U_1^{(i)} \le \alpha} \right) + (1-\beta) \left(\mathbf{1}_{\bar{U}_1^{(i)} \le \alpha} + \mathbf{1}_{1 - \bar{U}_1^{(i)} + \beta < \alpha} \right) \right) \tag{6}$$

Then it is straightforward to show

$$\begin{aligned}
\mathbb{E}[\hat{\alpha}_{PAV}] &= \frac{1}{2} \left[\beta \left(\mathbb{P}(U_1 \le \alpha) + 1 - \mathbb{P}(U_1 \le \beta - \alpha) \right) + (1-\beta) \left(\mathbb{P}(\bar{U}_1 \le \alpha) + 1 - \mathbb{P}(\bar{U}_1 \le 1 + \beta - \alpha) \right) \right] \\
&= \begin{cases} \frac{1}{2} \left[\beta \left(\frac{\alpha}{\beta} + 1 - \frac{\beta - \alpha}{\beta} \right) + (1-\beta)(0 + 1 - 1) \right], & \alpha \le \beta \\ \frac{1}{2} \left[\beta (1 + 1 - 0) + (1-\beta) \left(\frac{\alpha - \beta}{1-\beta} + 1 - \frac{1 + \beta - \alpha - \beta}{1-\beta} \right) \right], & \alpha > \beta \end{cases} \\
&= \alpha
\end{aligned} \tag{7}$$

We can then find the normalized variance as

$$n \operatorname{Var}(\hat{\alpha}_{PAV}) = \begin{cases}
(\alpha - \beta)(\beta - (2\alpha - 1)), & \alpha \le \frac{1}{2} \quad \beta \le \alpha \\
(\beta - \alpha)(2\alpha - \beta), & \qquad\quad \alpha < \beta \le 2\alpha \\
\alpha(\beta - 2\alpha), & \qquad\quad \beta > 2\alpha \\
(1-\alpha)((2\alpha - 1) - \beta), & \alpha > \frac{1}{2} \quad \beta \le \alpha \quad \beta \le 2\alpha - 1 \\
(\alpha - \beta)(\beta - (2\alpha - 1)), & \qquad\quad \beta \le \alpha \quad 2\alpha - 1 < \beta \\
(\beta - \alpha)(2\alpha - \beta), & \qquad\quad \beta > \alpha
\end{cases} \tag{8}$$

The normalized variance for both AV and parameterized antithetic variables (PAV) is shown in Figure 4.

Proposition 2.

$$\operatorname{Var}(\hat{\alpha}_{PAV}) < \operatorname{Var}(\hat{\alpha}_{MC}), \forall \alpha, \beta \in (0,1)$$

Proof. If $\alpha < \frac{1}{3}$, then $\max_\beta \operatorname{Var}(\hat{\alpha}_{PAV}) = \alpha - 2\alpha^2$; we are doing no worse than AV. We proved in Proposition 1 that AV is always better than MC. If $\frac{1}{3} < \alpha < \frac{1}{2}$, then $\max_\beta \operatorname{Var}(\hat{\alpha}_{PAV}) = \dfrac{(1-\alpha)^2}{4}$.

$$\frac{(1-\alpha)^2}{4} < \alpha(1-\alpha)$$

$$\frac{(1-\alpha)}{4} < \alpha$$

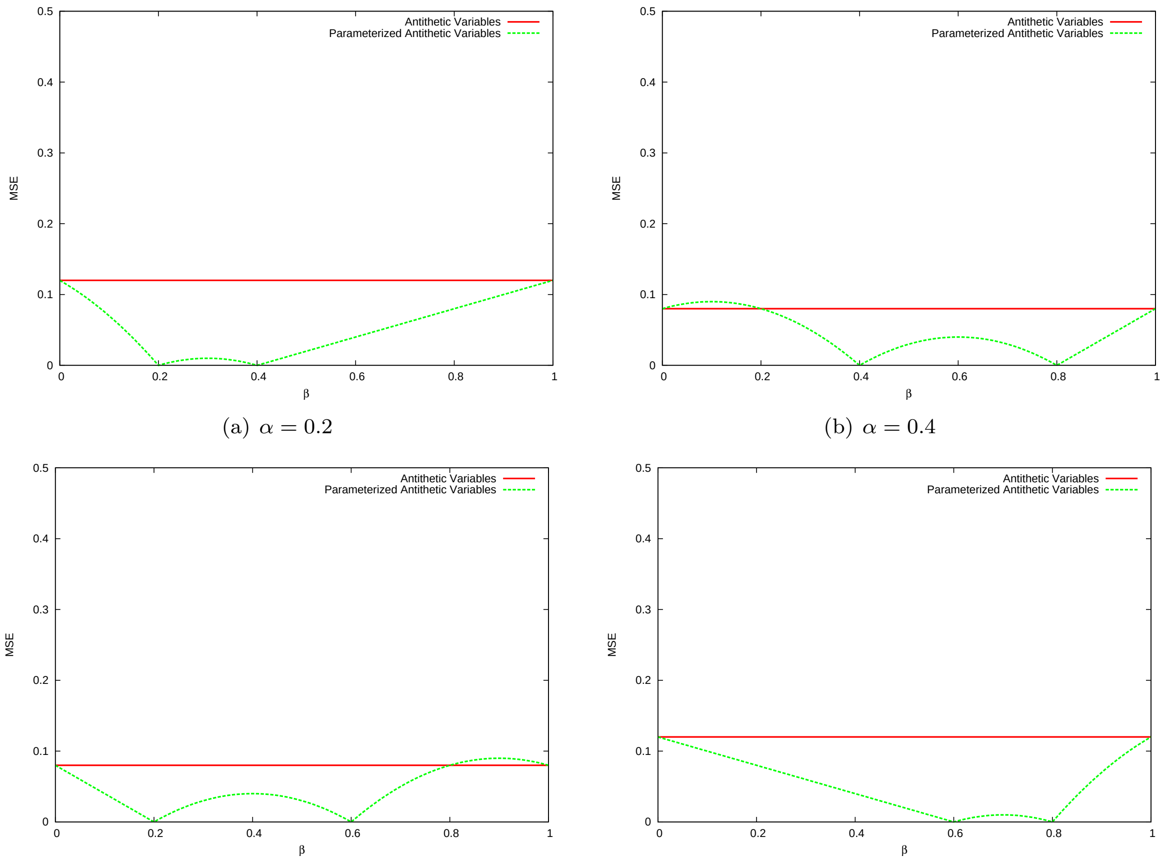
$$1 - \alpha < 4\alpha < 4\frac{1}{2}$$

$$1 - \alpha < 2$$

Because $\alpha \in [0,1]$, the above is true and we conclude that $\operatorname{Var}(\hat{\alpha}_{PAV}) < \operatorname{Var}(\hat{\alpha}_{AV})$ for $\alpha < \frac{1}{2}$.

If $\alpha > \frac{2}{3}$, then $\max_\beta \operatorname{Var}(\hat{\alpha}_{PAV}) = \alpha + 2(\alpha - \dfrac{1}{2}) - 2\alpha^2$; again we are doing no worse than AV. If $\frac{1}{3} < \alpha < \frac{1}{2}$, then $\max_\beta \operatorname{Var}(\hat{\alpha}_{PAV}) = \dfrac{\alpha^2}{4}$.

$$\frac{\alpha^2}{4} < \alpha(1-\alpha)$$



(a) $\alpha = 0.2$ (b) $\alpha = 0.4$

(c) $\alpha = 0.6$ (d) $\alpha = 0.8$

Figure 4: Normalized variance for parameterized antithetic variables (PAV): (a) $\alpha = 0.2$, (b) $\alpha = 0.4$, (c) $\alpha = 0.6$, and (d) $\alpha = 0.8$

$$\frac{\alpha}{4} < 1 - \alpha$$
$$\alpha < 4(1-\alpha) < 4\frac{1}{2}$$
$$\alpha < 2$$

Because $\alpha \in [0,1]$, the above is true and we conclude that $\mathrm{Var}\left(\hat{\alpha}_{PAV}\right) < \mathrm{Var}\left(\hat{\alpha}_{AV}\right)$ for $\alpha > \frac{1}{2}$.

$\square$

Corollary 1. *If $\alpha < \frac{1}{3}$ or $\alpha > \frac{2}{3}$, then*

$$\mathrm{Var}\left(\hat{\alpha}_{PAV}\right) < \mathrm{Var}\left(\hat{\alpha}_{AV}\right), \forall \beta \in (0,1)$$

Corollary 2. *If $\beta \geq 3\alpha - 1,\ \alpha \in (\frac{1}{3}, \frac{1}{2})$ or $\beta \leq 3\alpha - 1,\ \alpha \in (\frac{1}{2}, \frac{2}{3})$, then*

$$\mathrm{Var}\left(\hat{\alpha}_{PAV}\right) < \mathrm{Var}\left(\hat{\alpha}_{AV}\right)$$

Corollary 3. *If $\alpha \in (\frac{1}{3}, \frac{1}{2})$ and $\beta < 3\alpha - 1$, then the worst case choice of β is*

$$\frac{3\alpha - 1}{2}$$

which gives

$$\mathrm{Var}\left(\hat{\alpha}_{PAV}\right) = \left(\frac{1-\alpha}{2}\right)^2$$

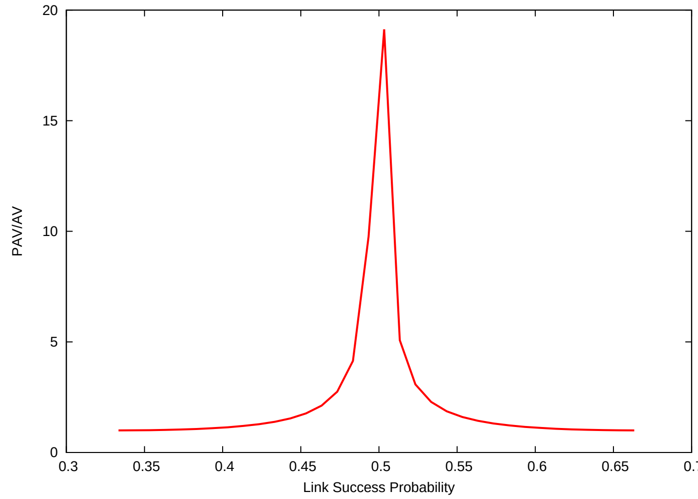


Figure 5: Ratio of parameterized antithetic variables (PAV) to antithetic variables (AV) as a function of link success probability, α

If $\alpha \in (\frac{1}{2}, \frac{2}{3})$ and $\beta > 3\alpha - 1$, then the worst case choice of β is

$$\frac{3\alpha}{2}$$

which gives

$$\mathrm{Var}\left(\hat{\alpha}_{PAV}\right) = \left(\frac{\alpha}{2}\right)^2$$

The ratio of PAV variance to AV variance is plotted in Figure 5.

1.1.4 Cramer Rao Lower Bounds

Monte Carlo If we let $m = \sum_{i=1}^{n} X_1^{(i)}$ denote the number of successful packet receptions, we can then write

$$p(\mathbf{x}; \alpha) = \prod_{i=1}^{m} \alpha \prod_{i=m+1}^{n} (1 - \alpha) \tag{9}$$

We first check to make sure the assumed probability density function (PDF) meets the condition that

$$\mathbb{E}\left[\frac{\partial \ln p(\mathbf{x}; \alpha)}{\partial \alpha}\right] = 0, \ \forall \alpha \tag{10}$$

This is known as the "regularity" condition.

$$\mathbb{E}\left[\frac{\partial\left(m \ln \alpha + (m-n)\ln(1-\alpha)\right)}{\partial \alpha}\right] = \mathbb{E}\left[\frac{m}{\alpha} + \frac{(m-n)}{1-\alpha}\right]$$
$$\frac{1}{\alpha}\mathbb{E}\left[m\right] + \frac{1}{1-\alpha}\mathbb{E}\left[m\right] - \frac{N}{1-\alpha}$$
$$\frac{\mathbb{E}[m]}{\alpha(1-\alpha)} - \frac{N}{1-\alpha} = \frac{N}{1-\alpha} - \frac{N}{1-\alpha}$$

The assumed PDF meets the regularity condition and therefore the variance of any unbiased estimator $\hat{\alpha}$ must satisfy

$$\mathrm{Var}\left(\hat{\alpha}\right) \geq \frac{1}{-\mathbb{E}\left[\frac{\partial^2 \ln p(\mathbf{x};\alpha)}{\partial \alpha^2}\right]} \tag{11}$$

$$\mathbb{E}\left[\frac{\partial^2 \ln p(\mathbf{x};\alpha)}{\partial\alpha^2}\right] = \mathbb{E}\left[-\frac{m}{\alpha^2} + \frac{(m-n)}{(1-\alpha)^2}\right]$$

$$-\frac{n}{\alpha} + \frac{n\alpha}{(1-\alpha)^2)} - \frac{n}{(1-\alpha)^2)} = -\frac{n}{\alpha(1-\alpha)}$$

$$\mathrm{Var}\left(\hat{\alpha}\right) \geq \frac{\alpha(1-\alpha)}{n} \tag{12}$$

Because the sample average attains this variance, we know that it is a minimum variance unbiased (MVU) estimator for α. Although this result my seem inconsistent with our results above for AV and PAV, it is not. The reason is that an estimator can have variance no lower than (12) for the assumed PDF in (9) - the assumed PDF is different for the AV and PAV and hence the variance for an unbiased estimator may be lower.

Antithetic Variables If we let

$$Z_i = \frac{h(U_1^{(i)}) + h(1-U_1^{(i)})}{2}$$

then we can rewrite (3) as

$$\hat{\alpha}_{AV} = \frac{2}{n}\sum_{i=1}^{n/2} Z_i$$

The PDF for Z_i is

$$\mathbb{P}[Z_i = z] = \begin{cases} (1-2\alpha)^+ & z = 0 \\ 2\min(\alpha, 1-\alpha) & z = \frac{1}{2} \\ (2\alpha-1)^+ & z = 1 \end{cases} \tag{13}$$

We can now calculate the Cramer Rao Lower Bound (CRLB) for the antithetic variable case. If $\alpha < \frac{1}{2}$ and we let m be the number of $Z_i = 0$, then

$$p(\mathbf{z};\alpha) = (1-2\alpha)^m(2\alpha)^{(\frac{n}{2}-m)}$$

$$\ln p(\mathbf{z};\alpha) = m\ln(1-2\alpha) + (\frac{n}{2}-m)\ln(2\alpha)$$

$$\frac{\partial \ln p(\mathbf{x};\alpha)}{\partial\alpha} = \frac{-2m}{1-2\alpha} + (\frac{n}{2}-m)\frac{1}{\alpha}$$

$$\mathbb{E}\left[\frac{\partial \ln p(\mathbf{x};\alpha)}{\partial\alpha}\right] = \frac{-2\mathbb{E}[m]}{1-2\alpha} + (\frac{n}{2}-\mathbb{E}[m])\frac{1}{\alpha}$$

$$\mathbb{E}[m] = (1-2\alpha)\frac{n}{2}$$

$$\mathbb{E}\left[\frac{\partial \ln p(\mathbf{x};\alpha)}{\partial\alpha}\right] = 0$$

The PDF meets the regularity conditions.

$$\frac{\partial^2 \ln p(\mathbf{x};\alpha)}{\partial\alpha^2} = \frac{-4m}{(1-2\alpha)^2} + (\frac{n}{2}-m)\frac{-4}{(2\alpha)^2}$$

$$\mathbb{E}\left[\frac{\partial^2 \ln p(\mathbf{x};\alpha)}{\partial\alpha^2}\right] = \frac{-2n}{(1-2\alpha)(2\alpha)}$$

The CRLB then is

$$n\,\mathrm{Var}\left(\hat{\alpha}_{AV}\right) \geq \alpha - 2\alpha^2 \tag{14}$$

Comparing (14) with (4) we see that the CRLB is attained. It can be shown in a similar manner that the bound is achieved when $\alpha > \frac{1}{2}$. Therefore (3) is efficient.

Parameterized Anithetic Variables

I need to find the assumed PDF for PAV before I can calculate the CRLB.

1.2 Delay Based Tomography

2 Two Links

We extend the previous analysis as before but now for the two link case.

2.1 Monte Carlo Estimation

We generate two sequences of unifrom random variables $U_1^{(n)}$ and $U_2^{(n)}$ and transforming them to Bernoulli random variables $X_i^{(n)} = h(U_i^{(n)}) = \mathbb{I}_{U_i^{(n)} < \alpha_i}$. The Monte Carlo estimator is then

$$\hat{\alpha}_{MC} = \frac{1}{n}\sum_{i=1}^{n} X_1^{(i)} X_2^{(i)} \tag{15}$$

This is clearly unbiased and has a normalized variance of

$$n \operatorname{Var}(\hat{\alpha}_{MC}) = \alpha_1\alpha_2(1-\alpha_1\alpha_2) \tag{16}$$

2.2 Antithetic Variables

The antithetic variable estimator is now

$$\hat{\alpha}_{AV} = \frac{1}{n}\sum_{i=1}^{n/2} h(U_1^{(i)})h(U_2^{(i)}) + h(1-U_1^{(i)})h(1-U_2^{(i)}) \tag{17}$$

As before, because expectation is linear and because $h(U_i^{(n)})$ and $h(1-U_i^{(n)})$ are identically distributed, it is easy enough to verify that the estimator is unbiased. For ease of notation, we replace $h(U_i^{(i)})h(U_2^{(i)})$ with $g(\mathbf{U}^{(i)})$. The normalized variance is given as

$$\begin{aligned}
n \operatorname{Var}(\hat{\alpha}_{AV}) &= \frac{1}{n}\sum_{i=1}^{n/2} \operatorname{Var}\left(g(\mathbf{U}^{(i)}) + g(1-\mathbf{U}^{(i)})\right) \\
&= \frac{\operatorname{Var}\left(g(\mathbf{U}^{(i)})\right) + \operatorname{Var}\left(g(1-\mathbf{U}^{(i)})\right) + 2\operatorname{Cov}\left(g(\mathbf{U}^{(i)}), g(1-\mathbf{U}^{(i)})\right)}{2} \\
&= \operatorname{Var}\left(g(\mathbf{U}^{(i)})\right) + \operatorname{Cov}\left(g(\mathbf{U}^{(i)}), g(1-\mathbf{U}^{(i)})\right) \\
&= \mathbb{E}\left[g(\mathbf{U}^{(i)})^2\right] + \mathbb{E}\left[g(\mathbf{U}^{(i)})g(1-\mathbf{U}^{(i)})\right] - 2\mathbb{E}\left[g(\mathbf{U}^{(i)})\right]^2 \\
&= \alpha_1\alpha_2 + 4\mathbb{I}_{\alpha_1>\frac{1}{2},\alpha_2>\frac{1}{2}}(\alpha_1 - \frac{1}{2})(\alpha_2 > \frac{1}{2}) - 2\alpha_1^2\alpha_2^2
\end{aligned}$$

2.3 Parameterized Antithetic Variables

When we extend the proposed technique of parameterized antithetic variables, we obtain the following estimator

$$\hat{\alpha}_{PAV} = \frac{1}{n}\sum_{i=1}^{n/2} \beta\left(\mathbf{1}_{U_1^{(i)}\le\alpha_1, U_2^{(i)}\le\alpha_2} + \mathbf{1}_{\beta-U_1^{(i)}\le\alpha_1, \beta-U_2^{(i)}\le\alpha_2}\right) + \tag{18}$$

$$(1-\beta)\left(\mathbf{1}_{\bar{U}_1^{(i)}\le\alpha_1, \bar{U}_2^{(i)}\le\alpha_2} + \mathbf{1}_{1-\bar{U}_1^{(i)}+\beta<\alpha_1, 1-\bar{U}_2^{(i)}+\beta<\alpha_2}\right) \tag{19}$$

Unfortunately, this is a biased estimator. The bias for this particular estimator is dependent upon both the choice of β and α and so it can not be adjusted for.

References

[1] Steven M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.

[2] Jian Ni and Sekhar Tatikonda. Network tomography based on additive metrics, 2008.

[3] Sheldon M. Ross. *Simulation*. Academic Press, fourth edition, 2006.