Random Matrices in Communications: Diagrams, Saddles and Replicas

Aris Moustakas
Physics Dept., University of Athens

Outline

- Diagrammatic approach to random matrix theory and applications to communications
- Saddle – point approach to random matrix theory
- Replicas and saddle - points

Outline

- Diagrammatic approach to random matrix theory and applications to communications
- Saddle – point approach to random matrix theory
- Replicas and saddle - points

Introduction

- Why do we need Random Matrix Theory in Communications?
  - To calculate statistics of random matrix-valued quantities
    - E.g. what is the PDF of $I\left(H\right)=\log\det\left(I+HH^{\dagger}\right)$
    - or what is the average of $SINR_k\left(S\right)=s_k^{\dagger}\left[I+SS^{\dagger}-s_k s_k^{\dagger}\right]^{-1}s_k$
  - Thus we need to calculate averages, variances, etc of functions of matrices
    - E.g. $\overline{f}_N=\frac{1}{N}E\left[Tr\left\{f\left(A_N\right)\right\}\right]$
  - For a given matrix (and its statistics), it is best to try to calculate the eigenvalue distributions (and then get statistics for all functions f)

$$\rho_N(z)=\frac{1}{N}E\left[\sum_{n=1}^{N}\delta\left(z-\lambda_n^N\right)\right]$$

$$\rho_{cN}(z,z')=\frac{1}{N^2}E\left[\sum_{n,m=1}^{N}\delta\left(z-\lambda_n^N\right)\delta\left(z'-\lambda_m^N\right)\right]-\rho_N(z)\rho_N(z')$$

    - etc
  - Then $\overline{f}_N=\int dz\ \ \rho_N(z)f(z)\quad Var_N(f)=\int dz\int dz'\ \ \rho_{cN}(z,z')f(z)f(z')$

Introduction

  - For finite N, $\rho(z)$ etc. have large fluctuations, therefore not always useful, and also more often than not impossible to calculate
  - Try limit $N = \infty$
    - Then densities do not depend on N and "harden".

- Diagrammatic Approach:
  - Simple intuitive method to calculate density of eigenvalues and other quantities without too much effort.
    - As examples will calculate densities of eigenvalues, but can be extended to other quantities
    - Variance of eigenvalues: Universal

- Saddle-point method to calculate general density of eigenvalues
  - Is valid for general probability distributions of matrices.

Diagrammatic Approach

- Want to calculate $\rho_N(z)$
  - Use identity $\rho_N(z) = \frac{1}{N} E\left[\sum_n \delta(z-\lambda_n)\right] = \frac{1}{\pi N} Im\left(E\left[Tr\frac{1}{z-i0^+-A_N}\right]\right)$
  - Thus need to calculate the Green´s function G(z) for complex z

$$g(z) = \frac{1}{N} E\left[Tr\frac{1}{z-A}\right] = \frac{1}{N}\sum_{n=0}^{\infty}\frac{Tr\left(E\left[A^n\right]\right)}{z^{n+1}}$$

    - Then take appropriate limits for z
    - Underlying Principle: Since A hermitian, g(z) analytic (hence expansion exists) except possibly on real axis, where it will have a branch cut. This principle breaks down for non-hermitian A, (not topic of this course)
  - How does the expectation E(.) apply on matrices?
    - Contraction of indices of matrices
    - Example: Gaussian statistics:

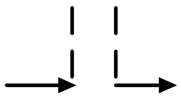
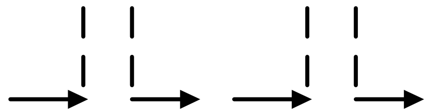
$$E\left[A_{ij}A_{kl}\right] = c\int \exp\left[-\frac{N}{2\sigma^2}Tr\left[\mathbf{A}^2\right]\right] = \frac{\sigma^2}{N}\delta_{il}\delta_{jk}$$

    - Similarly:

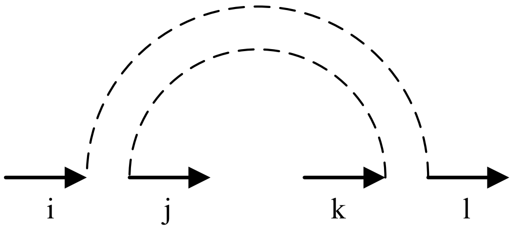
$$E\left[A_{ij}A_{kl}A_{mn}A_{pq}\right] = \frac{\sigma^4}{N^2}\left(\delta_{il}\delta_{jk}\delta_{mq}\delta_{np} + \delta_{in}\delta_{jm}\delta_{kq}\delta_{lp} + \delta_{iq}\delta_{jp}\delta_{kn}\delta_{lm}\right)$$

    - Wick: connect all possible index pairs (index democracy)
  - How can one see this diagrammatically?

Diagrammatic Approach

- Represent matrix element A_{ij} as one wiggly line (or two dashed ones) with two external lines
  - A_{ij} 
  - $A_{ij}A_{kl}$ 
  - Average over ***A*:**

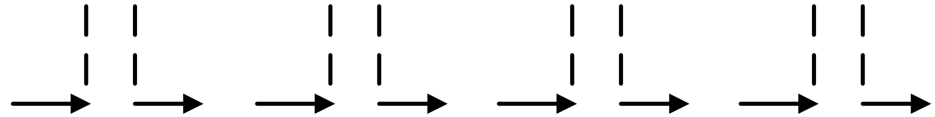
$$E\left[A_{ij}A_{kl}\right] = \frac{\sigma^2}{N}\delta_{il}\delta_{jk}$$



Diagrammatic Approach

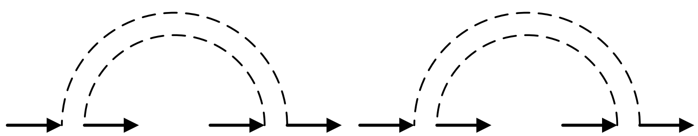
– Similarly 4 matrix average

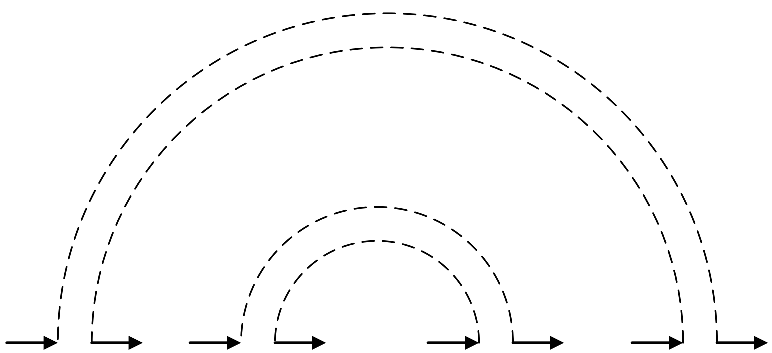
$A_{ij}A_{kl}A_{mn}A_{pq}$

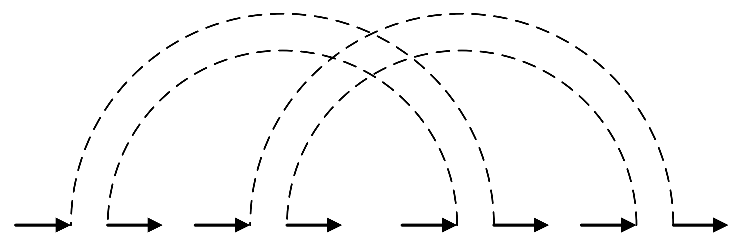


– becomes after averaging

$$E\left[A_{ij}A_{kl}A_{mn}A_{pq}\right] = \frac{\sigma^4}{N^2}\left(\delta_{il}\delta_{jk}\delta_{mq}\delta_{np} + \delta_{iq}\delta_{jp}\delta_{kn}\delta_{lm} + \delta_{in}\delta_{jm}\delta_{kq}\delta_{lp}\right)$$



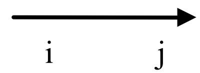




Diagrammatic Approach

– So far expressed A diagrammatically. Need to express 1/z

$$G_0(z)_{ij} = \frac{\delta_{ij}}{z}$$



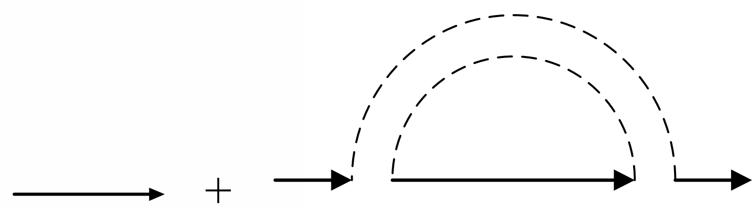
also $g_0(z) = \frac{1}{N} Tr G_0(z) = \frac{1}{z}$

– Then G becomes

$$G(z) \quad = \quad E\left[\frac{\mathbf{I}}{z\mathbf{I} - A}\right] = E\left[G_0(z) + G_0(z)AG_0(z) + G_0(z)AG_0(z)AG_0(z) + \ldots\right]$$

– To quadratic order in A we have

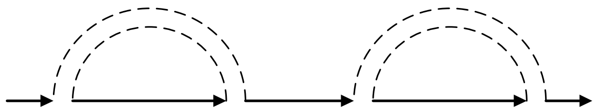
$$\begin{aligned} G_{ik}(z) \quad &= \quad G_{0,ik} + \sum_{mjnp}^{N} G_{0,im} E\left[A_{mj} G_{0,jn} A_{np}\right] G_{0,pk} \\ &= \quad \frac{\delta_{ik}}{z} + G_{0,ik}^2 \frac{\sigma^2}{N} Tr G_0 \\ &= \quad G_{0,ik} + G_{0,ik}^2 \sigma^2 g_0(z) \end{aligned}$$



- Double dashed line gives σ^2/N
- Loop (Trace) gives N
- Rainbow diagram

Diagrammatic Approach

- So far so good
  - The 4th order term is: $E\left[G_0(z)AG_0(z)AG_0(z)AG_0(z)AG_0(z)\right]$
  - First diagram (way of connecting dashed lines)



$$\frac{\sigma^4}{N^2}\delta_{il}\delta_{jk}\delta_{mq}\delta_{np}$$

- 2 matrix pairs: $1/N^2$
- 2 loops: N^2

$$E\left[G_0(z)AG_0(z)AG_0(z)AG_0(z)AG_0(z)\right]^1_{ii}$$

$$= \frac{\sigma^4}{N^2}\sum_{ijklmnpq}\frac{\delta_{ii}}{z}\delta_{jk}\frac{\delta_{jk}}{z}\delta_{il}\frac{\delta_{lm}}{z}\delta_{np}\frac{\delta_{np}}{z}\delta_{mq}\frac{\delta_{qi}}{z}$$

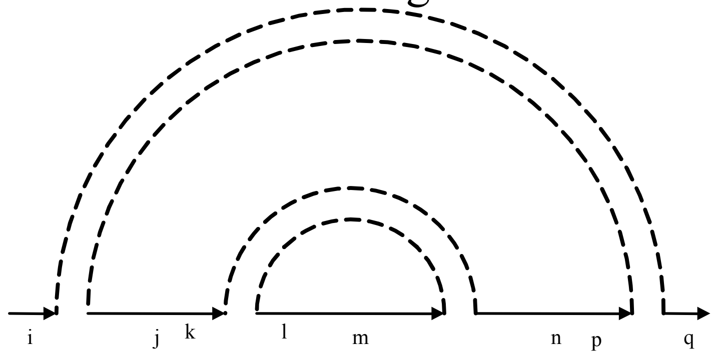
$$= G_0(z)^3_{iq}\frac{\sigma^4}{N^2}(TrG_0(z))^2 = G_0(z)^3_{iq}\sigma^4 g_0(z)^2$$

  - Can be disconnected into 2 pieces by breaking a single $G_0(z)$ line

$$G_0(z)E\left[AG_0(z)A\right]G_0(z)E\left[AG_0(z)A\right]G_0(z)$$

Diagrammatic Approach

– 4^{th} order - Second diagram



$$\frac{\sigma^4}{N^2}\delta_{iq}\delta_{jp}\delta_{kn}\delta_{lm}$$

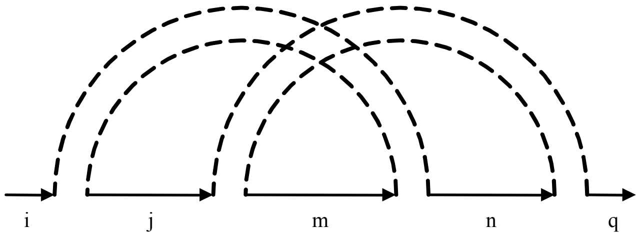
$$E\left[G_0(z)AG_0(z)AG_0(z)AG_0(z)AG_0(z)\right]_2$$

$$= \cdots = G_0(z)^2\frac{\sigma^4}{N^2}Tr\left[G_0(z)^2\right]Tr\left[G_0(z)\right] = G_0(z)^2_{iq}\sigma^4 g_0(z)^3$$

- 2 loops N^2
- 2 matrix pairs N^{-2}
- CANNOT be separated into 2 pieces by breaking a single $G_0(z)$ line

Diagrammatic Approach

– 4^{th} order - Third diagram



$$\frac{\sigma^4}{N^2}\delta_{in}\delta_{jm}\delta_{kq}\delta_{lp}$$

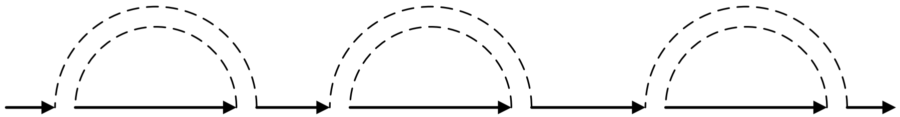
$$E\left[G_0(z)AG_0(z)AG_0(z)AG_0(z)AG_0(z)\right]_2$$
$$= \cdots = \frac{\sigma^4}{N^2}\left[G_0(z)^5\right]_{iq} = \frac{\sigma^4}{N^2}g_0(z)^5\delta_{iq}$$

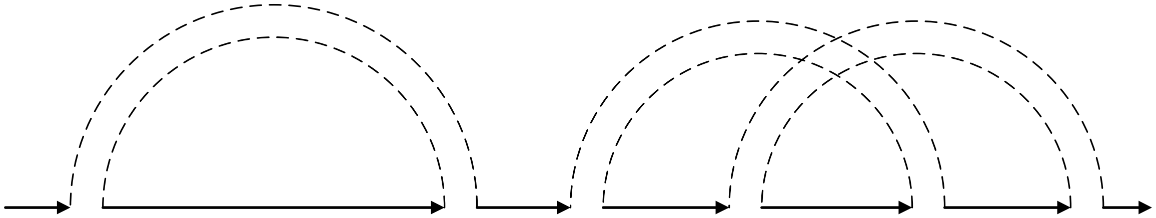
- No loops thus no summation (trace). All indices equal to the external ones (i,q)
- 2 matrix pairs N^{-2}
- SUBLEADING term for large N
- Also cannot be separated into 2 pieces by breaking a single $G_0(z)$ line
- Cannot uncross crossed lines by shifting lines around (crossing is a fundamental aspect of this diagram)
- An aspect of non-commutativity of matrices

Diagrammatic Approach

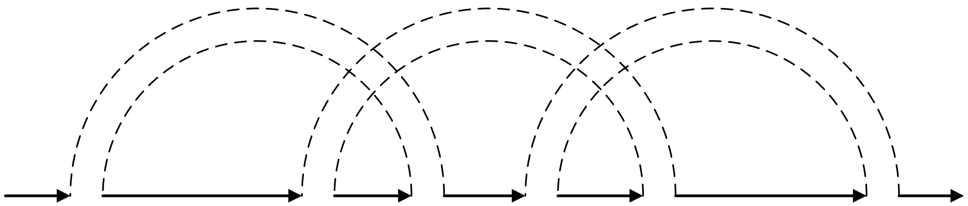
– Higher orders: What we expect

- Due to index democracy, all previous diagrams (ways of contracting dashed lines) will appear in separable form (cf diagram 1)





- plus more …

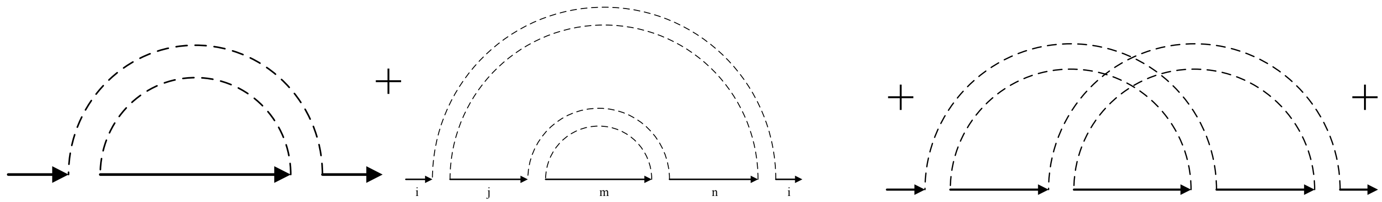


Diagrammatic Approach

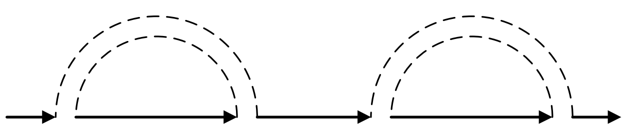
- Rules
  - Rearrange expansion in A as follows:

$$G(z) = G_0(z) +$$

$$G_0(z)\,\{ \quad + \quad + \quad + \ldots \}\, G_0(z)$$



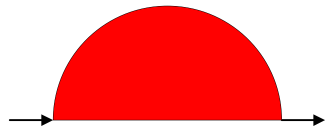
$$+G_0(z)\,\{\cdots\}\,G_0(z)\,\{\cdots\}\,G_0(z) + \cdots$$

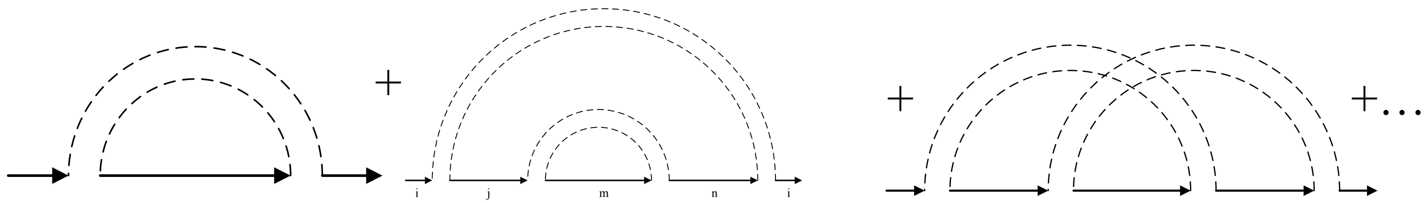


Diagrammatic Approach

- Rules (N still not large – finite)
  - Define following matrix Σ:

$\Sigma =$





$$= \left(\sigma^2 g_0(z) + \quad \sigma^4 g_0(z)^3 + \quad \frac{\sigma^4 g_0(z)^3}{N^2} + \cdots \right) \delta_{ik}$$

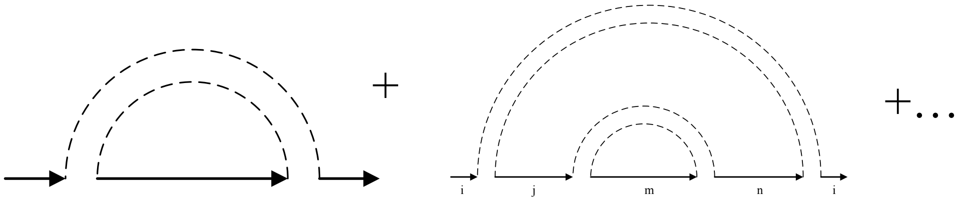
- We do not include the two "end" Green's functions
- Includes all diagrams that cannot be written as products of diagrams by breaking a single (solid) line ("one-particle irreducible" 1PI-diagrams)
- Σ is called "self-energy"
- Relation with R-transform (=Σ(g)) and S transform (=1/Σ) in free probability

Diagrammatic Approach

- Rules
  - Then G can be resummed simply in terms of Σ:

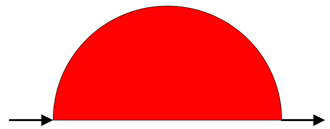
$$G = G_0 + G_0^2\Sigma + G_0^3\Sigma^2 + \cdots = \frac{1}{G_0^{-1} - \Sigma}$$

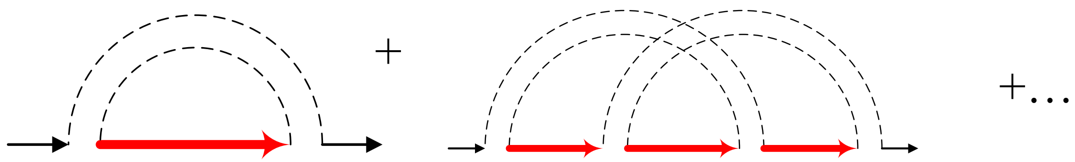
- So far the solid lines in Σ correspond to $G_0(z)$
- But, whenever we see a G_0 we will also have a $G_0^2\Sigma$+...
  - E.g.



$$\sigma^2 g_0(z) + \quad \sigma^2 g_0(z)^2\Sigma + \cdots = \sigma^2 g(z) = \sigma^2 \tfrac{1}{N} E\left[Tr\,\tfrac{I}{zI - A}\right]$$

- Rules
  - Thus the self-energy (diagonal) matrix Σ can be written in terms of G(z) as follows:

$\Sigma =$ 



$$= \sigma^2 g(z) + \frac{\sigma^4 g(z)^3}{N^2} + \cdots$$

  - All other diagrams have crossed lines
  - Our job has been simplified but, for finite N, we still need to find all diagrams which contribute to Σ

Diagrammatic Approach

- Asymptotics for N>>1
  - For large N, all diagrams with crossed lines can be neglected.
    - All diagrams with crossed lines contribute to terms in higher order of the small parameter 1/N.
    - The remaining diagrams are called planar, because the y can be drawn on a plane without any crossing lines
      - Diagrams with crossings can only be drawn without crossings in surfaces with "holes", the 1/N is a topological expansion (t'Hooft 60's – Nobel prize)

  - For large N (and Gaussian A), we are done!

$$\Sigma = \sigma^2 g(z)$$ and

$$G(z) = \frac{\mathbf{I}}{z-\Sigma} = \frac{\mathbf{I}}{z-\sigma^2 g(z)}$$ or $$g(z) = \frac{1}{z-\sigma^2 g(z)}$$

  - Solving for g gives $$g(z) = \frac{z-\sqrt{z^2-4\sigma^2}}{2\sigma^2}$$

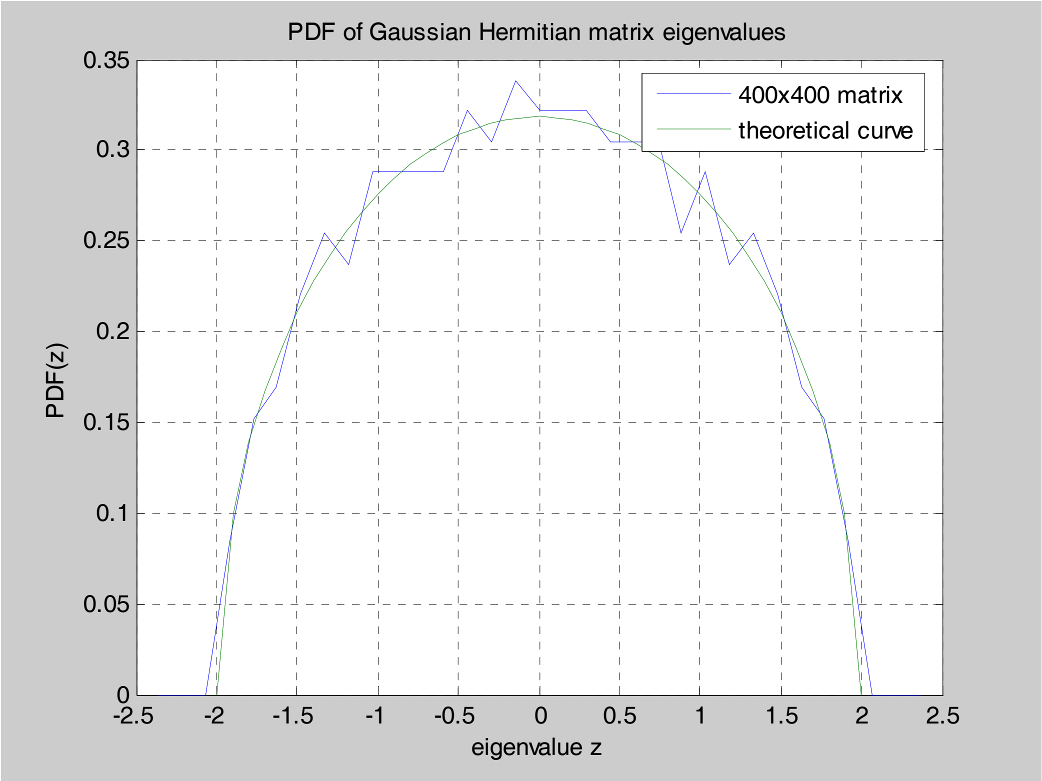
    - We keep the solution of the quadratic equation that behaves as 1/z for large z

Diagrammatic Approach

- Asymptotics for N>>1
  - We can now calculate the density of eigenvalues of A:

$$\rho(z) = \frac{\sqrt{4\sigma^2 - z^2}}{2\sigma^2 \pi}$$

  - Proof of Semicircle Law!!! (for Gaussian case)



Diagrammatic Approach

- How do we generalize to other systems of interest to communications?
  - E.g. take the case of the matrix $A = HRH^{\dagger}$
    - R is positive semi-definite matrix (correlation)
    - H is i.i.d. complex Gaussian NxN matrix
      - Applies to case where H is MxN (M<N) by setting according number of rows and columns of R to zero
  - Now we need to calculate the Green function

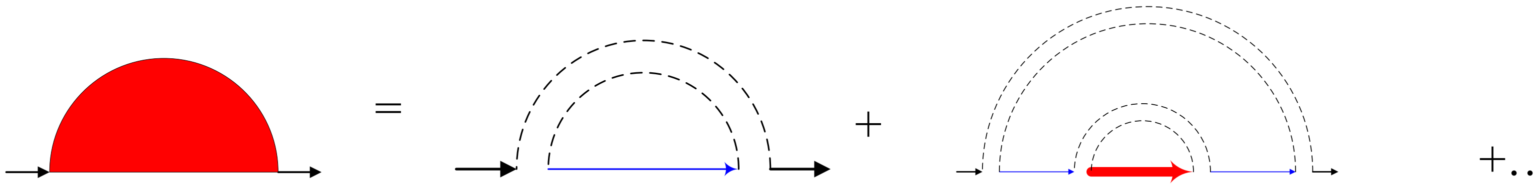
$$\begin{aligned} g(z) &= \frac{1}{N} ETr\left[\frac{1}{z - HRH^{\dagger}}\right] \\ &= \sum_{n=0}^{\infty} \frac{E\left[Tr\left(HRH^{\dagger}\right)^{n}\right]}{z^{n+1}} \end{aligned}$$

  - As before we can show that G(z) can be expressed as

$$g(z) = \frac{1}{z - \Sigma(z)}$$

Diagrammatic Approach

- $\Sigma(z)$ calculation



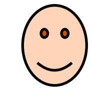
$$\begin{aligned}\Sigma(z) &= \frac{\sigma^2}{N} Tr\,[R] + \frac{\sigma^4}{N} Tr\left[R^2\right] g(z) + \cdots \\ &= \frac{\sigma^2}{N} \sum_{n=0}^{\infty} Tr\left[R\left(\sigma^2 R g(z)\right)^n\right]\end{aligned}$$

- Thus

$$\Sigma(z) = \frac{\sigma^2}{N} Tr\left[\frac{R}{1-\sigma^2 R g(z)}\right] \qquad g(z) = \frac{1}{z-\Sigma(z)}$$

- We see that Σ is not simply G, but another "Green" function
- For R=I we get $g(z) = \dfrac{z-\sqrt{z^2-4z\sigma^2}}{2z\sigma^2}$

Diagrammatic Approach

- We can still calculate the density of eigenvalues of $HRH^\dagger$ or $(H_0 + \sigma H)(H_0 + \sigma H)^\dagger$ 
  - More involved but straightforward
- Applications:
  - Imagine measuring N times a M-vector which has the (unknown) correlation matrix R. then the best estimate of R is $\hat{R} = HRH^\dagger$
    - One would hope that if M/N=β>≈1 one can recover the spectrum of R
    - For M,N>>1, but with a fixed ratio this can be done analytically
  - To calculate the ergodic mutual information of a Gaussian correlated channel, one needs to calculate

$$I = E\left[\log\det\left(\mathbf{I} + \rho HRH^\dagger\right)\right]$$

    - Which can be done if one knows the eigenvalue distribution of the matrix $HRH^\dagger$

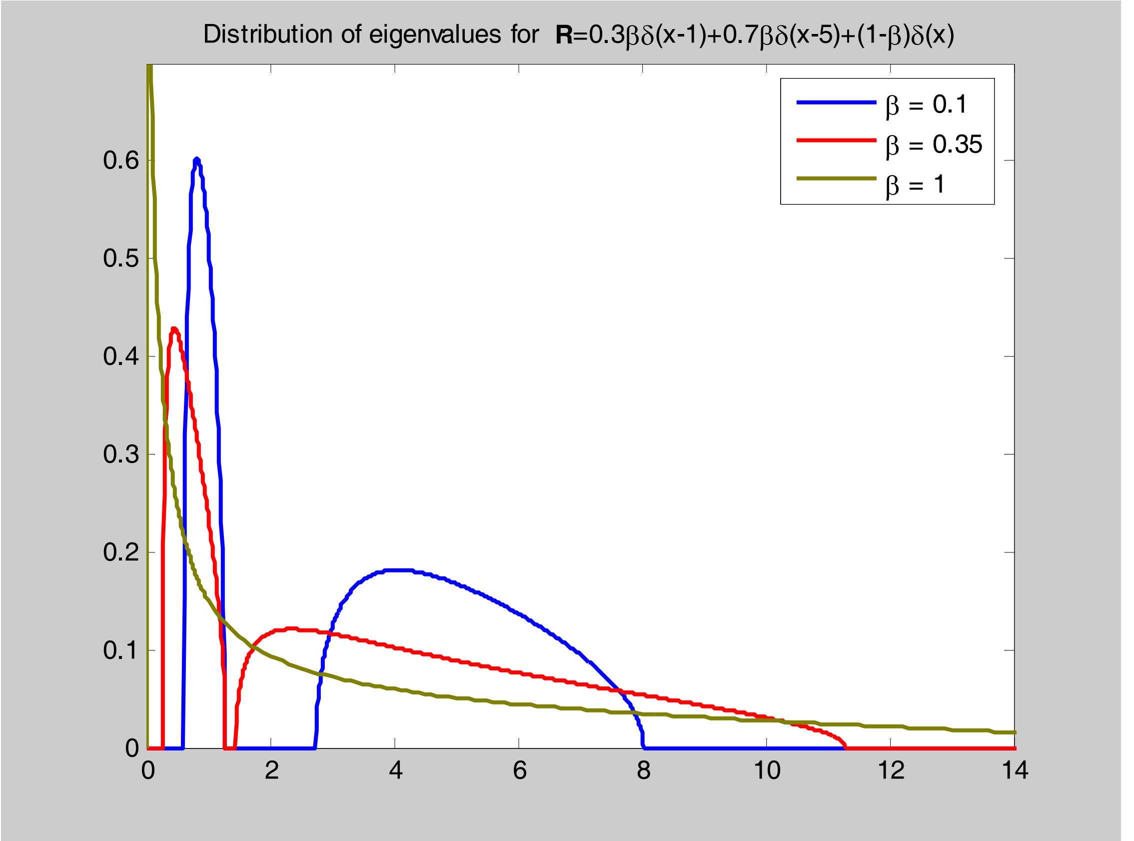
$$I = N\int dx\rho(x)\log\left(1 + \rho x\right)$$

Diagrammatic Approach

- Example: Assume the following eigenvalue distribution of R

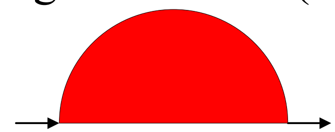
$$R(x) = 0.3\beta\delta(x-1) + 0.7\beta\delta(x-5) + (1-\beta)\delta(x)$$

  - Here need to solve cubic equation
  - Fraction of non-zero eigenvalues $\beta = \frac{M}{N}$
  - Even for N/M=10 (β=0.1) peak at x=5 not pronounced
  - For β=0.35 two blobs just broken apart
  - For β=1 absolutely no knowledge of the spectrum
  - Applications to econometric and communications data analysis

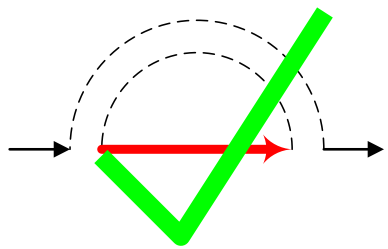


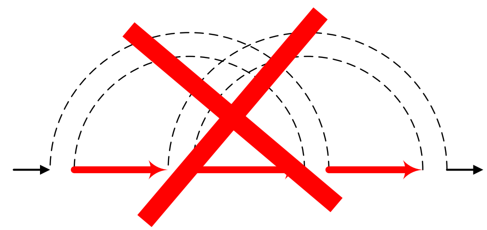
Diagrammatic Approach

- Recap:
  - Diagrams which cannot be separated by cutting a single solid line (1PI-diagrams) help us define and use self-energy Σ(z)



  - Large N approximation allows us to use only planar diagrams



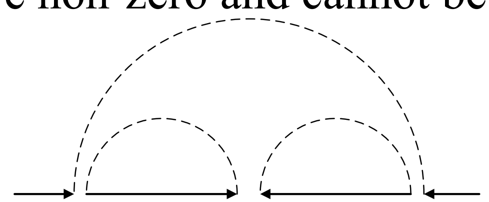


  - Gaussian Matrix approximation (implicit so-far):
    - Dashed lines always travel together
    - Results from having only non-zero second moments of matrices only, i.e.

$$E\left[A_{ij}A_{kl}\right] = \sigma_{ij}^2\delta_{il}\delta_{jk}/N$$

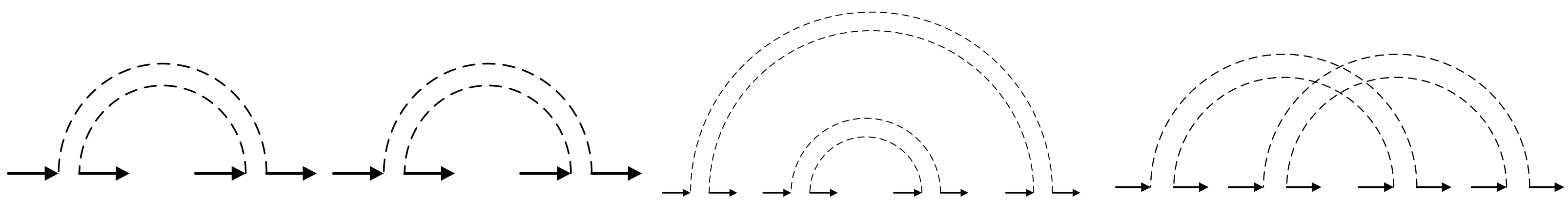
  - What happens when higher moments are non-zero and cannot be broken into products of second moments?, e.g.

$$E\left[A_{ij}A_{kl}A_{mn}\right] = f_{ijklmn}$$



Diagrammatic Approach

- Non-Gaussian RMT:
  - First recall Gaussian case:
    - To take the average of nth power of A, say $E\left[A_{ij}A_{kl}A_{mn}A_{pq}\right]$
    - When averaging over Gaussian matrices we connected (or contracted,or equalized) pairs of indices in a such a way that matrices are always paired together. This means that dashed lines travel together.

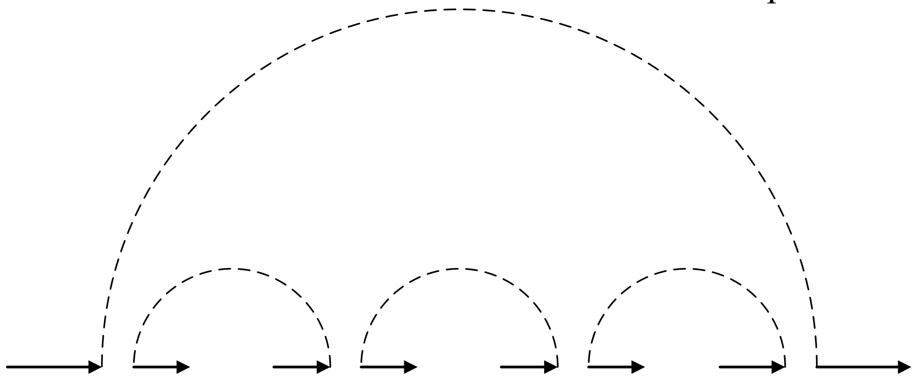


$$E\left[A_{ij}A_{kl}A_{mn}A_{pq}\right] = E\left[A_{ij}A_{kl}\right]E\left[A_{mn}A_{pq}\right] \text{ + two other pairings}$$

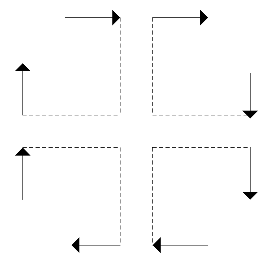
- Each of the expectations are usually just 1/N times something, e.g. σ_{ij}^2

Diagrammatic Approach

- Non-Gaussian RMT:
  - In contrast, averaging over a general (non-Gaussian) probability distribution of A will contract (equalize) all possible pairs of indices.
    - In the case of $E\left[A^4\right]$ we have now two types of terms:
      - Those, which can be written as products of expectations of pairs of A's as in the Gaussian case
      - Those that cannot be written as such products:



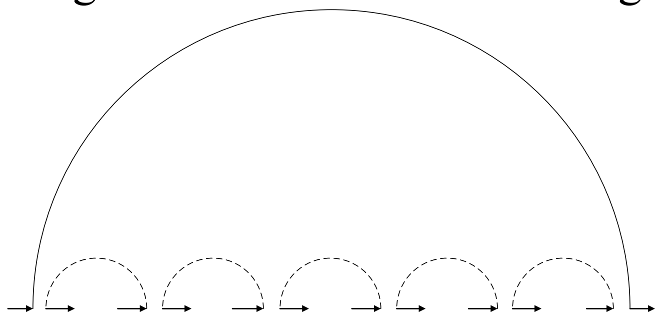
or



- These terms may have complicated behavior in N, but they will contribute in the large N limit, only if they scale as $1/N^3$. We can see this by connecting the solid lines: We create 4 loops, thus this diagram scales like $N^4 * 1/N^3 = N$, as we would get by connecting the lines of a quadratic term.

Diagrammatic Approach

- Non-Gaussian RMT:
  - For higher orders we have similar graphs appearing e.g.



- The leading order behavior of a term with 2k dashed lines is $\frac{\gamma_{2k}}{N^{2k-1}}$
- γ_{2k} is essentially the 2k-order correlation function of this probability distribution. (Usually we prefer it to be independent of indices)
- Again, only planar graphs will contribute to leading order in N (graphs with crossed lines will be de-facto subleading in N)
- Note that we have not discussed how to compute these coefficients.

Diagrammatic Approach

- Non-Gaussian RMT: Example: Show that $\gamma_{2k} = 0$ (k>1) for random matrix with i.i.d. elements $A_{ij} = \frac{\pm 1}{\sqrt{N}}$
  - For simplicity will only calculate γ_4
  - By definition

    $$\gamma_4 = \lim_{N\to\infty} \frac{1}{N} Tr\left\{E\left[AA^T AA^T\right]\right\} - \frac{2}{N^2} Tr^2\left\{E\left[AA^T\right]\right\}$$

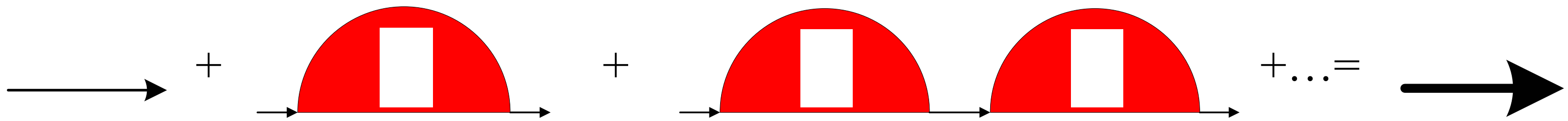
    - We subtract off the (2) disconnected diagrams
    - However,

    $$\gamma_2 = \frac{1}{N} Tr\left\{E\left[AA^T\right]\right\} = \frac{1}{N}\sum_{ij} E\left[A_{ij}^2\right] = 1$$

    - Also $\sum_{ijkp} E\left[A_{ij}A_{kj}A_{kp}A_{ip}\right] = \frac{2N^3 - N^2}{N^2}$
    - As a result $\gamma_4 = 0$!!! It can also be shown that this holds for higher k
  - This means that this matrix (a CDMA-code matrix) is equivalent to a Gaussian for large N
    - Note that this is generally not valid for finite N

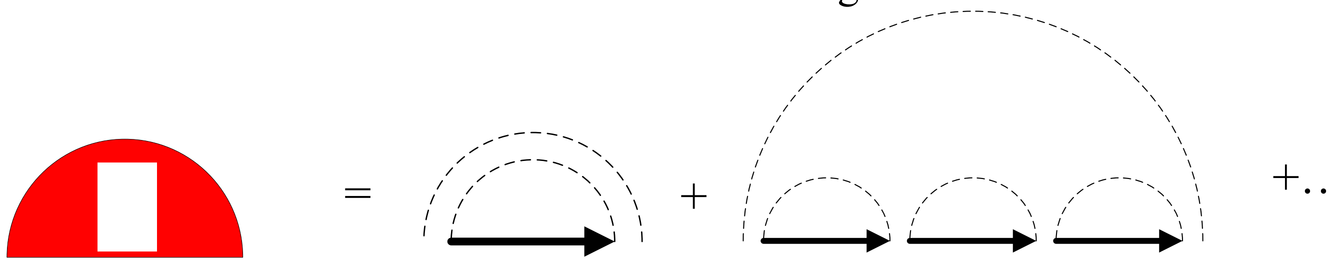
Diagrammatic Approach

- Non-Gaussian RMT:
  - Given the γ_{2k} we can now calculate the Green's function explicitly following the rules discussed earlier. First we have as before



$$G = G_0 + G_0 \Sigma G_0 + G_0 \Sigma G_0 \Sigma G_0 + = \frac{1}{G_0^{-1} - \Sigma}$$

  - Σ can now be calculated to leading order in N as follows



$$\Sigma = \gamma_2 g + \gamma_4 g^3 + = \sum_{k=1}^{\infty} \gamma_{2k} g^{2k-1}$$

  - Note that the 2k-th term has 2k-1 bubbles, resulting to $N^{1-2k} N^{2k-1} = O(1)$
- – Combining these two equations, we have a closed-form solution for g!

Diagrammatic Approach

- How do we calculate γ_{2k}??
  - Generally very hard but independent of problem one needs to calculate
    - Depend **only** on probability distribution
  - Usually we calculate them perturbatively assuming the distribution of A is close to a Gaussian, e.g.

$$P(A) \propto e^{-N\left\{\frac{1}{2}Tr\left[A^2\right]+\frac{g}{4}Tr\left[A^4\right]\right\}}$$

    - with $g<<1$ and expand the exponential in g
  - A few examples exist where they can be calculated explicitly:
    - e.g. the distribution of NxN unitary matrices **S** with $\mathbf{S}\mathbf{S}^\dagger = \mathbf{S}^\dagger\mathbf{S} = \mathbf{I_N}$
  - Implicit calculations
- Will show how to get these γ_{2k} implicitly for 2 case
  - Unitary case
  - Exponential polynomial distribution $P(A) \propto e^{-NTr[V_p(A)]}$
    - Even $V_p(x)$ polynomial of arbitrary degree p

Diagrammatic Approach

- Unitary case
  - Take $A = \mathbf{S} + \mathbf{S}^\dagger$
  - Since S is a unitary matrix its eigenvalues are $e^{i\theta_n}$ with θ_n i.i.d. Thus

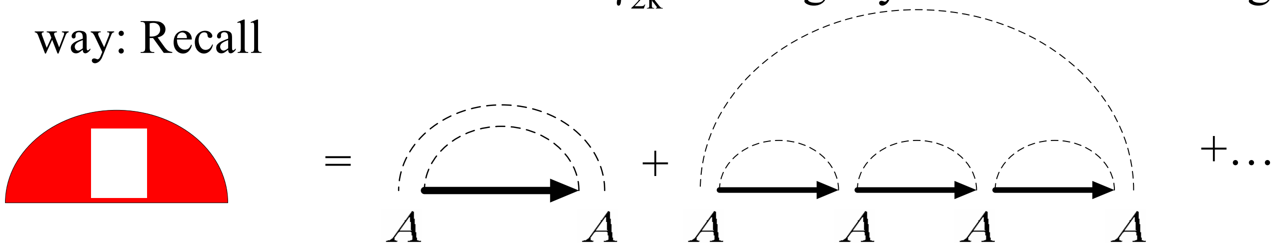
$$
\begin{aligned}
g(z) &= \frac{1}{N} E\left[Tr \frac{1}{z-A}\right] \\
&= \int_0^{2\pi} \frac{d\theta}{2\pi} \frac{1}{z - 2\cos\theta} = \frac{1}{\sqrt{z^2-4}} \qquad |z| > 2 \\
&= \frac{1}{z - \Sigma(z)}
\end{aligned}
$$

  - Eliminating z, we get

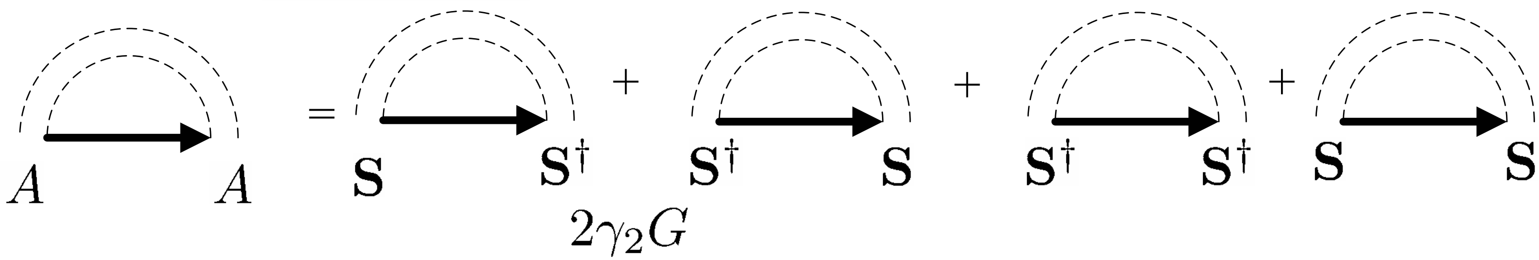
$$
\Sigma = \frac{\sqrt{1+4g^2}-1}{g} = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(2k)!}{(2k-1)k!^2} g^{2k-1}
$$

Diagrammatic Approach

- Unitary case
  - In this case Σ is related to the γ_{2k} in a slightly different but straightforward way: Recall



  - But since $A = \mathbf{S} + \mathbf{S}^{\dagger}$ we have for the first bubble



  - Last 2 terms cancel by symmetry. Similar cancellations at higher order, thus

$$\Sigma = 2 \sum_{k=1}^{\infty} \gamma_{2k} G^{2k-1}$$

$$\gamma_{2k} = \frac{c_k}{2} = \frac{(2k)!}{2(2k-1)k!^2}$$

Diagrammatic Approach

- Unitary case
  - May now use this result to calculate all possible unitary matrix averages.
  - For example the eigenvalue spectrum of $P_1 \mathbf{S} P_2 \mathbf{S}^\dagger P_1$
    - Here one needs to calculate the function

$$G(z) = E\left[\frac{1}{z - P_1 \mathbf{S} P_2 \mathbf{S}^\dagger P_1}\right]$$

    - This can be done using methods described above in the calculation of $HRH^\dagger$
    - For P_1 P_2 simple projection operators one looks for the spectrum of t_{21}

$$S = \begin{bmatrix} r_{11} & t_{31} & t_{21} \\ t_{13} & r_{33} & t_{32} \\ t_{12} & t_{23} & r_{22} \end{bmatrix}$$

-

$$\rho(x) = \frac{\sqrt{(x_{max} - x)(x - x_{min})}}{2\pi x(1-x)}$$

$$x_{max,min} = \frac{\left(\sqrt{(N_1(N-N_2)} \pm \sqrt{N_2(N-N_1)}\right)^2}{N^2}$$

Diagrammatic Approach

Recap

- This diagrammatic approach captures many (most) cases exactly
  - Similarities with other methods (free probability)
  - Works also for non-Gaussian cases (not all)
  - Can be used directly to calculate MMSE-type averages

- Other uses:
  - Can be used in very straightforward fashion (not here!!) to calculate second order moments (i.e. two eigenvalue distributions)
  - Often there is no more than 2^{nd} order (Gaussianity)
    - Thus can calculate full asymptotic distributions of mutual informations, SINRs etc.
  - Second order moments can be of more importance because the depend on less details: Universality

Outline

- Diagrammatic approach to random matrix theory and applications to communications
- Saddle – point approach to random matrix theory
- Replicas and saddle - points

Saddle-point Approach

- Exponential polynomial distribution
  - Even $V_p(x)$ polynomial of arbitrary degree p (even) $P(A) \propto e^{-NTr[V_p(A)]}$
    - E.g. Gaussian $V_2(x) = x^2/(2\sigma^2)$
  - Here will use different approach, based on an asymptotic approximation of integrals, which tend to be very peaked when one parameter is large (here N)

  - Recall saddle point approximation of integrals:
    - Take an integral of the form

      $$I_N[g] = \int dx e^{-Nf(x)} g(x)$$

    - For large N (assuming everything is "nice") we have (for $f'(x_0)=0$)

      $$I_N[g] \approx g(x_0) e^{-Nf(x_0)} \sqrt{\frac{2\pi}{Nf''(x_0)}}$$

    - Thus all the "action" of the whole integral is centered at the saddle point
    - If the exponential is a probability distribution, then the mean of any function g(x) is determined by its saddle point value

      $$E[g(x)] \approx g(x_0)$$

Saddle-point Approach

- Our task is to calculate G

$$G(z) = E\left[Tr\left\{\frac{1}{z-A}\right\}\right] = \sum_{n=1}^{N} E\left[\frac{1}{z-x_n}\right]$$

  - From this one can get ρ(z)
  - Behavior of G will be determined by saddle point properties of p(A)
    - Saddle pt behavior is characteristic of hardening/deterministic behavior
  - Since G=f(x_i) and TrV(A) = $\sum_i$ V(x_i) it is best to change coordinates from A_{ij} to eigenvalues x:
    - Use $A = \mathbf{U}\Lambda\mathbf{U}^\dagger$
    - Fact:

$$dA = \prod_i dA_{ii} \prod_{i>j} dReA_{ij} dImA_{ij} = d\mathbf{U} \prod_i dx_i \Delta(\{x_i\})^2$$

    - Where Δ is so-called Vandermonde determinant $\Delta(\{x_i\}) = \prod_{i>j}(x_i - x_j)$
    - **dU** is Haar measure (integrates over degrees of freedom of **U**)
    - Since nothing depends on **dU** can integrate out **U**
  - Then

$$\int p(A)dA \cdot \ldots = \prod_{i=1}^{N} \int dx_i p(\{x_i\}) \cdot \ldots$$

Saddle-point Approach

- Where $p(\{x_i\}) = e^{-NS(\{x_i\})}$

  $S = \sum_i V(x_i) - \frac{2}{N}\sum_{i>j} \ln|x_i - x_j|$
  - S is energy of N particles on a string located at positions x_i
    - In the presence of potential V(x)
    - Particles repel each other with logarithmic potential (2D electrostatic potential)
  - Total force on each particle is

    $$F_i = -\frac{\partial S}{\partial x_i} = -\frac{\partial V(x_i)}{\partial x_i} + \frac{2}{N}\sum_{j\neq i}\frac{1}{x_i - x_j}$$

    - No two particles can sit on top of each other: Prob. of equal eigenvalues = 0
      - Eigenvalue repulsion VERY IMPORTANT FEATURE OF RMT
    - Integrating over eigenvalue values $\int dx_i$ corresponds to integrating over all possible positions of particles
    - Saddle point solution corresponds to stationary energy w.r.t x_i , i.e. to zero force on each particle:

      $$\frac{\partial S}{\partial x_i} = 0 = F_i$$

      $$\frac{\partial V(x_i)}{\partial x_i} = \frac{2}{N}\sum_{j\neq i}\frac{1}{x_i - x_j}$$

Saddle-point Approach

- Analogy with thermodynamics
  - A (micro)state of a system is a vector of values of all variables in a system
    - E.g. positions, velocities, spins,…
  - Partition function is defined as

$$Z = \sum_{states} e^{-\frac{S(\{x_i\})}{k_B T}} \longleftrightarrow Z = \prod_i \int dx_i e^{-NS(\{x_i\})}$$

    - S energy of microstate, T temperature
  - Free energy contains all macroscopic information for system
    $F = -k_B T \ln Z$
  - Analogy $N \to \infty$ with T=0 (minimum energy state)
    - Careful usually take large N first and then T=0

Saddle-point Approach

- Thus for large N

$$g(z) = \frac{1}{N}\sum_{n=1}^{N} E\left[\frac{1}{z-x_n}\right] \approx \frac{1}{N}\sum_{n=1}^{N}\frac{1}{z-\tilde{x}_n}$$

  – $\tilde{x}_n$ are equilibrium positions of particles

- Now take square:

$$\begin{aligned} g(z)^2 &= \frac{1}{N^2}\sum_{i,j=1}^{N}\frac{1}{(z-\tilde{x}_i)(z-\tilde{x}_j)} \\ &\approx V'(z)g(z) - \frac{1}{N}\sum_{i=1}^{N}\frac{V'(z)-V'(\tilde{x}_i)}{z-\tilde{x}_i} \end{aligned}$$

  – Last term polynomial of order p-2, where order V(x) = p

$$g(z) = \frac{1}{2}\left(V'(z) - \sqrt{V'(z)^2 - 4Q(z)}\right)$$

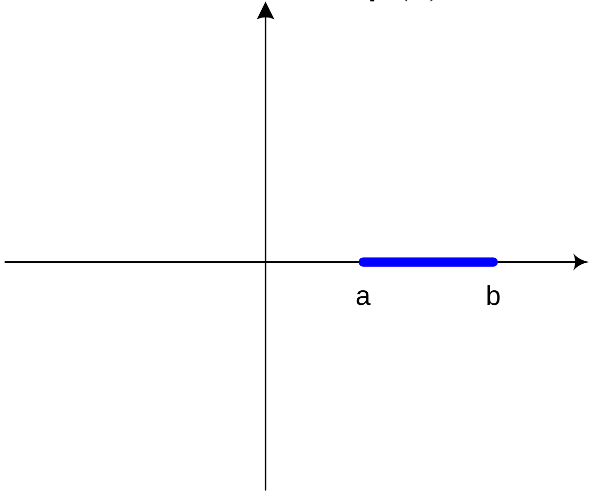
  - Have reduced problem to p-2 unknown numbers!
    – We know the p-1 order coefficient

Saddle-point Approach

- Need to think about what to expect for large N
  - Due to confining potential particles cannot wonder too far
  - Become packed at bottom of potential with distance ~1/N
  - Expect ρ(z) to be zero outside a certain range.
    - Within this range eigenvalues, at distance ~1/N form a continuum
- G(z) is analytic in z except on the real axis, where ρ(z)≠ 0.

$$\rho(z) = \frac{g(z - i0^+) - g(z + i0^+)}{2\pi i}$$

  - Branch cut where ρ(z) ≠ 0



Saddle-point Approach

- The form of $$g(z) = \frac{1}{2}\left(V'(z) - \sqrt{V'(z)^2 - 4Q(z)}\right)$$
  - Analytic except between (single) roots of polynomial under square root
    - Possibility of 2, 4, … 2(p-1) single roots, i.e. 1, 2, … (p-1) branch cuts where $\rho(z) \neq 0$
    - Start with 1 branch cut between a,b (unknown). If this solution is inconsistent then will have 2, 3, … cuts.
      $$V'(z)^2 - 4Q(z) = M(z)^2(z-a)(z-b)$$
    - Then
      $$g(z) = \frac{1}{2}\left(V'(z) - M(z)\sqrt{(z-a)(z-b)}\right)$$
    - Degree of polynomial M(z) = p-2, i.e. p-1 unknown coefficients, plus a,b
    - These coefficients can all be found by the condition that G(z)~1/z for large z
  - Density of eigenvalues takes form:
    $$\rho(z) = \Theta(z-a)\Theta(b-z)\frac{M(z)}{2}\sqrt{(z-a)(b-z)}$$
    - With M(z)>0 inbetween – otherwise one branch cut case not valid

Saddle-point Approach

- Examples:
  - Gaussian: $V(x) = \frac{x^2}{2\sigma^2}$
  - Degree P(z) =0, thus P(z)=A^2

    $$g(z) = \frac{1}{2}\left(\frac{z}{\sigma^2} - \sqrt{\frac{z^2}{\sigma^4} - 4A^2}\right)$$

  - At large z G(z) becomes $g(z) \approx \frac{1}{2\sigma^2}\left(z - z\left(1 - \frac{2A^2\sigma^4}{z^2}\right)\right) = \frac{A^2\sigma^2}{z}$
    - Thus A=1/σ resulting to

      $$g(z) = \frac{1}{2\sigma^2}\left(z - \sqrt{z^2 - 4\sigma^2}\right)$$

    - As before

Saddle-point Approach

- Examples:
  - Quartic: $$V(x) = \frac{a_2 x^2}{2} + \frac{g x^4}{4}$$
  - Degree M(z) =2, thus M(z)=Az^2 + Bz +C (also symmetry +-)

$$g(z) = \frac{1}{2}\left(a_2 z + g z^3 - (A z^2 + B z + C)\sqrt{(z^2 - a^2)}\right)$$

  - Taking the large z limit and thus determining A,B,C, a we get

$$g(z) = \frac{1}{2}\left(a_2 z + g z^3 - (g z^2 + a_2 + \frac{g a^2}{2})\sqrt{z^2 - a^2}\right)$$

    - Where $$a^2 = \frac{2}{3g}\left(\sqrt{a_2^2 + 12g} - a_2\right)$$
  - With density of eigenvalues $$\rho(z) = \left(g(z^2 + 0.5a^2) + a_2\right)\frac{\sqrt{a^2 - z^2}}{2\pi}$$
  - We see that for $a_2 < -2\sqrt{g}$
    - We get negative eigenvalues, ie. 1cut solution not valid (i.e. get 2 cut case)

Outline

- Diagrammatic approach to random matrix theory and applications to communications
- Saddle – point approach to random matrix theory
- Replicas and saddle - points

Introduction

- Statistical physics framework, replicas and its applications to multi-antenna systems and CDMA

- Random Matrix Theory of MIMO and CDMA
  - CDMA and multi-user detection
  - MIMO systems
  - Capacity

Idea behind CDMA

- Multi-User Detection
  - Use the same resources (channel, bandwidth, time) to receive (and successfully decode) information from several users
    - Divide and conquer: split resources "orthogonally" among different users
      - Use different frequency sub-bands for each user (OFDM)
      - Schedule users to transmit at different times (TDMA)
    - Orthogonality needs
      - Coordination/Synchronization (not always possible)
      - Many resources: over a period of N resources (subbands, time-slots, orthogonal codes) there can be up to N orthogonally transmitted users. The next one will "collide" with another one creating high interference (see OFDMA)
    - Non-orthogonal transmission: Spread the signal of each user over the whole resource pool, so that the presence of one user is felt in a mild way by all others, rather than strongly by a few:
      - No synchronization required
      - There are many more $\sim 2^N$ "random-looking" such ways to do it.
      - Random coding

Idea behind CDMA

- System Model

$$y_n = \sum_{m=1}^{K} S_{nm} x_n + z_n \qquad \text{or} \qquad \mathbf{y} = \mathbf{S}\mathbf{x} + \mathbf{z}$$

  - y_n received signal at chip n=1,…,N
  - z_n noise at chip n, assumed to be Gaussian distributed
  - x_m transmitted symbol from user m=1,…K
  - S_{nm} nth matrix element of code m
    - $\mathbf{S} = [\mathbf{s_1 s_2} \dots \mathbf{s_K}]$
    - Each element is ± 1, i.e. $S_{nm} = \pm 1$
    - For orthogonal codes elements are *not* i.i.d. (e.g. Haar matrices)
  - Each transmitted signal is normalized so that the total energy over the whole duration N is unity (on average)
    - For BPSK, $x_m = \frac{\pm 1}{N^{1/2}}$
    - For Gaussian transmitted signals $x_m \sim \mathcal{N}(0, 1/N)$
      - Can be seen as approximation of higher modulations, e.g. 16QAM etc

Capacity of CDMA

- Mutual Information of above model:
  - Assumption: all powers of transmitting users are equal
    - Rate of each user will be the same (on average)

$$c = \frac{1}{K} I(\mathbf{x}, \mathbf{y}|\mathbf{S})$$

  - Interested for large N, K limit, keeping their ratio fixed $\beta = \frac{N}{K}$
    - In this limit expect C to become equal with its S-average
    - S-average is also due to the changing of codes at every symbol

$$c = \lim_{K \leftarrow \infty} \frac{1}{K} I(\mathbf{x}, \mathbf{y}|\mathbf{S}) \approx \lim_{K \leftarrow \infty} \frac{1}{K} E\left[I(\mathbf{x}, \mathbf{y}|\mathbf{S})\right]_{\mathbf{S}}$$

  - 2 cases:
    - BPSK input
    - Gaussian input
  - Note: CDMA = for either input case, “channel” is binary $S_{nm} = \pm 1$
    - Due to “universality” this is equivalent to Gaussian i.i.d. channel

Analogy with MIMO

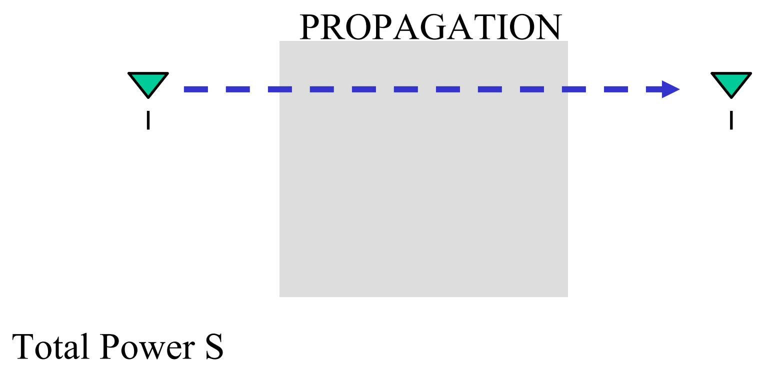
- System Model

$$y_n = \sum_{m=1}^{K} G_{nm} x_n + z_n$$

  - y_n received signal at antenna n=1,…,N
  - z_n noise at antenna n, assumed to be Gaussian distributed
  - x_m transmitted symbol from antenna m=1,…K
  - G_{nm} nth matrix element of antenna m

    - Each element is Gaussian (Rayleigh fading) generally correlated, but for simplicity assumed to be i.i.d $G_{nm} \sim \mathcal{CN}(0,1)$
    - Each transmitted signal is normalized so that the total energy from all antennas is S (on average)
    - For BPSK, $x_m = \frac{\pm 1}{N^{1/2}}$
    - For Gaussian transmitted signals $x_m \sim \mathcal{CN}(0, 1/N)$
      - Can be seen as approximation of higher modulations, e.g. 16QAM etc
      - Generally signals are correlated (beamforming etc), but here will assume i.i.d.

Analogy with MIMO (SISO)



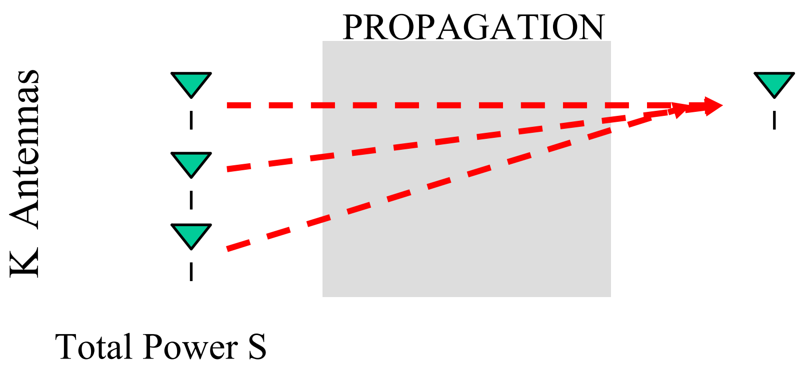
- Take transmission between 2 antennas
  - Analogy to CDMA = 1 user – no need to spread, i.e. code is 11111…

$$y = Gx + z$$

- Mutual information (for Gaussian input)

$$C = I(x, y|G) = \log(1 + S|G|^2)$$

Analogy with MIMO (MISO)



Assume Tx knows the instant. channel

K Tx Antennas – 1 Rx Antenna:

$$y = \mathbf{G}^{\dagger}\mathbf{x} + z$$

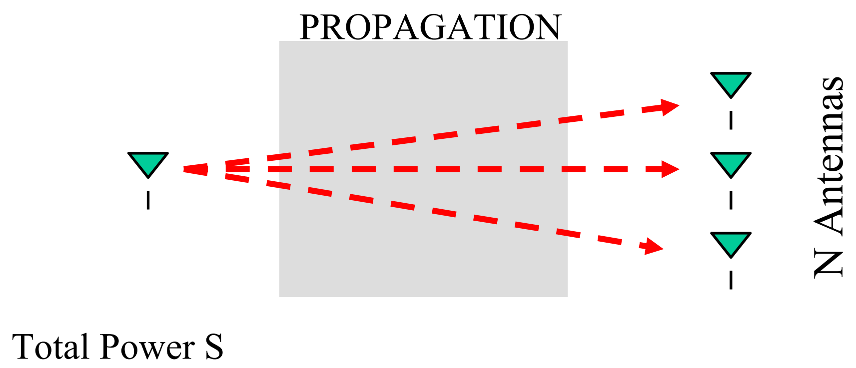
Optimal Transmission: Beamforming $\mathbf{x} = \mathbf{v}x = \dfrac{\mathbf{G}}{\sqrt{\mathbf{G}^{\dagger}\mathbf{G}}}x$

- Analogy with CDMA: K users– BF = cooperation between users

Resulting Capacity: $C = I(\mathbf{x}, y|\mathbf{G}) = \log(1 + S\mathbf{G}^{\dagger}\mathbf{G}) \approx \log(1 + KS)$

- Logarithmic capacity gain
- More robust system (lower realistic bit error rate)

Idea behind MIMO (SIMO)



1 Tx Antenna – N Rx Antennas:

Assume Tx knows the instant. channel

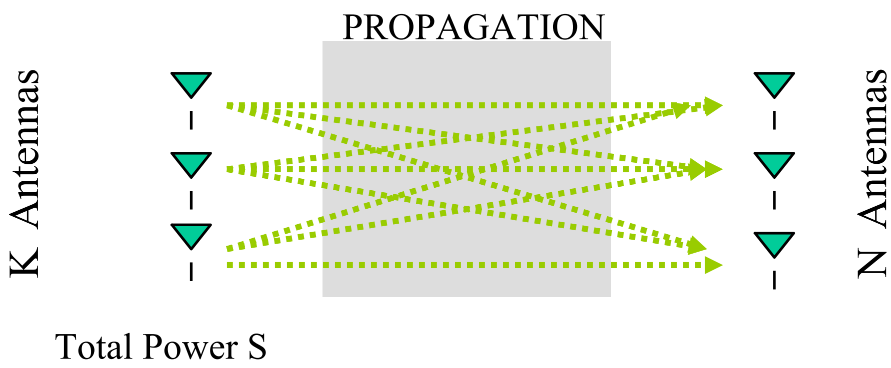
$$\mathbf{y} = \mathbf{G}x + \mathbf{z}$$

Optimal Reception: Beamforming $\hat{y} = \frac{\mathbf{G}^{\dagger}\mathbf{y}}{\sqrt{\mathbf{G}^{\dagger}\mathbf{G}}}$

•Analogy with CDMA: 1 user spreading – no gain other than log-SNR

Resulting Capacity: $C = I(\mathbf{y}, x|\mathbf{G}) = \log(1 + S\mathbf{G}^{\dagger}\mathbf{G}) \approx \log(1 + NS)$

Idea behind MIMO



K Tx Antennas – N Rx Antennas:

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{z}$$

Resulting Capacity:

- Only for scattering medium: G full rank
- Fixed Total Power S

Linear Increase with Antenna Number – similar to multiuser gain in CDMA

$$C = I(\mathbf{y}, x|\mathbf{G}) = \log\det(\mathbf{I} + \frac{S\mathbf{G}\mathbf{G}^{\dagger}}{K}) \approx K\log(1 + \frac{NS}{K})$$

Capacity of MIMO

- Mutual Information of MIMO model:
  - Assumption: all powers of transmitting antennas are equal, i.i.d

$$c = \frac{1}{K} I(\mathbf{x}, \mathbf{y}|\mathbf{G})$$

  - Ergodic capacity/antenna $c = \frac{1}{K} E\left[I(\mathbf{x}, \mathbf{y}|\mathbf{G})\right]_{\mathbf{G}}$
    - Channel “rides” all the waves within one symbol
  - Interested for large N, K limit
    - In this limit expect C to become equal with its S-average

$$c = \lim_{K\to\infty} \frac{1}{K} I(\mathbf{x}, \mathbf{y}|\mathbf{G}) \approx \lim_{K\to\infty} \frac{1}{K} E\left[I(\mathbf{x}, \mathbf{y}|\mathbf{G})\right]_{\mathbf{G}}$$

    - Would also like to variance around mean for outage statistics
  - 2 cases:
    - BPSK input
    - Gaussian input
  - G is usually assumed Gaussian – need not be Gaussian
    - We have seen that this is equivalent to Gaussian i.i.d. channel

MIMO – CDMA Analogy

CDMA		MIMO
$$y_n = \sum_{m=1}^{K} S_{nm} x_n + z_n$$		$$y_n = \sum_{m=1}^{K} G_{nm} x_n + z_n$$
K users		K Tx antennas
User cooperation		Tx Beamforming
N chips/code length		N Rx antennas
	Gaussian Noise (assume unit variance)	
Code matrix ± 1		Channel matrix
Entries i.i.d.	equivalent (large N)	Gaussian i.i.d.
	Input binary/Gaussian	

- NEXT: will calculate i.i.d. Gaussian channel, Gaussian input capacity
  - Similar in spirit with binary input case
  - Use (another) saddle-point approximation ("strange" math)

Capacity of MIMO/CDMA System with Gaussian input

- Will calculate average capacity and then assume ergodicity for large N

$$c = \lim_{K\to\infty} \frac{1}{K} I(\mathbf{x},\mathbf{y}|\mathbf{G}) \approx \lim_{K\to\infty} \frac{1}{K} E\left[I(\mathbf{x},\mathbf{y}|\mathbf{G})\right]_{\mathbf{G}}$$

  – Start with definition

$$C = E\left[I(\mathbf{x},\mathbf{y}|\mathbf{G})\right]_{\mathbf{G}} = E\left[h(\mathbf{y}|\mathbf{G})\right]_{\mathbf{G}} - E\left[I(\mathbf{y}|\mathbf{x},\mathbf{G})\right]_{\mathbf{G}}$$

  – For Gaussian input this takes simple (familiar) form:

$$C = E\left[\log\det\left(\mathbf{I}_N + \rho\mathbf{G}^\dagger\mathbf{G}\right)\right]_{\mathbf{G}}$$

    - Here take variance of i.i.d. elements of G =1/N
    - For non-Gaussian (binary) input arguments are similar, except algebra harder
    - Looks like $C = E\left[\log Z(\mathbf{G})\right]_{\mathbf{G}}$
    - Z partition function (counts number of states of system)
      – Here det counts words that can be sent
    - Average outside log means averaging over quenched variables

  – Difficult to do analytically

    - Although this simple form has been done exactly in many different ways (!)
    - Resort to scary math tricks (that work)

Capacity of MIMO/CDMA System with Gaussian input

- Need to calculate average over logs
  - Use simple "replica" trick:

$$\log Z = -\frac{d}{dx} Z^{-x}\Big|_{x=0}$$

- So now need to do:

$$\begin{aligned} E\left[\log\det\left(\mathbf{I}_N + \rho\mathbf{G}^\dagger\mathbf{G}\right)\right]_{\mathbf{G}} &= -\frac{d}{dx} E\left[\det\left(\mathbf{I}_N + \rho\mathbf{G}^\dagger\mathbf{G}\right)^{-x}\right]_{\mathbf{G}}\Big|_{x=0} \\ &= -\frac{d}{dx} g(x)\Big|_{x=0} \end{aligned}$$

  - g(x) is moment generating function of C
  - If now differentiate w.r.t. ρ get resolvent
    - methods equivalent with diagrammatics
  - Still difficult: x needs to vary continuously at 0
  - Is there anything doable?
    - Calculate at x=integer might be (more) doable
    - Then analytically continue to x=0

Capacity of MIMO/CDMA System with Gaussian input

- Analytic continuation of g(x) to x=0
  - This is a valid mathematical argument for this particular g(x)

- Thus need to calculate g(n) for n=1,2,3,…
- 6 easy steps

- Step 1
  - Identity 1 (for **X** vector Nx1, **M** NxN matrix)

$$[\det \mathbf{M}]^{-1} = \int d\mathbf{X} \exp\{-\frac{1}{2}\mathbf{X}^{\dagger}\mathbf{M}^{-1}\mathbf{X}\}$$

  - Use this identity n times
    - Make **X** an Nxn matrix

$$\left[\det\left(\mathbf{I}+\mathbf{G}^{\dagger}\mathbf{G}\right)\right]^{-n} = \int d\mathbf{X} \exp\{-\frac{1}{2}Tr\left[\mathbf{X}^{\dagger}\mathbf{X}+\mathbf{X}^{\dagger}\mathbf{G}^{\dagger}\mathbf{G}\mathbf{X}\right]\}$$

Capacity of MIMO/CDMA System with Gaussian input

- Step 2:
  - Identity 2 (completing the square)

$$\exp\{-\frac{1}{2}\mathbf{A}^\dagger\mathbf{A}\} = \int d\mathbf{Y}\,\exp\{-\frac{1}{2}Tr\left[\mathbf{Y}^\dagger\mathbf{Y} + \mathbf{A}^\dagger\mathbf{Y} - \mathbf{Y}^\dagger\mathbf{A}\right]\}$$

  - Apply it to

$$\exp\{-\frac{1}{2}Tr\left[\mathbf{X}^\dagger\mathbf{G}^\dagger\mathbf{G}\mathbf{X}\right]\}$$

$$\left[\det\left(\mathbf{I} + \mathbf{G}^\dagger\mathbf{G}\right)\right]^{-n} =$$
$$\int d\mathbf{X}\int d\mathbf{Y}\,\exp\{-\frac{1}{2}Tr\left[\mathbf{X}^\dagger\mathbf{X} + \mathbf{Y}^\dagger\mathbf{Y} + \mathbf{X}^\dagger\mathbf{G}^\dagger\mathbf{Y} - \mathbf{Y}^\dagger\mathbf{G}\mathbf{X}\right]\}$$

  - Now we only have linear terms in G in the exponent
    - May now average over Gaussian G

Capacity of MIMO/CDMA System with Gaussian input

- Step 3:
  - Use Identity 2 (again, but the other way around)

$$\exp\{-\frac{1}{2}Tr\left[\mathbf{X}^\dagger\mathbf{X}\mathbf{Y}^\dagger\mathbf{Y}\right]\} = \int d\mathbf{G}\exp\{-\frac{1}{2}Tr\left[\mathbf{G}^\dagger\mathbf{G} - \mathbf{X}\mathbf{Y}^\dagger\mathbf{G} + \mathbf{G}^\dagger\mathbf{Y}\mathbf{X}^\dagger\right]\}$$

  - Thus

$$E\left[\left[\det\left(\mathbf{I}+\mathbf{G}^\dagger\mathbf{G}\right)\right]^{-n}\right] =$$
$$\int d\mathbf{X}\int d\mathbf{Y}\exp\{-\frac{1}{2}Tr\left[\mathbf{X}^\dagger\mathbf{X}+\mathbf{Y}^\dagger\mathbf{Y}+\frac{1}{2}\mathbf{X}^\dagger\mathbf{X}\mathbf{Y}^\dagger\mathbf{Y}\right]\}$$

  - Not there yet: have exchanged problem of det with problem of quartic term
  - Make ansatz: not all terms $\left[\mathbf{X}^\dagger\mathbf{X}\right]_{ab}$ contribute the same in the integral
  - Smells like saddle point: Tr in exponent gives O(N)
  - Try to separate fast from slow variables

Capacity of MIMO/CDMA System with Gaussian input

- Step 4:
  - Try to separate fast from slow variables
  - Multiply with integrals:

$$\int \prod_{a,b=1}^{n} dQ_{ab}\delta\left(Q_{ab} - \mathbf{X}_a^{\dagger}\mathbf{X}_b\right)$$

  - Write δ-function as Fourier – integral (note the 2πi in denom)

$$\int \prod_{a,b=1}^{n} \frac{dQ_{ab}dR_{ab}}{2\pi i} \exp\left(Tr\left[R\left(Q - \mathbf{X}^{\dagger}\mathbf{X}\right)\right]\right)$$

  - Thus

$$E\left[\left[\det\left(\mathbf{I}+\mathbf{G}^{\dagger}\mathbf{G}\right)\right]^{-n}\right] = \int d\mathbf{Q}d\mathbf{R}\int d\mathbf{X}\int d\mathbf{Y}\,\cdot$$

$$\exp\{-\frac{1}{2}Tr\left[\mathbf{X}^{\dagger}\mathbf{X}+\mathbf{Y}^{\dagger}\mathbf{Y}+\mathbf{Q}\mathbf{Y}^{\dagger}\mathbf{Y}+\mathbf{R}\mathbf{X}^{\dagger}\mathbf{X}-2\mathbf{R}\mathbf{T}\right]\}$$

Capacity of MIMO/CDMA System with Gaussian input

- Step 5:
  - Integrate over X, Y using Identity 1:

$$E\left[\left[\det\left(\mathbf{I}+\mathbf{G}^{\dagger}\mathbf{G}\right)\right]^{-n}\right]=\int d\mathbf{Q}d\mathbf{R}e^{-N\mathcal{S}}$$

$$\mathcal{S}=\log\det\left[\mathbf{I_n}+\mathbf{Q}\right]+\log\det\left[\mathbf{I_n}+\frac{\rho}{N}\mathbf{R}\right]-\frac{1}{N}Tr\left[\mathbf{RT}\right]$$

  - S is O(n), but due to N in front, can use saddle-point methods
    - Remember eventually n=0, but for now assume it is some number
  - Need to find saddle-point:
    - Could search for all points that satisfy saddle point equations, i.e.

$$\frac{\partial\mathcal{S}}{\partial Q_{ab}}=\frac{\partial\mathcal{S}}{\partial R_{ab}}=0$$

    - Instead, will cheat, by looking at symmetry of Q,R. Since: $Q_{ab}=\mathbf{X}_a^{\dagger}\mathbf{X}_b$
    - Q should have the transformation properties of $X_a^{+}X_b$
    - Since X_b is complex Gaussian (vector), so should $U_{ba}X_a$ for U unitary transformation U(n). Thus Q should be U(n) symmetric (similar for R). Therefore look for saddle points with this symmetry
    - **This is not true in the binary case, where the symmetry is just +-**

Capacity of MIMO/CDMA System with Gaussian input

- Step 6:
  - Insert U(n) symmetric saddle point solution into S and find saddle-point equations
    $$\begin{aligned} Q^0_{ab} &= q\delta_{ab} \\ R^0_{ab} &= rN\delta_{ab} \end{aligned}$$
  - Then
    $$\mathcal{S} = n\left(\log\left[1+q\right] + \log\left[1+\rho r\right] - rq\right) + O(n^2) = n\mathcal{S}_0 + O(n^2)$$
  - q, r satisfy equations:
    $$\begin{aligned} r &= \frac{1}{1+q} \\ q &= \frac{\rho}{1+\rho r} \end{aligned}$$

Capacity of MIMO/CDMA System with Gaussian input

- Done:

$$g(n) \quad = \quad E\left[\left[\det\left(\mathbf{I}+\mathbf{G}^{\dagger}\mathbf{G}\right)\right]^{-n}\right] \approx e^{-Nn\mathcal{S}_0}$$

  - Analytically continuing n to real values and taking derivative give us

$$\begin{aligned} E\left[\log\det\left(\mathbf{I}_N+\rho\mathbf{G}^{\dagger}\mathbf{G}\right)\right]_{\mathbf{G}} &= -\frac{d}{dx}g(x)\Big|_{x=0} \\ &= N\mathcal{S}_0 \end{aligned}$$

    - After inserting the correct values of q,r we get:

$$C = N\left(\log\left[\rho\frac{\sqrt{1+4\rho}+1}{\sqrt{1+4\rho}-1}\right] - \frac{\sqrt{1+4\rho}-1}{\sqrt{1+4\rho}+1}\right)$$

- Can generalize for:
  - Correlated channels
  - Binary input
  - Variance of distribution (Gaussian entries)

Capacity of MIMO/CDMA System with Gaussian input

- Variance of distribution:

$$g(n) = \int d\mathbf{Q}d\mathbf{R}e^{-N\mathcal{S}} \approx e^{-Nn\mathcal{S}_0}\int d\delta\mathbf{Q}d\delta\mathbf{R}e^{-\mathcal{S}_2}$$

$$\mathcal{S}_2 = \frac{1}{2}\sum_{a,b=1}^{n}[\delta Q_{ba}\delta R_{ba}]\begin{pmatrix} -Nr^2 & -1 \\ -1 & -q^2/N \end{pmatrix}[\delta Q_{ab}\delta R_{ab}]^T$$

  – Recall factor of *i* in integral measure

$$\int\prod_{a,b=1}^{n}\frac{dQ_{ab}dR_{ab}}{2\pi i} = \int\prod_{a,b=1}^{n}\frac{d\delta Q_{ab}d\delta R_{ab}}{2\pi i}$$

  – To absorb the *i*'s take

$$\delta\mathbf{Q} = \delta\mathbf{Q}^\dagger$$
$$\delta\mathbf{R} = -\delta\mathbf{R}^\dagger$$

  – After making a rotation to diagonalize the Hessian and taking real and complex parts separately we get

Capacity of MIMO/CDMA System with Gaussian input

- MGF of mutual information

$$
\begin{aligned}
g(n) &\approx e^{-Nn\mathcal{S}_0} \prod_{a,b}^{n} \left(1 - q^2 r^2\right)^{1/2} \\
&= e^{-Nn\mathcal{S}_0 + \frac{n^2}{2} \log\left(1 - q^2 r^2\right)}
\end{aligned}
$$

  – Second derivative of log(g(n)) at n=0 will give variance, hence

$$
Var(C) = -\log\left[1 - q^2 r^2\right] = -\log\left[1 - \left(\frac{q}{1+q}\right)^2\right] > 0
$$

  – Higher moments are O(1/N)
  – C is Gaussian to very good accuracy even for N=2!
- Stability of Replica Symmetry Assumption (related to the Hessian eigenvalues) is equivalent to above result!!!
  – Thus Replica Symmetry is always stable (not surprising here due to continuous symmetry of **Q**)

Recap

- Replica approach gives results for first few moments of channel capacities (for Gaussian channels).
  - It works!!!
    - When it doesn’t there is a very good reason why (and there ways to know it).
  - Easy to implement to different cases
  - Connections with other approaches (diagrammatic, free probability)
  - Second order statistics trivial once found saddle point

References

- A. M. Tulino and S. Verdú. "*Random matrix theory and wireless communications,*" Foundations and Trends in Communications and Information Theory, **1**(1), June 2004.
- P. W. Brouwer and C. W. J. Beenakker, "*Diagrammatic method of integration over the unitary group, with applications to quantum transport,*" J. Math. Phys. **37** (10), p. 4904, Oct. 1996.
- N. Argaman and A. Zee, "*Diagrammatic theory of random scattering matrices for normal-metal-superconducting mesoscopic junctions*," Phys. Rev. B **54**, p. 7406, Sept. 1996.
- S. Verdú, "*Multiuser Detection,*" Cambridge University Press, New York, 1998.
- M. L. Mehta, "Random Matrices", Elsevier Academic Press, Amsterdam, 2004.
- A. L. Moustakas, S. H. Simon and A. M. Sengupta, "MIMO Capacity Through Correlated Channels in the Presence of Correlated Interferers and Noise: A (Not So) Large N Analysis," IEEE Trans. Information Theory, vol. 49 (10), p. 2545, Oct. 2003.