CAPACITY OF MIMO WIRELESS CHANNELS VIA VIRTUAL REPRESENTATION

Venu Veeravalli
National Science Foundation
&
Dept of ECE and Coordinated Science Lab
University of Illinois at Urbana Champaign

vveerava@nsf.gov, vvv@uiuc.edu
http://www.comm.csl.uiuc.edu/~vvv

(Joint work with Yingbin Liang and Akbar Sayeed)

Problem Statement

- **MIMO channel model** (over one symbol period)

$$\underline{\boldsymbol{y}} = \sqrt{\frac{\Gamma}{n_t}}\, \boldsymbol{H}\, \underline{\boldsymbol{x}} + \underline{\boldsymbol{w}}$$

  - ◦ n_t transmit and n_r receive antennas
  - ◦ power constraint: $\mathsf{E}[\boldsymbol{x}^\dagger \boldsymbol{x}] \leq n_t$

- **Channel Statistics**
  - ◦ elements of $\boldsymbol{H}$ identically distributed with:
    - ☞ finite variance (normalized to 1)
    - ☞ $\boldsymbol{H}_{k,\ell}$ having same distribution as $-\boldsymbol{H}_{k,\ell}$
    - ☞ possible correlation among elements

- **Coherent channel assumption**
  - ◦ $\boldsymbol{H}$ known at receiver; statistics of $\boldsymbol{H}$ known at Tx

- Ergodic Capacity and Optimal Input Distribution?

- Asymptotics as $n_t, n_r \to \infty$?

Previous Work

☞ **I.i.d model for $\boldsymbol{H}$**

- Capacity; closed-form asymptotic capacity [Telatar '99]
- Accuracy of asymptotics [Kamath&Hughes '02]

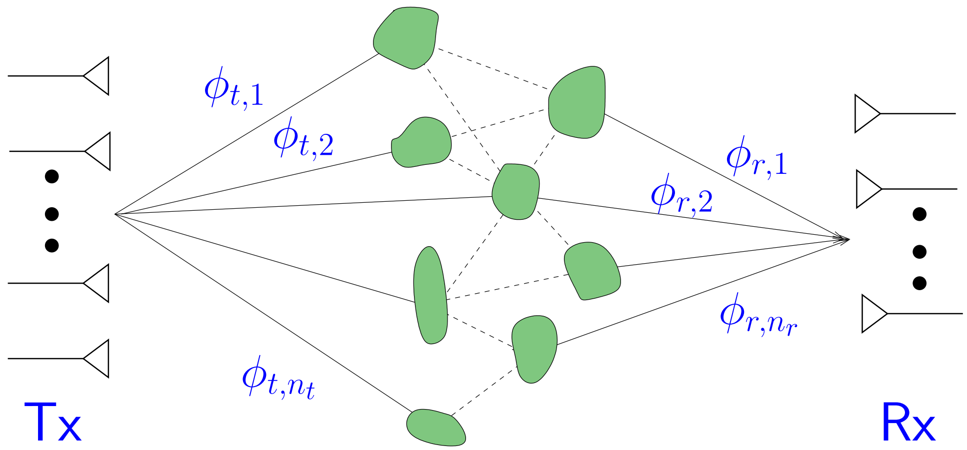
☞ **Correlated model for $\boldsymbol{H}$**

- Characterization of optimal input; optimality of beamforming
  Visotsky & Madhow (MISO case)
  Jafar & Goldsmith
  Jorswieck & Boche
  Simon & Moustakas
- Asymptotics and capacity scaling [Chua et al '02]

☞ Product form correlation

$$\mathsf{E}\left[\boldsymbol{H}_{k,\ell}\,\boldsymbol{H}^{\star}_{p,q}\right] = \rho^{t}_{i,p}\,\rho^{r}_{j,q}$$

Channel Model for ULA's



- **Scaled (virtual) angle**

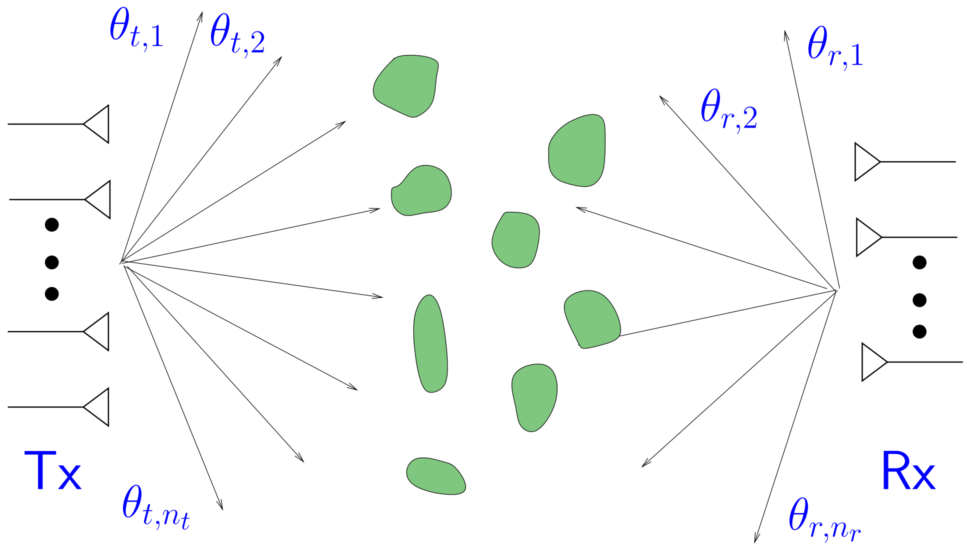
$$\theta = \frac{d(\sin(\phi)+1)}{\lambda_c} = 0.5(\sin(\phi)+1) \quad \text{(half-wavelength spacing)}$$

- **Physical Channel Model**

$$\boldsymbol{H} = \sqrt{n_t, n_r} \int_0^1 \int_0^1 \boldsymbol{\alpha}(\theta_r, \theta_t)\, \underline{a}_r(\theta_r)\, \underline{a}_t^{\dagger}(\theta_t)\, d\theta_r\, d\theta_t$$

$$\underline{a}_t(\theta_t) = \frac{1}{\sqrt{n_t}} \left[1,\ e^{-j2\pi(\theta_t-0.5)}, \cdots, e^{-j2\pi(n_t-1)(\theta_t-0.5)}\right]^{\top}$$

Virtual Representation



- Steering along n_t and n_r virtual angles spaced equally in $[0, 1]$

$$\begin{aligned} A_r &= [\underline{a}_r(\theta_{r,1}), \underline{a}_r(\theta_{r,2}), \cdots, \underline{a}_r(\theta_{r,n_r})] \\ A_t &= [\underline{a}_t(\theta_{t,1}), \underline{a}_t(\theta_{t,2}), \cdots, \underline{a}_t(\theta_{t,n_t})] \end{aligned}$$

- A_r and A_t are unitary discrete Fourier transform matrices

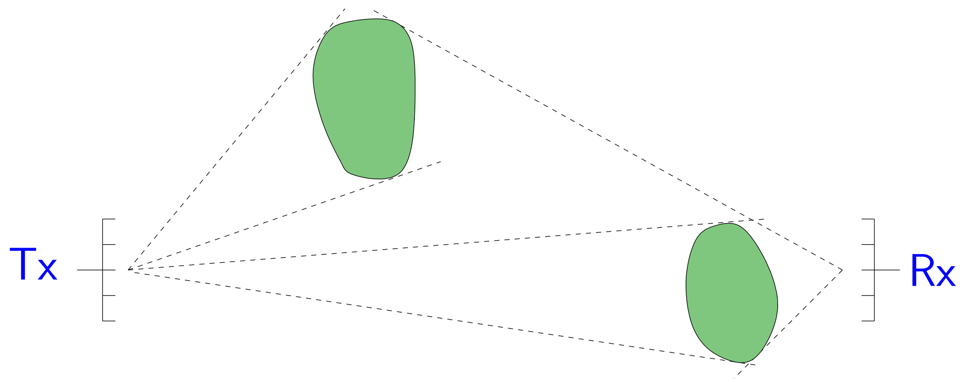
- **Virtual Channel Model**

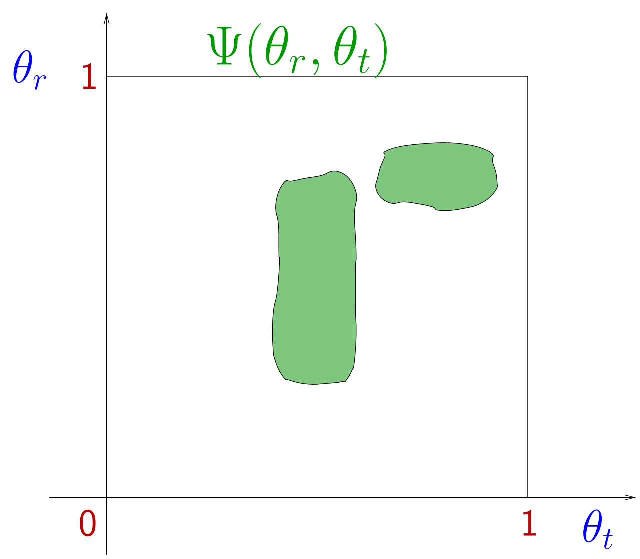
$$\boldsymbol{H} = A_r \, \tilde{\boldsymbol{H}} A_t^{\dagger}$$

Spatial Scattering Function

- Uncorrelated scattering assumption

$$\mathsf{E}[\boldsymbol{\alpha}(\theta_r,\theta_t)\ \boldsymbol{\alpha}^*(\theta_r',\theta_t')] = \Psi(\theta_r,\theta_t)\ \delta(\theta_r-\theta_r')\ \delta(\theta_t-\theta_t')$$





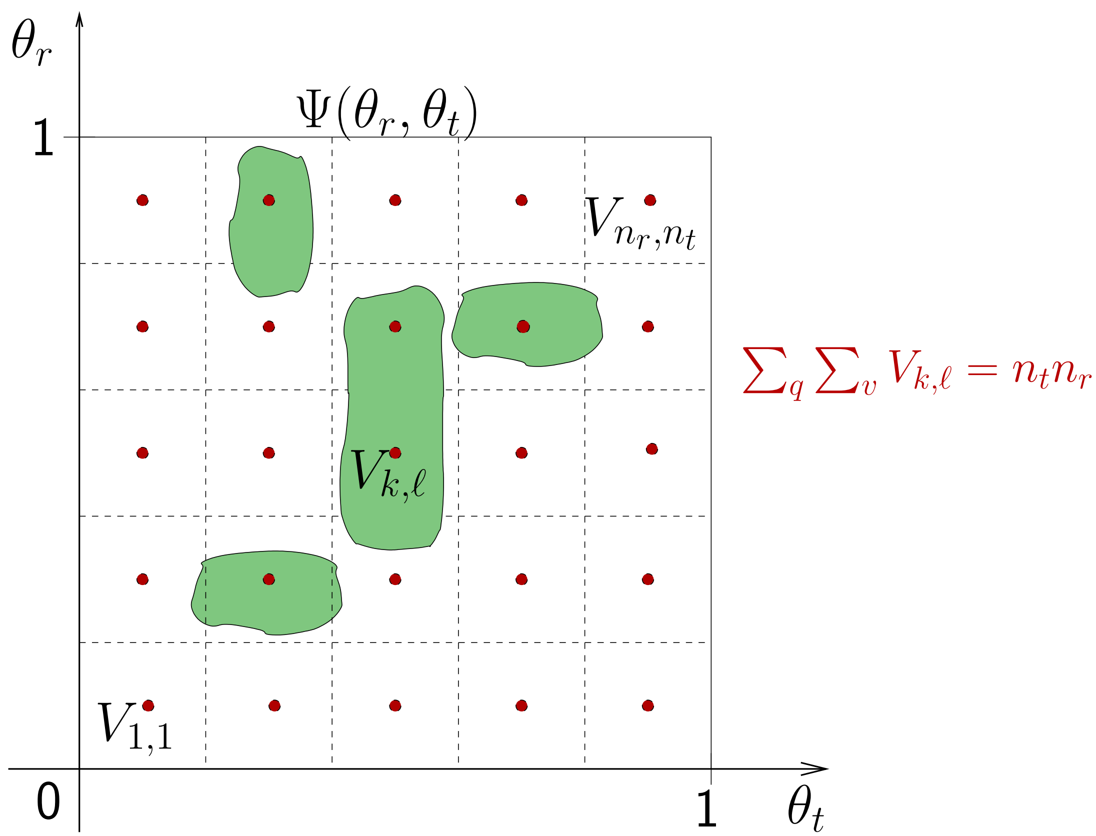
$$\int\int \Psi(\theta_r,\theta_t)\ d\theta_r\ d\theta_t = 1$$

Virtual Channel Coefficients

$$\boldsymbol{H} = A_r \, \tilde{\boldsymbol{H}} A_t^\dagger$$

- ☞ **Key Property:** $\tilde{\boldsymbol{H}}_{k,\ell}$ approximately uncorrelated

- ☞ **Variance Matrix:** V with $V_{k,\ell} = \mathsf{E}[|\tilde{\boldsymbol{H}}_{k,\ell}|^2] \approx \Psi(\theta_{r,k}, \theta_{t,\ell})$



Ergodic Capacity in Virtual Domain

$$\underline{\boldsymbol{y}} = \sqrt{\frac{\Gamma}{n_t}}\, \boldsymbol{H}\ \underline{\boldsymbol{x}} + \underline{\boldsymbol{w}}$$

☞ Capacity achieved by zero-mean proper complex Gaussian $\underline{\boldsymbol{x}}$ with covariance Q that satisfies $\mathrm{Tr}(Q) \leq n_t$

$$C = \max_{Q:\mathrm{Tr}(Q)\leq n_t} \mathsf{E}\left[\log\det\left(I + \frac{\Gamma}{n_t}\boldsymbol{H}Q\boldsymbol{H}^{\dagger}\right)\right]$$

$$\underline{\tilde{\boldsymbol{y}}} = \sqrt{\frac{\Gamma}{n_t}}\tilde{\boldsymbol{H}}\underline{\tilde{\boldsymbol{x}}} + \underline{\tilde{\boldsymbol{w}}}$$

☞ $\underline{\tilde{\boldsymbol{x}}} = A_t^{\dagger}\underline{\boldsymbol{x}}$; $\underline{\tilde{\boldsymbol{y}}} = A_r^{\dagger}\underline{\boldsymbol{y}}$; $\underline{\tilde{\boldsymbol{w}}} = A_r^{\dagger}\underline{\boldsymbol{w}}$

☞ Capacity achieved by zero-mean proper complex Gaussian $\underline{\tilde{\boldsymbol{x}}}$ with covariance $\tilde{Q}$ that satisfies $\mathrm{Tr}(\tilde{Q}) \leq n_t$

$$C = \max_{\tilde{Q}:\mathrm{Tr}(\tilde{Q})\leq n_t} \mathsf{E}\left[\log\det\left(I + \frac{\Gamma}{n_t}\tilde{\boldsymbol{H}}\tilde{Q}\tilde{\boldsymbol{H}}^{\dagger}\right)\right]$$

Optimal Input Distribution

$$C = \max_{\tilde{Q}:\mathrm{Tr}(\tilde{Q})\leq n_t} \mathsf{E}\left[\log\det\left(I + \frac{\Gamma}{n_t}\tilde{\boldsymbol{H}}\tilde{Q}\tilde{\boldsymbol{H}}^{\dagger}\right)\right]$$

- **Result:** Optimal $\tilde{Q}$ is a diagonal matrix Λ°
  - signals sent along the different Tx angles $\theta_{t,i}$ independent
  - λ_i power allocated to angle $\theta_{t,i}$
  - optimal λ_i's easily found numerically

- Optimal Q satisfies $Q^{\circ} = A_t \Lambda^{\circ} A_t^{\dagger}$

- **Result:** At sufficiently low SNR, only one element of Λ° is nonzero
  - beamforming along Tx angle $\theta_{t,i}$ with largest total power gain is optimal
  - can also find necessary and sufficient condition for beamforming to be optimal in terms of variance matrix V

Asymptotically Optimal Power Allocation at Low SNR

Theorem. At low SNR, optimal Λ° has all elements equal to zero except that $\lambda_i^\circ = n_t$ with index i identified by $i = \arg \max_{1 \leq \ell \leq n_t} \sum_{k=1}^{n_r} V_{k,\ell}$.

If maximizing index is not unique, define index set

$$\mathcal{T} = \left\{ i : i = \arg \max_{1 \leq \ell \leq n_t} \sum_{k=1}^{n_r} V_{k,\ell} \right\}$$

Then Λ° is such that

$$\sum_{i:i \in \mathcal{T}} \lambda_i^\circ = n_t, \ \lambda_i^\circ \geq 0 \quad \text{for } i \in \mathcal{T}$$
$$\text{and} \quad \lambda_i^\circ = 0, \quad \text{for } i \notin \mathcal{T}$$

i.e., power is arbitrarily assigned to diagonal elements corresponding to those maximizing indexes without changing capacity as long as total power is n_t.

Condition for Beamforming to be Optimal

Theorem. A necessary and sufficient condition for beamforming to i-th virtual angle to be optimal is given by

$$\Gamma \sum_{k=1}^{n_r} (1 - \mu_{k,i})\, V_{k,\ell^\circ} - \sum_{k=1}^{n_r} \mu_{k,i} \leq 0,$$

where $\ell^\circ = \arg \max_{\substack{1 \leq \ell \leq n_t \\ \ell \neq i}} \sum_{k=1}^{n_r} (1 - \mu_{k,i})\ V_{k,\ell}$. The functions $\mu_{k,i}$ are defined as

$$\mu_{k,i} := \mu_{k,i}\left(V_{1,i}, V_{2,i}, \ldots, V_{n_r,i}\right) = \mathsf{E}\left[\frac{\Gamma\ |\tilde{\boldsymbol{H}}_{k,i}|^2}{1 + \Gamma \|\underline{\tilde{\boldsymbol{h}}}_i\|^2}\right]$$

Among above n_t conditions corresponding to $1 \leq i \leq n_t$, at most one can be satisfied.

Example of Correlated MIMO Channel

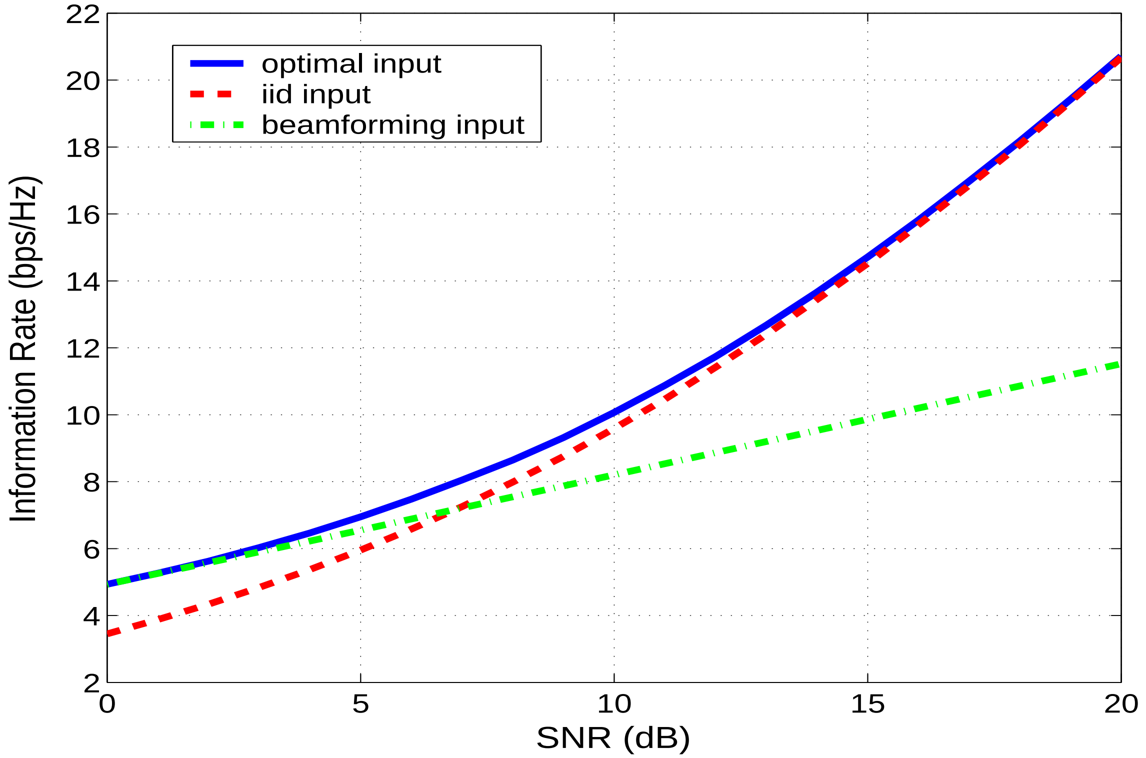
- An example: 5×5 MIMO channel with variance matrix

$$V = \frac{25}{5.7} \begin{pmatrix} 0.1 & 0 & 1 & 0 & 0 \\ 0 & 0.1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0.25 & 0 \\ 0 & 0 & 1 & 0 & 0.25 \end{pmatrix}$$

- Variance matrix V represents physical environment with two small scatterers, two bigger scatterers, and one large scattering cluster

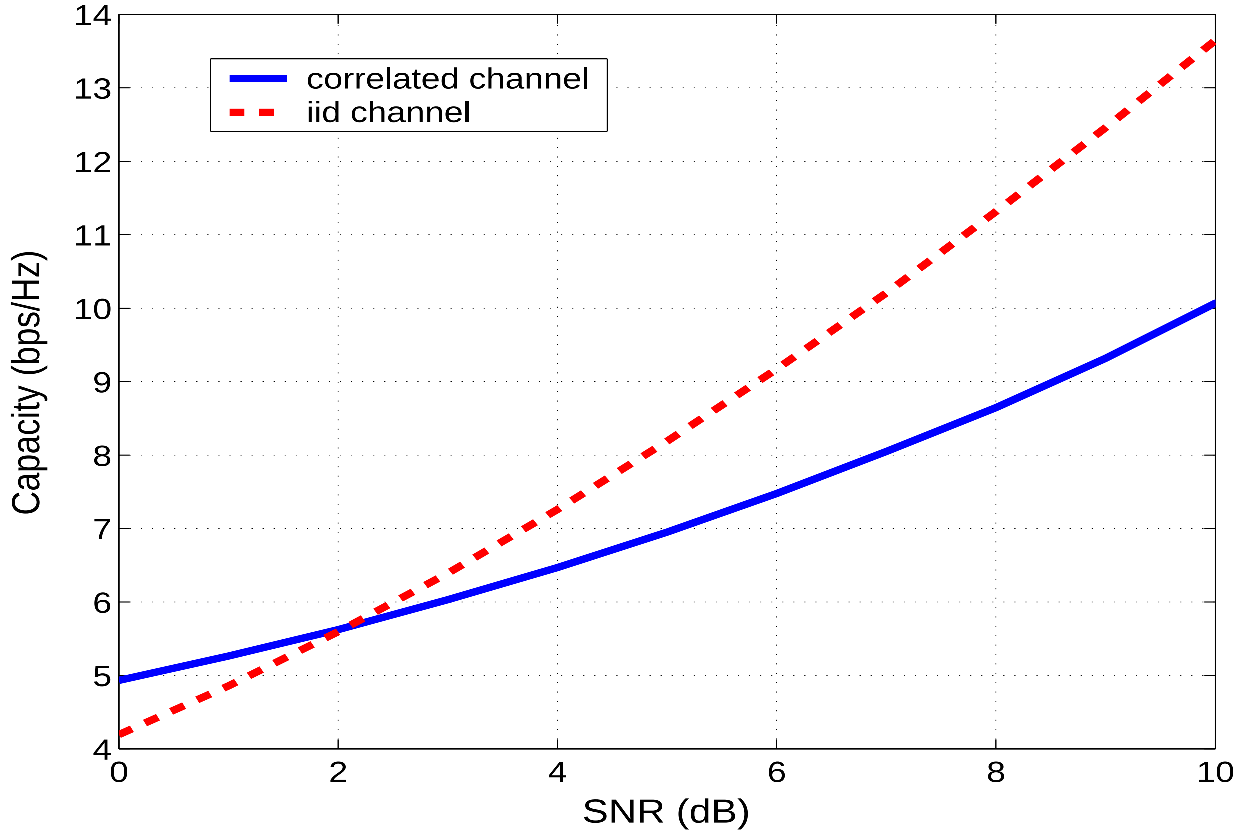
- Original $\boldsymbol{H}$ has highly correlated entries, because V has small fraction of dominant entries

Numerical Results for Channel in Example



- Information rate is improved by using optimal inputs; improvement becomes less significant as SNR increases
- Beamforming is optimal for SNR's below 0.29 dB, and remains close to optimal for SNR's below 5 dB

Effect of Correlation on Capacity



- Corr. channel has larger capacity than i.i.d. channel for SNR's below 2 dB
- Multiplexing gain offered by i.i.d. channel manifests itself only at sufficiently high SNR's

Asymptotics for Large Number of Antennas

$$C = \mathsf{E}\left[\log\det\left(I + \frac{\Gamma}{n_t}\tilde{\boldsymbol{H}}\Lambda^{\circ}\tilde{\boldsymbol{H}}^{\dagger}\right)\right]$$

- For simplicity, let $n_t = n_r = n$, and let $n \to \infty$

- **Normalized asymptotic capacity**

$$\bar{C} = \frac{1}{n}\left[\log\det\left(I + \frac{\Gamma}{n}\tilde{\boldsymbol{H}}\Lambda^{\circ}\tilde{\boldsymbol{H}}^{\dagger}\right)\right]$$

- **Key quantity:** limiting eigenvalue distribution of $\tilde{\boldsymbol{H}}\Lambda^{\circ}\tilde{\boldsymbol{H}}^{\dagger}/n$
  - ◦ i.i.d. case is relatively easy – Wishart distribution [Teletar '99]
  - ◦ we exploit independence of entries of $\tilde{\boldsymbol{H}}$ in correlated case to apply random matrix result of Girko
  - ◦ Λ° being diagonal is necessary to apply Girko's result

Limiting Input Power Profile and Stieltjes Transform

Assumption. For each n, define the function $s_n : [0, \tau] \to \Re$ by

$$s_n(v) = \lambda_\ell \,, \quad \text{for } v \in \left[\frac{\ell - 1}{n}, \frac{\ell}{n}\right]$$

where $\ell = 1, \cdots, \lfloor n\tau \rfloor$. Then $s_n(v)$ is bounded for each n, and converges uniformly to a limiting bounded function $s(v)$ as $n \to \infty$.

Definition. The Stieltjes transform m_A of a $n \times n$ Hermitian matrix A is defined as

$$m_A(z) = \frac{1}{n} \mathrm{Tr}\{(A - zI)^{-1})\}$$

Asymptotic Capacity

☞ **Result 1:** Assume that the Stieltjes transform of $\tilde{\boldsymbol{H}}\Lambda\tilde{\boldsymbol{H}}^{\dagger}/n$ converges to a limit denoted by $m(z)$ as $n \to \infty$. Then

$$\lim_{n\to\infty} \bar{C} = \int_0^1 \frac{1}{t}\left(1 - \frac{1}{t\Gamma} m\left(-\frac{1}{t\Gamma}\right)\right) dt \qquad (1)$$

☞ **Result 2 [Girko]:** The limiting Stieltjes transform of $\tilde{\boldsymbol{H}}\Lambda\tilde{\boldsymbol{H}}^{\dagger}/n$ exists as $n \to \infty$, and it is given by:

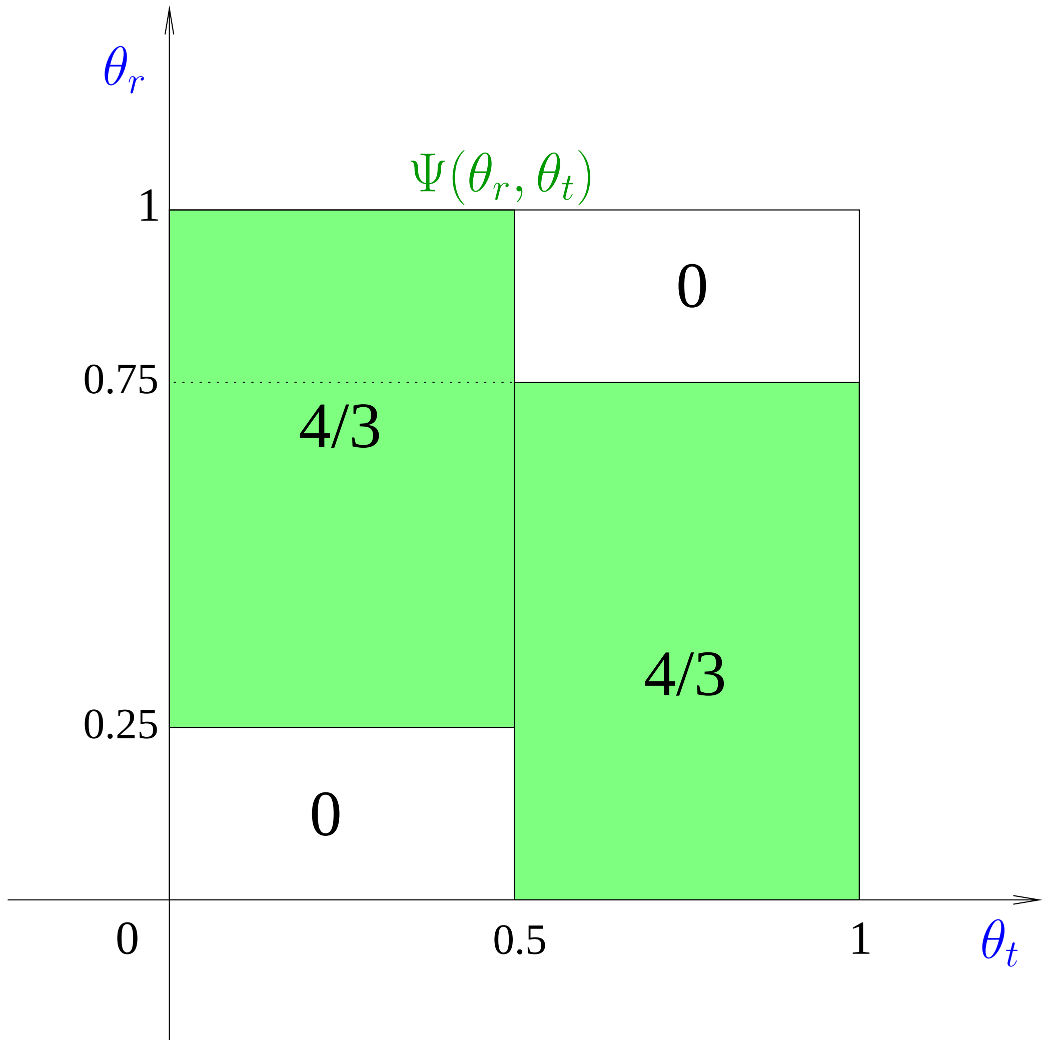
$$m(z) = \int_0^1 e(u, z) du \qquad (2)$$

where $e(u, t)$ satisfies

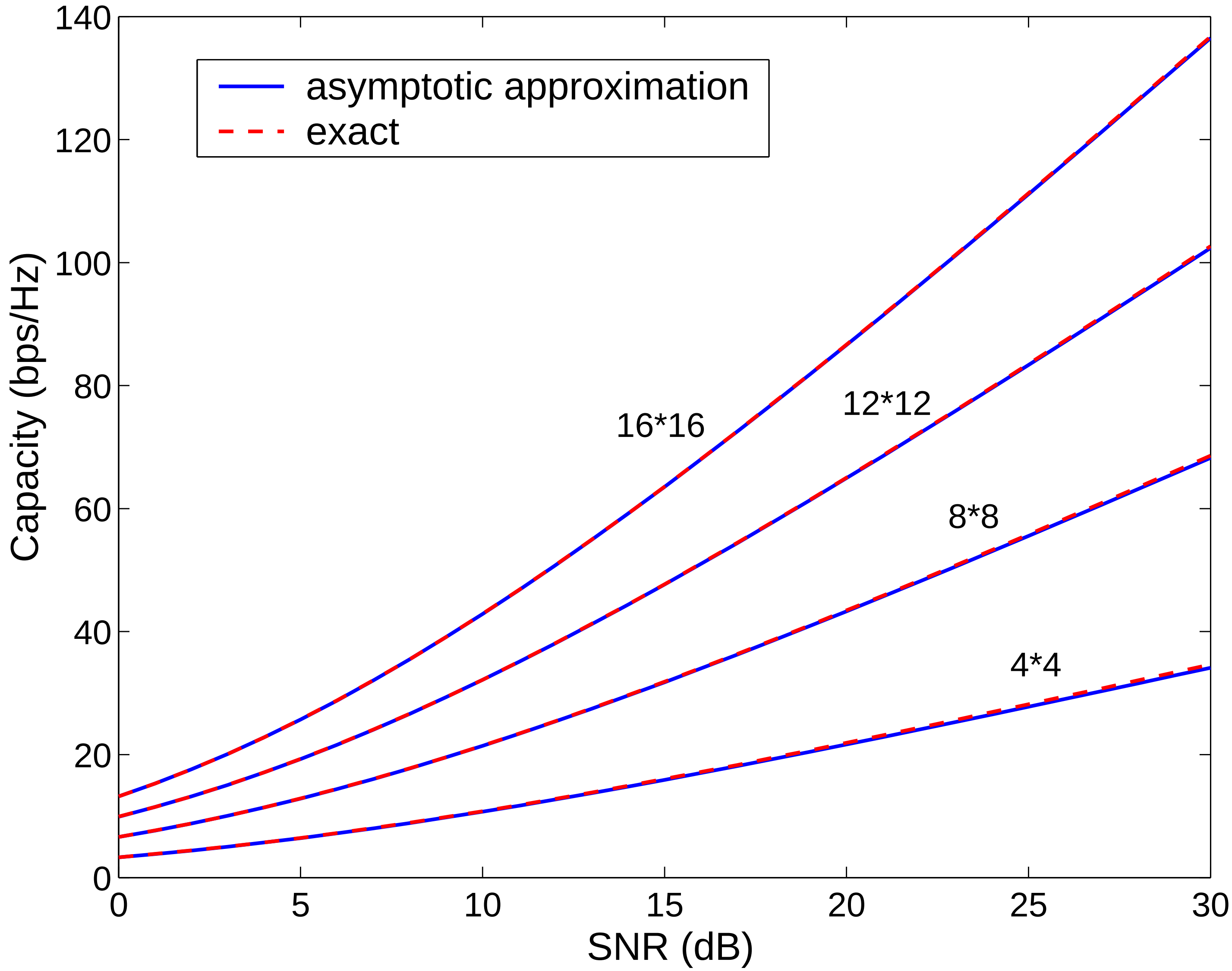
$$e(u, z) = \left[-z + \int_0^1 \frac{s(v)\Psi(u, v)dv}{1 + \int_0^1 e(w, z)s(v)\Psi(w, v)dw}\right]^{-1} \qquad (3)$$

☞ Using (1), (2), (3), we get limiting $\bar{C}$ directly in terms of Ψ

Example 1

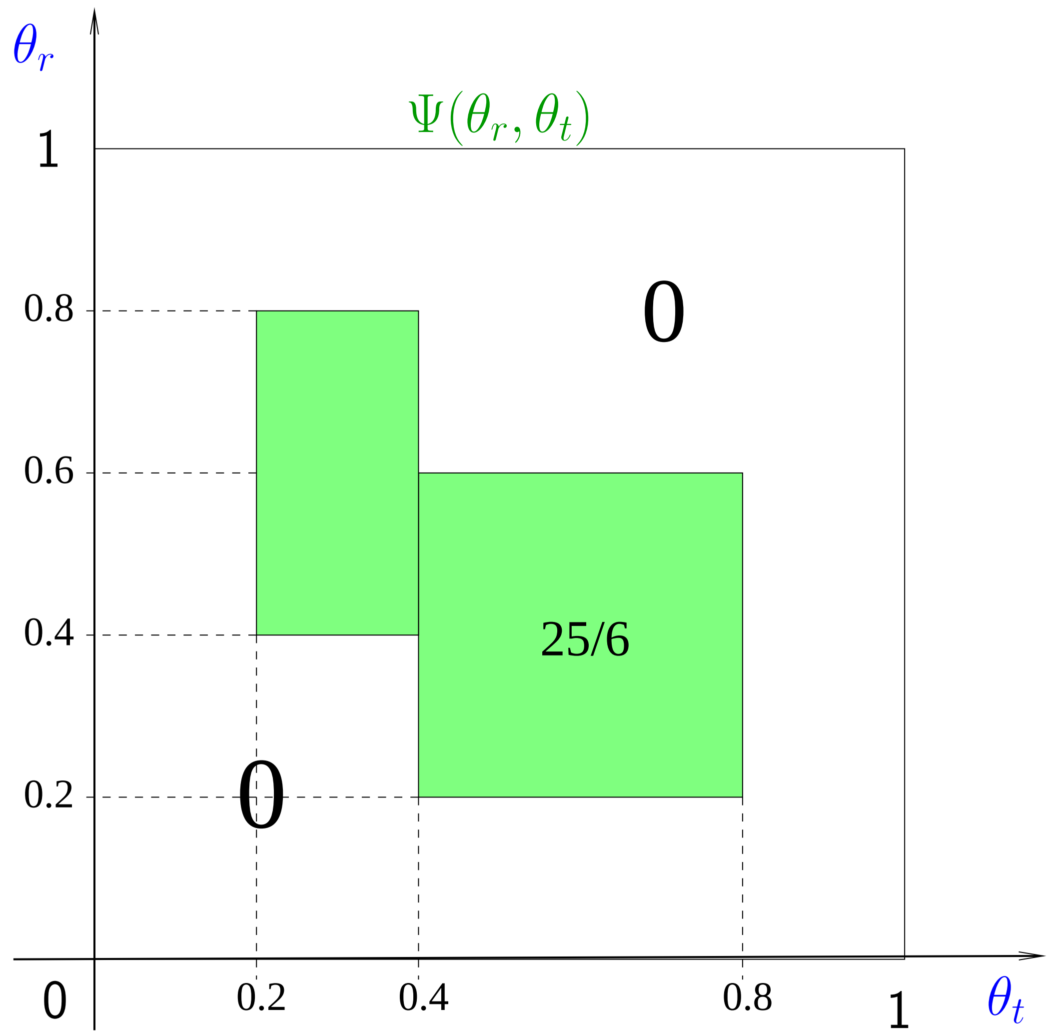


Results – 1

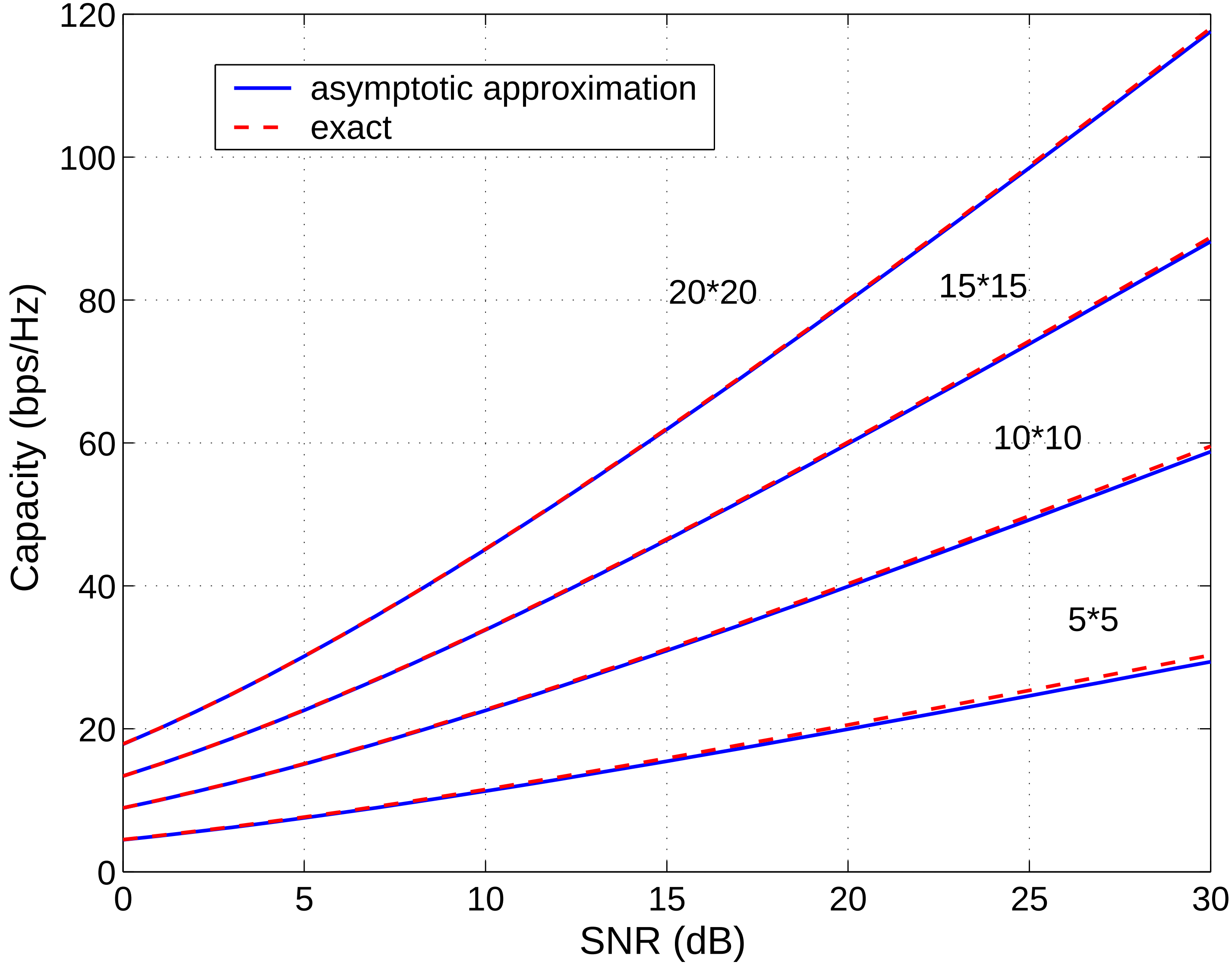


asymptotic approximation
exact
16*16
12*12
8*8
4*4
Capacity (bps/Hz)
SNR (dB)
0
20
40
60
80
100
120
140
0
5
10
15
20
25
30

Example 2



Results – 2



120
100
80
60
40
20
0
Capacity (bps/Hz)
0
5
10
15
20
25
30
SNR (dB)
asymptotic approximation
exact
20*20
15*15
10*10
5*5

Conclusions

- Exploited virtual representation of MIMO Rayleigh fading channel to analyze ergodic capacity under coherent assumption
- General model for channel statistics
- Results valid for arbitrary spatial scattering functions – no product form assumption is required
- Optimal input covariance is diagonal in virtual domain
- Beamforming along one of n_t fixed angles is optimal at low SNR
- Asymptotic normalized capacity expressed in terms of spatial scattering function
- Asymptotics accurate even for 4 or 5 antennas at Tx and Rx

Reference

V. V. Veeravalli, Y. Liang and A. M. Sayeed. "Correlated MIMO Rayleigh Fading Channels: Capacity, Optimal Signaling, and Scaling Laws." Submitted to the *IEEE Transactions on Information Theory*, September 2003.

Preprint available at http://www.comm.csl.uiuc.edu/ ∼vvv