Channel capacity estimation using free probability theory

Øyvind Ryan and Merouane Debbah

January 2008

Problem at hand

The capacity per receiving antenna of a channel with $n \times m$ channel matrix $\mathbf{H}$ and signal to noise ratio $\rho = \frac{1}{\sigma^2}$ is given by

$$C = \frac{1}{n} \log_2 \det \left(\mathbf{I}_n + \frac{1}{m\sigma^2} \mathbf{H}\mathbf{H}^H \right) = \frac{1}{n} \sum_{l=1}^{n} \log_2 (1 + \frac{1}{\sigma^2} \lambda_l) \quad (1)$$

where λ_l are the eigenvalues of $\frac{1}{m}\mathbf{H}\mathbf{H}^H$. We would like to estimate C.

To estimate C, we will use free probability tools to estimate the eigenvalues of $\frac{1}{m}\mathbf{H}\mathbf{H}^H$ based on some observations $\hat{\mathbf{H}}_i$

Observation model 1

The following is a much used observation model:

$$\hat{\mathbf{H}}_i = \mathbf{H} + \sigma \mathbf{X}_i \tag{2}$$

where

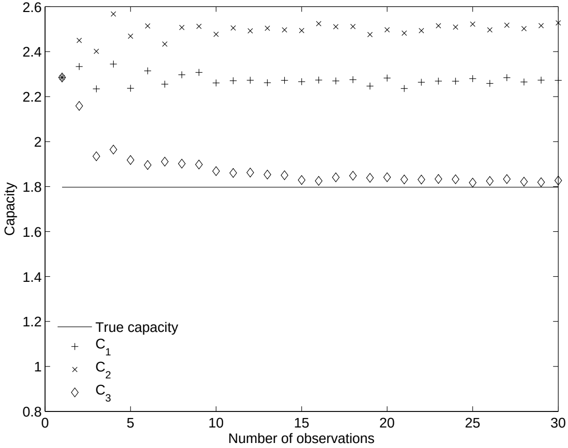
- The matrices are $n \times m$ (n is the number of receiving antennas, m is the number of transmitting antennas)
- $\hat{\mathbf{H}}_i$ is the measured MIMO matrix,
- $\mathbf{X}_i$ is the noise matrix with i.i.d standard complex Gaussian entries.

Existing ways to estimate the channel capacity

Several channel capacity estimators have been used in the literature:

$$
\begin{aligned}
C_1 &= \frac{1}{nL}\sum_{i=1}^{L}\log_2\det\left(\mathbf{I}_n + \frac{1}{m\sigma^2}\hat{\mathbf{H}}_i\hat{\mathbf{H}}_i^H\right) \\
C_2 &= \frac{1}{n}\log_2\det\left(\mathbf{I}_n + \frac{1}{L\sigma^2 m}\sum_{i=1}^{L}\hat{\mathbf{H}}_i\hat{\mathbf{H}}_i^H\right) \\
C_3 &= \frac{1}{n}\log_2\det\left(\mathbf{I}_n + \frac{1}{\sigma^2 m}\left(\frac{1}{L}\sum_{i=1}^{L}\hat{\mathbf{H}}_i\right)\left(\frac{1}{L}\sum_{i=1}^{L}\hat{\mathbf{H}}_i\right)^H\right)
\end{aligned}
\tag{3}
$$

Why not try to formulate an estimator based on free probability instead?



Comparison of the classical capacity estimators for various number of observations. $\sigma^2 = 0.1$, $n = 10$ receive antennas, $m = 10$ transmit antennas. The rank of $\mathbf{H}$ was 3.

The Marčhenko Pastur law

The Marčhenko Pastur law μ_c:

$$f^{\mu_c}(x) = (1 - \frac{1}{c})^+ \delta_0(x) + \frac{\sqrt{(x-a)^+(b-x)^+}}{2\pi c x}, \qquad (4)$$

where $(z)^+ = \max(0, z)$, $a = (1-\sqrt{c})^2$, $b = (1+\sqrt{c})^2$, and $\delta_0(x)$ is dirac measure (point mass) at 0.

- free cumulants: $1, c, c^2, c^3, \ldots$.
- μ_c is the limit eigenvalue distribution of $\frac{1}{N}\mathbf{X}\mathbf{X}^H$, with $\mathbf{X}$ an $n \times N$ with independent standard complex Gaussian entries as $N \to \infty$, and $\frac{n}{N} \to c$.

Main free probability result we will use

Define

$$
\begin{aligned}
\Gamma_n &= \frac{1}{N}\mathbf{R}_n\mathbf{R}_n^H \\
\mathbf{W}_n &= \frac{1}{N}(\mathbf{R}_n + \sigma\mathbf{X}_n)(\mathbf{R}_n + \sigma\mathbf{X}_n)^H,
\end{aligned}
$$

where $\mathbf{R}_n$ and $\mathbf{X}_n$ are independent $n \times N$ random matrices, $\mathbf{X}_n$ is complex, standard, Gaussian.

Theorem

If e.e.d.$(\Gamma_n) \rightarrow \nu_\Gamma$, then e.e.d.$(\mathbf{W}_n) \rightarrow \nu_W$ where ν_W is uniquely identified by

$$\nu_W \boxslash \mu_c = (\nu_\Gamma \boxslash \mu_c) \boxplus \delta_{\sigma^2}$$

($\boxslash$ = "the opposite of $\boxtimes$").

Realization of the theorem for the problem at hand

Form the compound observation matrix

$$\begin{aligned}
\hat{\mathbf{H}}_{1\ldots L} &= \mathbf{H}_{1\ldots L} + \frac{\sigma}{\sqrt{L}}\mathbf{X}_{1\ldots L}, \text{ where} \\
\hat{\mathbf{H}}_{1\ldots L} &= \frac{1}{\sqrt{L}}\left[\hat{\mathbf{H}}_1, \hat{\mathbf{H}}_2, \ldots, \hat{\mathbf{H}}_L\right], \\
\mathbf{H}_{1\ldots L} &= \frac{1}{\sqrt{L}}\left[\mathbf{H}, \mathbf{H}, \ldots, \mathbf{H}\right], \\
\mathbf{X}_{1\ldots L} &= \left[\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_L\right].
\end{aligned}$$

For the problem at hand, the theorem takes the form

$$\nu_{\frac{1}{m}\hat{\mathbf{H}}_{1\ldots L}\hat{\mathbf{H}}_{1\ldots L}^H} \boxslash \mu_{\frac{n}{mL}} \approx \left(\nu_{\frac{1}{m}\mathbf{H}_{1\ldots L}\mathbf{H}_{1\ldots L}^H} \boxslash \mu_{\frac{n}{mL}}\right) \boxplus \delta_{\sigma^2} \qquad (5)$$

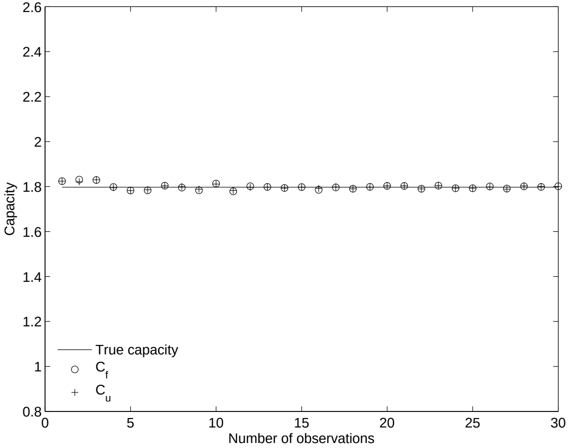
Since $\frac{1}{m}\mathbf{H}_{1\ldots L}\mathbf{H}_{1\ldots L}^H = \frac{1}{m}\mathbf{H}\mathbf{H}^H$, we can now estimate the moments of $\frac{1}{m}\mathbf{H}\mathbf{H}^H$ from the moments of the observation matrix $\frac{1}{m}\hat{\mathbf{H}}_{1\ldots L}\hat{\mathbf{H}}_{1\ldots L}^H$, and thereby estimate the eigenvalues, and hence the channel capacity.

Free probability based estimator for the moments of the channel matrix

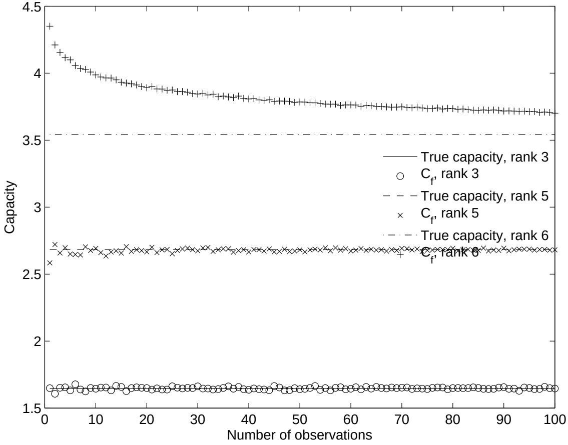
Can also be written in the following way for the first four moments:

$$
\begin{aligned}
\hat{h}_1 &= h_1 + \sigma^2 \\
\hat{h}_2 &= h_2 + 2\sigma^2(1+c)h_1 + \sigma^4(1+c) \\
\hat{h}_3 &= h_3 + 3\sigma^2(1+c)h_2 + 3\sigma^2 c h_1^2 \\
&\quad +3\sigma^4\left(c^2+3c+1\right)h_1 \\
&\quad +\sigma^6\left(c^2+3c+1\right) \\
\hat{h}_4 &= h_4 + 4\sigma^2(1+c)h_3 + 8\sigma^2 c h_2 h_1 \\
&\quad +\sigma^4(6c^2+16c+6)h_2 \\
&\quad +14\sigma^4 c(1+c)h_1^2 \\
&\quad +4\sigma^6(c^3+6c^2+6c+1)h_1 \\
&\quad +\sigma^8\left(c^3+6c^2+6c+1\right),
\end{aligned}
\tag{6}
$$

where $\hat{h}_i$ are the moments of the observation matrix $\frac{1}{m}\hat{\mathbf{H}}_{1\ldots L}\hat{\mathbf{H}}_{1\ldots L}^H$, h_i are the moments of $\frac{1}{m}\mathbf{H}\mathbf{H}^H$.



Comparison of C_f and C_u for various number of observations. $\sigma^2 = 0.1$, $n = 10$ receive antennas, $m = 10$ transmit antennas. The rank of $\mathbf{H}$ was 3.



C_f for various number of observations. No phase off-set/phase drift. $\sigma^2 = 0.1$, $n = 10$ receive antennas, $m = 10$ transmit antennas. The rank of $\mathbf{H}$ was 3, 5 and 6.

How would an algorithm for free convolution look?

Definition

A family of unital $*$-subalgebras $(A_i)_{i\in I}$ is called a free family if

$$\left.\begin{array}{c} a_j \in A_{i_j} \\ i_1 \neq i_2, i_2 \neq i_3, \cdots, i_{n-1} \neq i_n \\ \phi(a_1) = \phi(a_2) = \cdots = \phi(a_n) = 0 \end{array}\right\} \Rightarrow \phi(a_1 \cdots a_n) = 0. \quad (7)$$

A family of random variables a_i is called a free family if the algebras they generate form a free family.

- How do we implement this in terms of moments?
- From the previous result, we are basically interested in computing the moments of ab, when ab are free, and b is free Poisson.

Implementation of main result

The following formula can be used for incremental calculation of the moments of the measure $\mu \boxtimes \mu_c$, from the moments of the measure μ:

$$[coef_m](cM_{\mu\boxtimes\mu_c}) = \sum_{k=1}^{m} [coef_k](cM_\mu)[coef_{m-k}](1 + cM_{\mu\boxtimes\mu_c})^k. \quad (8)$$

Here,

- $M_\mu(z) = \mu_1 z + \mu_2 z^2 + ...$, where μ_i are the moments of μ.
- $coef_k$ means the coefficient of z^k in the polynomial.
- The power series coefficient can be computed through k-fold discrete (classical) convolution.
- (8) is proved by first proving that $cM_{\mu\boxtimes\mu_c} = (cM_\mu)\boxed{\star}Zeta$.

Observation model 2

A more general observation model is:

$$\hat{\mathbf{H}}_i = \mathbf{D}_i^r \mathbf{H} \mathbf{D}_i^t + \sigma \mathbf{X}_i, \tag{9}$$

where $\mathbf{D}_i^r$ and $\mathbf{D}_i^t$ are $n \times n$ and $m \times m$ diagonal matrices which represent phase off-sets and phase drifts (impairments due to the antennas and not the channel) at the receiver and transmitter given respectively by

$$\begin{aligned} \mathbf{D}_i^r &= \mathrm{diag}[e^{j\phi_1^i}, ..., e^{j\phi_n^i}], \text{ and} \\ \mathbf{D}_i^t &= \mathrm{diag}[e^{j\theta_1^i}, ..., e^{j\theta_m^t}] \end{aligned}$$

where the phases ϕ_j^i and θ_j^i are random. We assume all phases independent and uniformly distributed.

Problem when extending to phase off-set and phase drift

- In the compund observation matrix we now put

$$\mathbf{H}_{1\ldots L} = \frac{1}{\sqrt{L}} \left[\mathbf{D}_i^r \mathbf{H} \mathbf{D}_i^t, \mathbf{D}_i^r \mathbf{H} \mathbf{D}_i^t, \ldots, \mathbf{D}_i^r \mathbf{H} \mathbf{D}_i^t \right],$$

  The moments of $\frac{1}{m}\mathbf{H}_{1\ldots L}\mathbf{H}_{1\ldots L}^H$ are now in general different from the moments of $\frac{1}{m}\mathbf{H}\mathbf{H}^H$!
- In other words stacking the observations and using the free convolution framework does not give us what we want

A way to resolve this:

- Don't stack the observations at all.
- Perform convolution through exact formulas for the mixed moments of matrices and Gaussian matrices of lower order.
- Unbiased capacity estimator.

Unbiased estimator for the moments of the channel matrix

Let $\hat{h}_i$ be the first moments of the sample covariance matrix $\frac{1}{m}\hat{\mathbf{H}}_i\hat{\mathbf{H}}_i^H$. An unbiased estimator for the first moments h_i of $\frac{1}{m}\mathbf{H}\mathbf{H}^H$ is given by

$$\begin{aligned}
\hat{h}_1 &= h_1 + \sigma^2 \\
\hat{h}_2 &= h_2 + 2\sigma^2(1+c)h_1 + \sigma^4(1+c) \\
\hat{h}_3 &= h_3 + 3\sigma^2(1+c)h_2 + 3\sigma^2 c h_1^2 \\
&\quad +3\sigma^4\left(c^2+3c+1+\tfrac{1}{m^2}\right)h_1 \\
&\quad +\sigma^6\left(c^2+3c+1+\tfrac{1}{m^2}\right) \\
\hat{h}_4 &= h_4 + 4\sigma^2(1+c)h_3 + 8\sigma^2 c h_2 h_1 \\
&\quad +\sigma^4(6c^2+16c+6+\tfrac{16}{m^2})h_2 \\
&\quad +14\sigma^4 c(1+c)h_1^2 \\
&\quad +4\sigma^6(c^3+6c^2+6c+1+\tfrac{5(c+1)}{m^2})h_1 \\
&\quad +\sigma^8\left(c^3+6c^2+6c+1+\frac{5(c+1)}{m^2}\right),
\end{aligned} \tag{10}$$

Exact formulas for expecations of mixed moments of Gaussian and deterministic matrices

We have that

$$\begin{aligned}
E\left[tr_n\left(\mathbf{W}_n\right)\right] &= m_1+\sigma^2\\
E\left[tr_n\left(\mathbf{W}_n^2\right)\right] &= m_2+2\sigma^2(1+c)m_1+\sigma^4(1+c)\\
E\left[tr_n\left(\mathbf{W}_n^3\right)\right] &= m_3+3\sigma^2(1+c)m_2+3\sigma^2 c m_1^2\\
&\quad +3\sigma^4\left(c^2+3c+1+\frac{1}{N^2}\right)m_1\\
&\quad +\sigma^6\left(c^2+3c+1+\frac{1}{N^2}\right)\\
E\left[tr_n\left(\mathbf{W}_n^4\right)\right] &= m_4+4\sigma^2(1+c)m_3+8\sigma^2 c m_2 m_1\\
&\quad +\sigma^4(6c^2+16c+6+\frac{16}{N^2})m_2\\
&\quad +14\sigma^4 c(1+c)m_1^2\\
&\quad +4\sigma^6(c^3+6c^2+6c+1+\frac{5(c+1)}{N^2})m_1\\
&\quad +\sigma^8\left(c^3+6c^2+6c+1+\frac{5(c+1)}{N^2}\right),
\end{aligned} \tag{11}$$

where $m_j = tr_n\left(\left(\frac{1}{N}\mathbf{R}_n\mathbf{R}_n^H\right)^j\right)$.

Derivation of the limiting distribution for $\frac{1}{N}\mathbf{X}\mathbf{X}^H$

When x is standard complex Gaussian, we have that

$$E\left(|x|^{2p}\right) = p!.$$

A more general statement concerns a random matrix $\frac{1}{N}\mathbf{X}\mathbf{X}^H$, where $\mathbf{X}$ is an $n \times N$ random matrix with independent standard complex Gaussian entries. It is known [HT] that

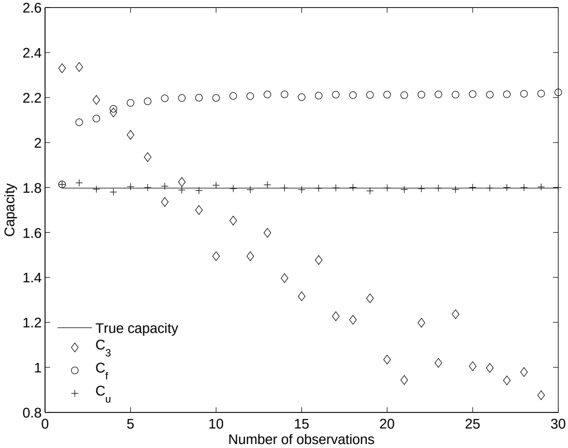
$$\tau_n\left(\left(\frac{1}{N}\mathbf{X}\mathbf{X}^H\right)^p\right) = \frac{1}{N^p n}\sum_{\pi \in S_p} N^{k(\hat{\pi})} n^{l(\hat{\pi})},$$

where $\hat{\pi}$ is a permutation in S_{2p} constructed in a certain way from π, and $k(\hat{\pi}), l(\hat{\pi})$ are functions taking values in $\{0, 1, 2, ...\}$.

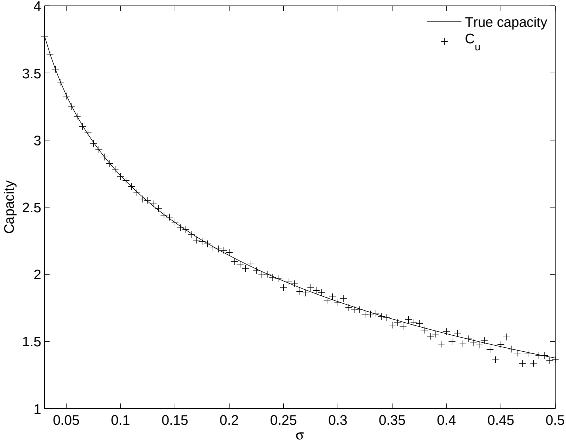
One can show that this equals

$$\tau_n\left(\left(\frac{1}{N}\mathbf{X}\mathbf{X}^H\right)^p\right) = \sum_{\hat{\pi}\in NC_{2p}} c^{l(\hat{\pi})-1} + \sum_k \frac{a_k}{N^{2k}}.$$

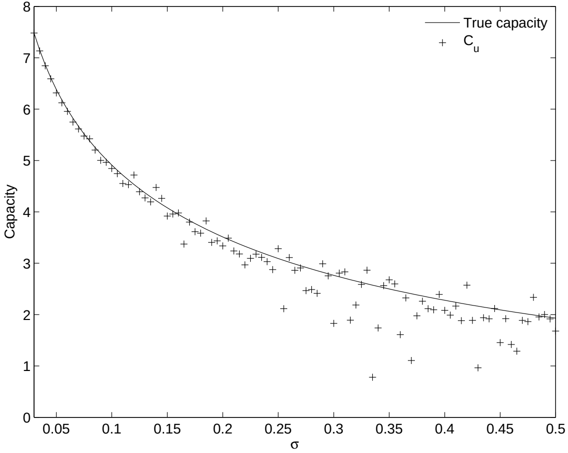
The convergence is "almost sure".



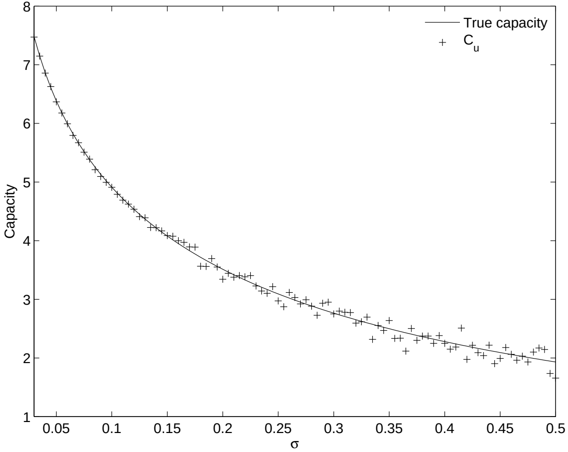
Comparison of capacity estimators which worked when no phase off-set/drift was present, for increasing number of observations. With phase drift and phase off-set. $\sigma^2 = 0.1$, $n = 10$ receive antennas, $m = 10$ transmit antennas. The rank of $\mathbf{H}$ was 3.



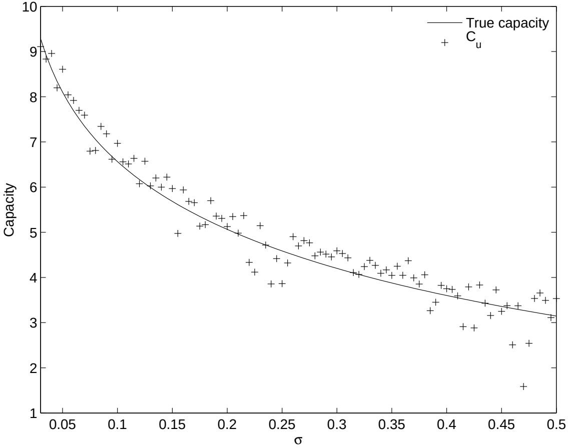
C_u for $L = 1$ observation, $n = 10$ receive antennas, $m = 10$ transmit antennas, with varying values of σ. With phase drift and phase off-set. The rank of $\mathbf{H}$ was 3.



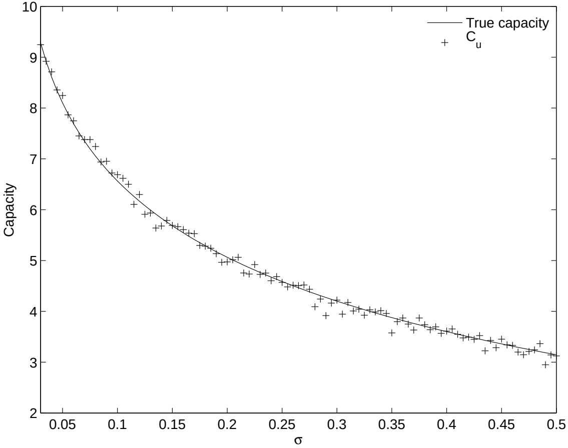
C_u for $L = 1$ observation, $n = 4$ receive antennas, $m = 4$ transmit antennas, with varying values of σ. With phase drift and phase off-set. The rank of $\mathbf{H}$ was 3.



C_u for $L = 10$ observations, $n = 4$ receive antennas, $m = 4$ transmit antennas, with varying values of σ. With phase drift and phase off-set. The rank of $\mathbf{H}$ was 3.



C_u for $L = 50$ observations, $n = 4$ receive antennas, $m = 4$ transmit antennas, with varying values of σ. With phase drift and phase off-set. The rank of $\mathbf{H}$ was 4.



C_u for $L = 1600$ observations, $n = 4$ receive antennas, $m = 4$ transmit antennas, with varying values of σ. With phase drift and phase off-set. The rank of $\mathbf{H}$ was 4.

References

[HT]: "Random Matrices and K-theory for Exact C^*-algebras". U. Haagerup and S. Thorbjørnsen. citeseer.ist.psu.edu/114210.html. 1998.

This talk is available at
http://heim.ifi.uio.no/∼oyvindry/talks.shtml.
My publications are listed at
http://heim.ifi.uio.no/∼oyvindry/publications.shtml