Rigorous Analysis of Random Linear Estimation - via spatial Coupling

Nicolas Macris, Les Houches 3 March 2017

Joint with Jean Barbier, Mohamad Dia, Florent Krzakala

... coming back to our almost 10 year old work with Satish Korada - (Tight Bounds on the Capacity of Binary Input Random CDMA Systems – in Transactions Inf. Theory)

... independent analysis without spatial coupling of Henry Pfister and Gallen Reeves

PLAN

I - Gaussian random linear estimation: setting

II - Rigorous results on mutual information and MMSE (replica)

III - Optimality range of Approximate Message Passing

IV - Spatial coupling and comparison of thresholds

V - Upper bound on mutual information by interpolation method

VI - Exploiting sub-optimality of AMP: integration arguments

VII - Conclusion, remarks

I. Gaussian random linear estimation: setting

Reconstruct a signal $s_i \sim P_0(.)$, $i = 1, \dots, N$ from M linear measurements

$$y_\mu = \sum_{i=1}^{N} \phi_{\mu i} s_i + z_\mu \sqrt{\Delta}, \quad \mu \in \{1, \dots, M\}$$

Random measurement matrix $\phi_{\mu i} \sim \mathcal{N}(0, \frac{1}{N})$ and noise $z_\mu \sim \mathcal{N}(0, 1)$.

Set $\alpha = \frac{M}{N}$ for the measurement rate.

Find estimator $\widehat{\mathbf{s}}(\mathbf{y})$ with a "good" $\mathrm{MSE}[\widehat{\mathbf{s}}(\cdot)] = \mathbb{E}\|\mathbf{s} - \widehat{\mathbf{s}}(\mathbf{y})\|^2$. The optimal one is

$$\widehat{\mathbf{s}}^{\mathrm{MMSE}}(\mathbf{y}) = \mathbb{E}_{\mathbf{x}|\mathbf{y}}[\mathbf{x}|\mathbf{y}]$$

in other words

$$\underbrace{\mathbb{E}\|\mathbf{s} - \mathbb{E}_{\mathbf{x}|\mathbf{y}}[\mathbf{x}|\mathbf{y}]\|^2}_{\mathrm{MMSE}} \leq \mathrm{MSE}[\widehat{\mathbf{s}}(\cdot)]$$

From the Bayes formula $P(\mathbf{x}|\mathbf{y}) = P(\mathbf{y}|\mathbf{x})P_0(\mathbf{x})/P(\mathbf{y})$:

$$P(\mathbf{x}|\mathbf{y}) = \frac{1}{\mathcal{Z}(\mathbf{y})} \exp\Big(-\frac{1}{2\Delta} \sum_{\mu=1}^{M} ([\phi\mathbf{x}]_\mu - y_\mu)^2 \Big) \prod_{i=1}^{N} P_0(x_i)$$

where the partition function is

$$\mathcal{Z}(\mathbf{y}) = \int \exp\Big(-\frac{1}{2\Delta} \sum_{\mu=1}^{M} ([\phi\mathbf{x}]_\mu - y_\mu)^2 \Big) \prod_{i=1}^{N} P_0(x_i)\, dx_i$$

So the MMSE estimator

$$\widehat{\mathbf{s}}^{\text{MMSE}}(\mathbf{y}) = \mathbb{E}_{\mathbf{x}|\mathbf{y}}[\mathbf{x}|\mathbf{y}] = \langle\, \mathbf{x} \,\rangle$$

is a Gibbs average $\langle - \rangle$ with quenched disorder $\mathbf{y} \equiv \{\phi, \mathbf{s}, \mathbf{z}\}$.

Key objects.

Mutual information / free energy

$$i_N = \frac{1}{N}\mathbb{E}_{\boldsymbol{\Phi}}\mathbb{E}_{\mathbf{X},\mathbf{Y}}\left[\ln\left(\frac{P(\mathbf{X},\mathbf{Y})}{P_0(\mathbf{X})P(\mathbf{Y})}\right)\right] = -\frac{\alpha}{2}\underbrace{-\frac{1}{N}\mathbb{E}_{\phi,\mathbf{s},\mathbf{z}}\ln(\mathcal{Z})}_{\text{average free energy}}$$

where

$$\mathcal{Z} = \int \exp\left(-\frac{1}{2\Delta}\sum_{\mu=1}^{M}([\phi\mathbf{x}]_\mu - [\phi\mathbf{s}]_\mu - z_\mu\sqrt{\Delta})^2\right)\prod_{i=1}^{N} P_0(x_i)\,dx_i$$

MMSE (per variable)

$$\text{mmse}_N = \frac{1}{N}\mathbb{E}_{\phi,\mathbf{s},\mathbf{z}}\|\mathbf{s}-\langle\mathbf{x}\rangle\|^2$$

Measurement MMSE

$$\text{ymmse}_M = \frac{1}{M}\mathbb{E}_{\phi,\mathbf{s},\mathbf{z}}\|\phi(\mathbf{s}-\langle\mathbf{x}\rangle)\|^2 \underbrace{= \frac{2}{\alpha}\frac{di_N}{d\Delta^{-1}}}_{\text{I}-\text{MMSE formula}}$$

II. The replica formula and rigorous results

Physics calculations (Kabashima - J. Phys. A 2003, Tanaka - Tran. Inf. Th. 2004) yield:

$$\lim_{N\to+\infty} i_N = \min_E i_{\rm RS}(E;\Delta)$$

To describe $i_{\rm RS}$ introduce a scalar “denoising model”:

$$\tilde{y} = \tilde{s} + \tilde{z}\Sigma, \qquad \tilde{z} \sim \mathcal{N}(0,1)$$

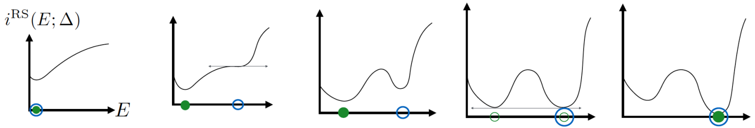
with averaged free energy or mutual information $I(\widetilde{S};\widetilde{Y})$:

$$\widetilde{f}(\Sigma) = -\mathbb{E}_{s,z} \ln \int dx\, P_0(x) \exp\left\{-\frac{(\widetilde{x} - (\widetilde{s} + \widetilde{z}\Sigma))^2}{2\Sigma^2}\right\}$$

Effective signal-to-noise ratio $\Sigma(E;\Delta)^{-2} = \alpha/(\Delta + E)$ and

$$i_{\rm RS}(E;\Delta) = \frac{1}{2}\left(\alpha \ln(1 + E/\Delta) - \frac{E}{\Sigma(E;\Delta)^2}\right) + \widetilde{f}(\Sigma(E;\Delta)) - \frac{1}{2}$$

Assumption 1: P_0 *such that* $i_{\mathrm{RS}}(E;\Delta)$ *has at most* 3 *stationary points.*



Definition: Δ_{RS} is the unique value s.t $i_{\mathrm{RS}}(E;\Delta)$ has 2 minimisers.
It is the unique non-analyticity point of $\min_E i_{\mathrm{RS}}(E;\Delta)$.

Assumption 2: P_0 *is a discrete mass distribution.*

Remark: assumptions $1+2$ can be removed by working a bit harder. Or by asking Henry Pfister and Gallen Reeves (arXiv:1607.02524 [cs.IT]).

Theorem 1: the replica calculation gives correct results.

- *For all* Δ,

$$\lim_{N\to+\infty} i_N = \min_E i_{\mathrm{RS}}(E;\Delta)$$

- *The mutual information* $\lim_{N\to+\infty} i_N$ *has a unique non-analyticity threshold equal to* Δ_{RS}.
- *For all* $\Delta \neq \Delta_{\mathrm{RS}}$

$$\lim_{N\to+\infty} \mathrm{mmse}_N = \operatorname*{argmin}_E i_{\mathrm{RS}}(E;\Delta) \equiv \widetilde{E}(\Delta)$$

and

$$\lim_{M\to+\infty} \mathrm{ymmse}_M = \frac{\widetilde{E}(\Delta)}{1+\widetilde{E}(\Delta)/\Delta}$$

Slightly more general result by independent methods:

(Pfister, Reeves - ISIT 2016 - arXiv:1607.02524 [cs.IT])

Previous results:

– Case with no phase transitions rigorosuly (Montanari, Tse - ITW 2006).

– Upper bound for binary signals (Korada, Macris - Allerton 2007 - Trans. Inf.Th 2010).

The following theorem is proved independently (and essentially before) the replica formula.

Theorem 2: relations between mmse's

For any discrete prior P_0 and almost every Δ (there might be several phase transitions here)

$$\lim_{N\to+\infty} \mathrm{ymmse}_N = \lim_{N\to+\infty} \frac{\mathrm{mmse}_N}{1+\mathrm{mmse}_N/\Delta}$$

Proof methods: based on concentration of "overlap parameters".

- Concentration of the free energy (e.g. use Ledoux-Talagrand for Gaussian variables $\phi_{\mu i}$, z_μ and McDiarmid for the signal s_i)

$$\mathbb{P}(\ln \mathcal{Z} - \mathbb{E}[\ln \mathcal{Z}] \geq \epsilon N) \leq e^{-c\epsilon^2 N}$$

- + "Nishimori" identities for Bayesian inference (with known prior).

III. Optimality range of Approximate Message Passing

Min-Sum equs on dense graph $+$ quadratic approximation $\Rightarrow$ AMP.
Algorithm updates estimates $\widehat{\mathbf{s}}^{(t)}(\mathbf{y})$, $t \geq 0$, with initialisation $\widehat{\mathbf{s}}^{(0)} = 0$.

Theorem (Bayati-Montanari 2011): *State Evolution tracks the square error* $E^{(t)} = \lim_{N\to+\infty} \frac{1}{N}\|\mathbf{s} - \widehat{\mathbf{s}}^{(t)}\|^2$. *Almost surely:*

$$E^{(t+1)} = \mathrm{mmse}(\Sigma(E^{(t)};\Delta)^{-2}), \qquad E^{(0)} = \mathbb{E}[s_i^2]$$

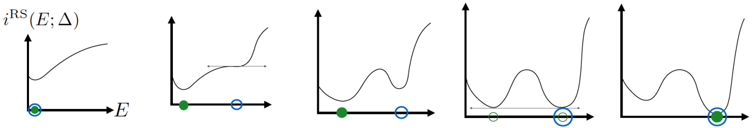
where $\mathrm{mmse}(\Sigma^{-2})$ *is associated to scalar channel* $\tilde{y} = \tilde{s} + \tilde{z}\Sigma$,

$$\mathrm{mmse}(\Sigma^{-2}) = \mathbb{E}[(\tilde{s} - \mathbb{E}[\tilde{x}|\tilde{y}])^2].$$

Recall: $\Sigma(E;\Delta)^{-2} = \alpha/(E+\Delta)$

Fixed point equation: $E = \mathrm{mmse}(\Sigma(E;\Delta)^{-2}) \iff \frac{\partial}{\partial E} i_{\mathrm{RS}}(E;\Delta) = 0.$

$\Rightarrow$ state evolution is governed by the potential function $i_{\mathrm{RS}}(E;\Delta)$:



- easy phase $\Delta < \Delta_{\mathrm{AMP}}$: $E^{(\infty)} = \widetilde{E}(\Delta)$ unique "good" min of i_{RS}.
- hard phase $\Delta \in [\Delta_{\mathrm{AMP}}, \Delta_{\mathrm{RS}}]$: $E^{(\infty)} > \widetilde{E}(\Delta)$ stuck in local min.
- bad phase $\Delta > \Delta_{\mathrm{RS}}$: $E^{(\infty)} = \widetilde{E}(\Delta)$ global "bad" min of i_{RS}.

Recall $\widetilde{E}(\Delta) = \underset{E}{\mathrm{argmin}}\, i_{\mathrm{RS}}(E;\Delta)$.

Theorem: optimality range of AMP.

For $\Delta < \Delta_{\rm AMP}$ (easy phase) and $\Delta > \Delta_{\rm RS}$ (bad phase) AMP is optimal in the sense:

$$E^{\infty} = \lim_{t\to+\infty} E^{(t)} = \lim_{N\to\infty} \mathrm{mmse}_N$$

$$\lim_{t\to+\infty} \mathrm{ymse}_{\rm AMP}^{(t)} = \lim_{M\to\infty} \mathrm{ymmse}_M$$

Recall:

$$E^{(t)} = \lim_{N\to+\infty} \frac{1}{N}\|\mathbf{s} - \widehat{\mathbf{s}}^{(t)}\|^2, \quad \mathrm{mmse}_N = \frac{1}{N}\mathbb{E}\|\mathbf{s} - \langle \mathbf{x}\rangle\|^2$$

and:

$$\mathrm{ymse}_{\rm AMP}^{(t)} = \lim_{N\to+\infty} \frac{1}{M}\|\phi(\mathbf{s} - \widehat{\mathbf{s}}^{(t)})\|^2, \quad \mathrm{ymmse}_M = \frac{1}{M}\mathbb{E}\|\phi(\mathbf{s} - \langle \mathbf{x}\rangle)\|^2$$

Proof strategy

First define the static phase transition threshold as:
$\Delta_{\rm Opt} = \sup\{\Delta \text{ s.t } \lim_{N\to+\infty} i_N \text{ is analytic in }]0,\Delta[\}$

Main steps:

1- Prove that $\Delta_{\rm RS} \leq \Delta_{\rm Opt}$. Via a detour: spatial coupling.

2- Guerra-Tonninelli interpolation type upper bound on the mutual information (S.Korada-N.M 2010 for binary signals)

$$\lim_{N\to+\infty} i_N \leq \min_E i_{\rm RS}(E;\Delta)$$

3- Sub-optimality of AMP:

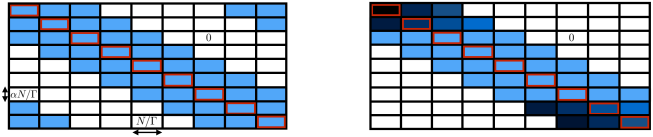
$$E^{\infty} \geq \limsup_{N\to\infty} \mathrm{mmse}_N$$

4- By integration and analyticity arguments:

$$(1) + (3) + (2) \text{ imply } \lim_{N\to+\infty} i_N = \min\ i_{\rm RS}(E;\Delta).$$

IV. Spatial coupling and the proof of $\Delta_{\mathrm{RS}} \leq \Delta_{\mathrm{Opt}}$

Random linear estimation problem with structured measurement matrix (in CS by Krzakala, Mézard, Sausset, Sun, Zdeborova).



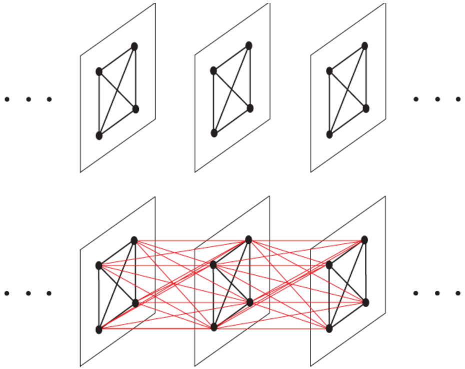
$\mathbb{R}^{M\times N}$ matrix made of $\Gamma\times\Gamma$ *blocks* (r,c) each with N/Γ columns and M/Γ rows. The i.i.d entries in block (r,c) are $\mathcal{N}(0, \frac{J_{r,c}}{N/\Gamma})$.

Periodic model (left): w forward and w backward *coupling blocks* with $J_{r,c}=\frac{1}{2w+1}$ if $|r-c|\leq w$.

Open model (right): at the boundaries the periodicity is broken. The boundaries are "seeded", e.g. the signal is known for the first few blocks.

Alternate view:

Chain of length Γ with dense graphs, each with N/Γ nodes, positioned at $c \in \{1, \ldots, \Gamma\}$ and coupled to neighboring graphs $\{c - w, \ldots, c + w\}$.



$$\sum_{\mu=1}^{M} ([\phi \mathbf{x}]_\mu - y_\mu)^2 = \sum_{|c-c'| \leq w} \sum_{i \in c, j \in c'} [\phi^T \phi]_{ij} \, x_i x_j + \ldots$$

Spatially coupled systems have two crucial properties:

- Invariance: The phase transition threshold is the same as the one of the uncoupled underlying system.
- Threshold saturation: The AMP threshold equals Δ_{RS}. With "seeded" boundary condition: no more metastable states. This is algorithmically favourable.

This is very generic: sparse graphical codes (LDPC), sparse superposition codes, compressed sensing, constraint satisfaction (to some extent), spiked Wigner model (matrix factorisation), spatially coupled Curie-Weiss models...

Theorem: Invariance of the mutual information

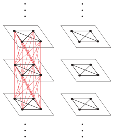
For fixed Γ *and* w*:* $\lim_{N\to\infty} i^{\rm per}_{\Gamma,w} = \lim_{N\to\infty} i_N.$

For fixed w*:* $\lim_{\Gamma\to\infty} \lim_{N\to\infty} i^{\rm open}_{\Gamma,w} = \lim_{N\to\infty} i_N.$

Corollary: $\Delta^{\rm coupled}_{\rm Opt} = \Delta_{\rm Opt}$

Proof idea: new interpolation method connecting decoupled $w = 0$, spatially coupled $0 < w < (\Gamma - 1)/2$ and fully coupled $w = (\Gamma - 1)/2$ systems:

$$\lim_{N\to\infty} i^{\rm per}_{\Gamma,(\Gamma-1)/2} \le \lim_{N\to\infty} i^{\rm per}_{\Gamma,w} \le \lim_{N\to\infty} i^{\rm per}_{\Gamma,0}$$



(methodology is related to Jean's talk)

State evolution for the spatially coupled system:

Profile of MSE for positions $r = 1, \ldots, \Gamma$:

$$E_r^{(t)} = \lim_{N\to\infty} \frac{1}{(2w+1)N/\Gamma} \sum_{c=r-w}^{r+w} \|\mathbf{s}_c - \widehat{\mathbf{s}}_c^{(t)}\|^2$$

Boundary seed: imposes $E_r^{(t)} = 0$ for all times if $r \in \mathcal{B}$.

For $r \notin \mathcal{B}$: recursive SE equation for the whole profile $(E_r^{(t)})_{r=1}^{\Gamma}$.
Initialization $E_r^{(0)} = \mathbb{E}[s^2]$.

Definition:

$$\Delta_{\rm AMP}^{\rm coupled} = \liminf_{w\to\infty} \liminf_{\Gamma\to\infty} \sup\{\Delta > 0 \mid E_r^{(\infty)} \le E_{\rm good}(\Delta)\, \forall\, r = 1, \cdots, \Gamma\}$$

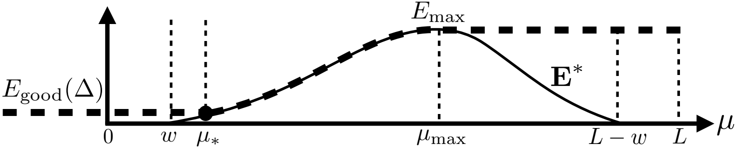
where $E_{\rm good}(\Delta)$ the smallest minimum of $i_{\rm RS}(E; \Delta)$.

Theorem - threshold saturation:

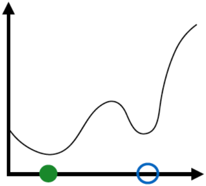
$$\Delta_{\rm AMP}^{\rm coupled} = \Delta_{\rm RS}$$

Rough idea of proof: analyse a nucleation phenomenon

- Profile is driven to flat E_{good} for all positions $r \in \{1, \cdots, \Gamma\}$ because it is energetically favourable to grow the seed.
- A reconstruction "wave" develops:



- The front of the wave is a "ball" rolling in the convex envelope of the potential as long as $\Delta < \Delta_{\mathrm{RS}}$. $\Rightarrow \Delta_{\mathrm{AMP}}^{\mathrm{coupled}} = \Delta_{\mathrm{RS}}$.



End of proof of $\Delta_{\rm RS} \leq \Delta_{\rm Opt}$:

$$\Delta_{\rm RS} \underset{(a)}{=} \Delta_{\rm AMP}^{\rm coupled} \underset{(b)}{\leq} \Delta_{\rm Opt}^{\rm coupled} \underset{(c)}{=} \Delta_{\rm Opt} \quad \left(\underset{(d)}{\leq} \Delta_{\rm RS} \right)$$

- (a) threshold saturation.
- (b) sub-optimality of AMP.
- (c) invariance of mutual information under coupling.

- (d) needs extra arguments using interpolation upper bound.

V. Upper bound by the interpolation method

Theorem: For any discrete prior P_0 we have

$$\lim_{N\to+\infty} i_N \leq \min_E i_{\rm RS}(E;\Delta)$$

Proof sketch: non-linear interpolation between the denoising model at $t=0$ and original system at $t=1$.

$$\begin{cases} \mathbf{y} = \phi\mathbf{s} + \mathbf{z}\frac{1}{\sqrt{\gamma(t)}}, & \gamma(0)=0, \qquad \gamma(1) = 1/\Delta \\ \widetilde{\mathbf{y}} = \mathbf{s} + \widetilde{\mathbf{z}}\frac{1}{\sqrt{\lambda(t)}}, & \lambda(0) = \underbrace{\Sigma(E;\Delta)^{-2}}_{\frac{\alpha}{\Delta+E}}, \qquad \lambda(1) = 0 \end{cases}$$

Total "effective SNR" should be conserved for all t:

$$\frac{\alpha}{\gamma(t)^{-1}+E} + \lambda(t) = \frac{\alpha}{\Delta+E}$$

This induces an interpolation of mutual informations for $t \in [0,1]$:

$$i_N(t=1) = i_N(t=0) + \int_0^1 dt\, \frac{di_N(t)}{dt}$$

$$i_N = i_{\rm RS}(E;\Delta) + \int_0^1 dt \Big\{ \underbrace{\frac{di_N}{dt}}_{\mathcal{A}_N + \mathcal{B}_N} - \frac{\alpha}{2}\frac{d\gamma(t)}{dt}\Big(\frac{E}{1+E\,\gamma(t)} - \frac{E}{(1+E\,\gamma(t))^2}\Big)\Big\}$$

Using Gaussian integrations by parts:

$$\mathcal{A}_N = \frac{d\gamma(t)}{dt}\frac{1}{2N}\sum_{\mu=1}^{M}\mathbb{E}[([\Phi\langle\mathbf{x}\rangle_t]_\mu - [\Phi\mathbf{s}]_\mu)^2] = \underbrace{\frac{\alpha}{2}\frac{d\gamma(t)}{dt}\mathrm{ymmse}_{N,t}}$$

(concentration of free energy and overlaps) $\approx \frac{\alpha}{2}\frac{d\gamma(t)}{dt}\frac{\mathrm{mmse}_{N,t}}{1+\mathrm{mmse}_{N,t}/\Delta}$

$$\mathcal{B}_N = \frac{d\lambda(t)}{dt}\frac{1}{2N}\sum_{i=1}^{N}\mathbb{E}[(\langle x_i\rangle_t - s_i)^2] = \underbrace{\frac{1}{2}\frac{d\lambda(t)}{dt}\mathrm{mmse}_{N,t}}$$

$$= -\frac{\alpha}{2}\frac{d\gamma(t)}{dt}\frac{\mathrm{mmse}_{N,t}}{(1+E\,\gamma(t))^2}$$

Finally:

$$i_N \approx i_{\mathrm{RS}}(E;\Delta) - \frac{\alpha}{2}\int_0^1 dt\, \underbrace{\frac{d\gamma(t)}{dt}\frac{\gamma(t)(E-\mathrm{mmse}_{N,t})^2}{(1+E\,\gamma(t))^2(1+\mathrm{mmse}_{N,t}\,\gamma(t))}}_{\text{positive}}$$

which implies

$$i_N \leq \min_E i_{\mathrm{RS}}(E;\Delta).$$

VI. Exploiting sub-optimality of AMP - integration argument and proof completion

We showed $\Delta_{\rm RS} \leq \Delta_{\rm Opt}$ (by spatial coupling)
and $i_N \leq \min_E i_{\rm RS}(E;\Delta)$ (by interpolation).

Now use sub-optimality of AMP: $E^\infty \geq \limsup \mathrm{mmse}_N \quad \Rightarrow$

$$\underbrace{\limsup_{M\to+\infty} \mathrm{ymmse}_M}_{\limsup_{N\to+\infty} \frac{di_N}{d\Delta^{-1}}} = \limsup_{N\to+\infty} \frac{\mathrm{mmse}_N}{1+\mathrm{mmse}_N/\Delta} \leq \underbrace{\frac{E^{(\infty)}}{1+E^{(\infty)}/\Delta}}_{\frac{\partial i_{\rm RS}(E^\infty;\Delta)}{\partial \Delta^{-1}}}$$

Easy phase $\Delta < \Delta_{\rm AMP}$:

$E^\infty = \mathrm{argmin}\, i_{\rm RS}(E;\Delta)$ so $\liminf_{M\to+\infty} \frac{di_N}{d\Delta} \geq \frac{d}{d\Delta} \min\ i_{\rm RS}(E;\Delta)$.

Integrating over $[0,\Delta] \subset [0,\Delta_{\rm AMP}]$: $\liminf_{N\to+\infty} i_N \geq \min_E i_{\rm RS}(E;\Delta)$ thus

$$\lim_{N\to+\infty} i_N = \min_E i_{\rm RS}(E^\infty;\Delta) \quad \text{for} \quad \Delta < \Delta_{\rm AMP}$$

Interval $\Delta_{\text{AMP}} < \Delta < \Delta_{\text{Opt}}$: We know $\underbrace{\Delta_{\text{AMP}} \le \Delta_{\text{RS}}}_{\text{hard phase}} \le \Delta_{\text{Opt}}$.

- By unicity of analytic continuation: $\lim_{N\to\infty} i_N = \min_E i_{\text{RS}}(E;\Delta)$ for $\Delta < \Delta_{\text{RS}}$ and by concavity we have continuity so equality at Δ_{RS}.

- For $\underbrace{\Delta_{\text{RS}} \le \Delta \le \Delta_{\text{Opt}}}_{\text{"non-existent interval"}}$: Repeat the integration argument.

$$\int_{\Delta_{\text{RS}}}^{\Delta} d\Delta \frac{d}{d\Delta} \min_E i_{\text{RS}}(E;\Delta) \le \liminf_{N\to+\infty} \frac{1}{N} \int_{\Delta_{\text{RS}}}^{\Delta} d\Delta \frac{di_N}{d\Delta}.$$

which leads again to

$$\min_E i_{\text{RS}}(E;\Delta) \le \liminf_{N\to+\infty} i_N$$

and to equality for $\Delta \in [0, \Delta_{\text{Opt}}]$ and also $\Delta_{\text{RS}} = \Delta_{\text{Opt}}$.

Summarizing $\lim_{N\to\infty} i_N = \min_E i_{\text{RS}}(E;\Delta)$ for $\Delta \le \Delta_{\text{Opt}} = \Delta_{\text{RS}}$.

Bad phase: $\Delta > \Delta_{\rm RS} = \Delta_{\rm Opt}$

Use again an integration of

$$\int_{\Delta_{\rm RS}}^{\Delta} d\Delta \frac{d}{d\Delta} \min_E i_{\rm RS}(E;\Delta) \leq \liminf_{N\to+\infty} \frac{1}{N} \int_{\Delta_{\rm RS}}^{\Delta} d\Delta \frac{di_N}{d\Delta}.$$

and combine with the interpolation bound.

Remark:

the integration argument does not work directly for $\Delta \in [\Delta_{\rm AMP}, \Delta_{\rm RS}]$

$$\underbrace{\frac{\partial i_{\rm RS}(E^\infty;\Delta)}{\partial \Delta^{-1}}}_{\neq \frac{d}{d\Delta^{-1}} \min_E i_{\rm RS}(E;\Delta)} \geq \limsup_{N\to+\infty} \frac{1}{N} \frac{di_N}{d\Delta^{-1}}$$

VII. Conclusion

- Similar analysis works for other Bayesian inference problems, e.g., LDPC codes, low-rank matrix factorisation, etc...
- Method is pretty generic except for the interpolation upper bound.
- New methods have appeared which don't use spatial coupling: see other talks today.