

メッセージ伝播法入門

第1回講義資料

導入

九州大学

平成30年10月10日～10月12日

豊橋技術科学大学
電気・電子情報工学系
准教授 竹内啓悟

本講義の目標

1. 初日 (第1回から第3回)

D. L. Donoho, A. Maleki, and A. Montanari, “Message-passing algorithms for compressed sensing,” *Proc. Nat. Acad. Sci.*, vol. 106, no. 45, pp. 18914–18919, Nov. 2009.

2. 二日目 (第4回から第6回)

M. Bayati and A. Montanari, “The dynamics of message passing on dense graphs, with applications to compressed sensing,” *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 764–785, Feb. 2011.

3. 三日目 (第7回と第8回)

D. L. Donoho, A. Javanmard, and A. Montanari, “Information theoretically optimal compressed sensing via spatial coupling and approximate message passing,” *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7434–7464, Nov. 2013.

本講義の心得

数学的議論は最高難度

世界トップの研究者が見ている世界は理解しがたい？

本講義資料を下に「試験」を実施したら、学生全員不合格？

木ではなく、森を見よ

数学的議論の詳細ではなく、議論のアイデアを把握せよ。

途中でついてこれなくなっても諦めるな

各回の講義は基本的に独立である。

一回分の講義を理解したら、それで論文1本書ける？

数理モデル

線形観測モデル

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}, \quad \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$$

$\mathbf{y} = (y_1, \dots, y_M)^T \in \mathbb{R}^M$: M 次元観測ベクトル

$\mathbf{x} = (x_1, \dots, x_N)^T \in \mathbb{R}^N$: N 次元信号ベクトル

$\mathbf{w} = (w_1, \dots, w_M)^T \in \mathbb{R}^M$: M 次元ノイズベクトル

$\mathbf{A} = [A_{mn}] \in \mathbb{R}^{M \times N}$: M 行 N 列の観測行列

目標

\mathbf{y} 、 \mathbf{A} 、未知の確率変数の統計性に関する情報を使って、
最小計算量かつ最良な性能で信号ベクトル \mathbf{x} を推定せよ。

本講義を通じての仮定

信号ベクトル

\mathbf{x} は平均0分散1の独立同一分布に従う(i.i.d.)成分からなる。

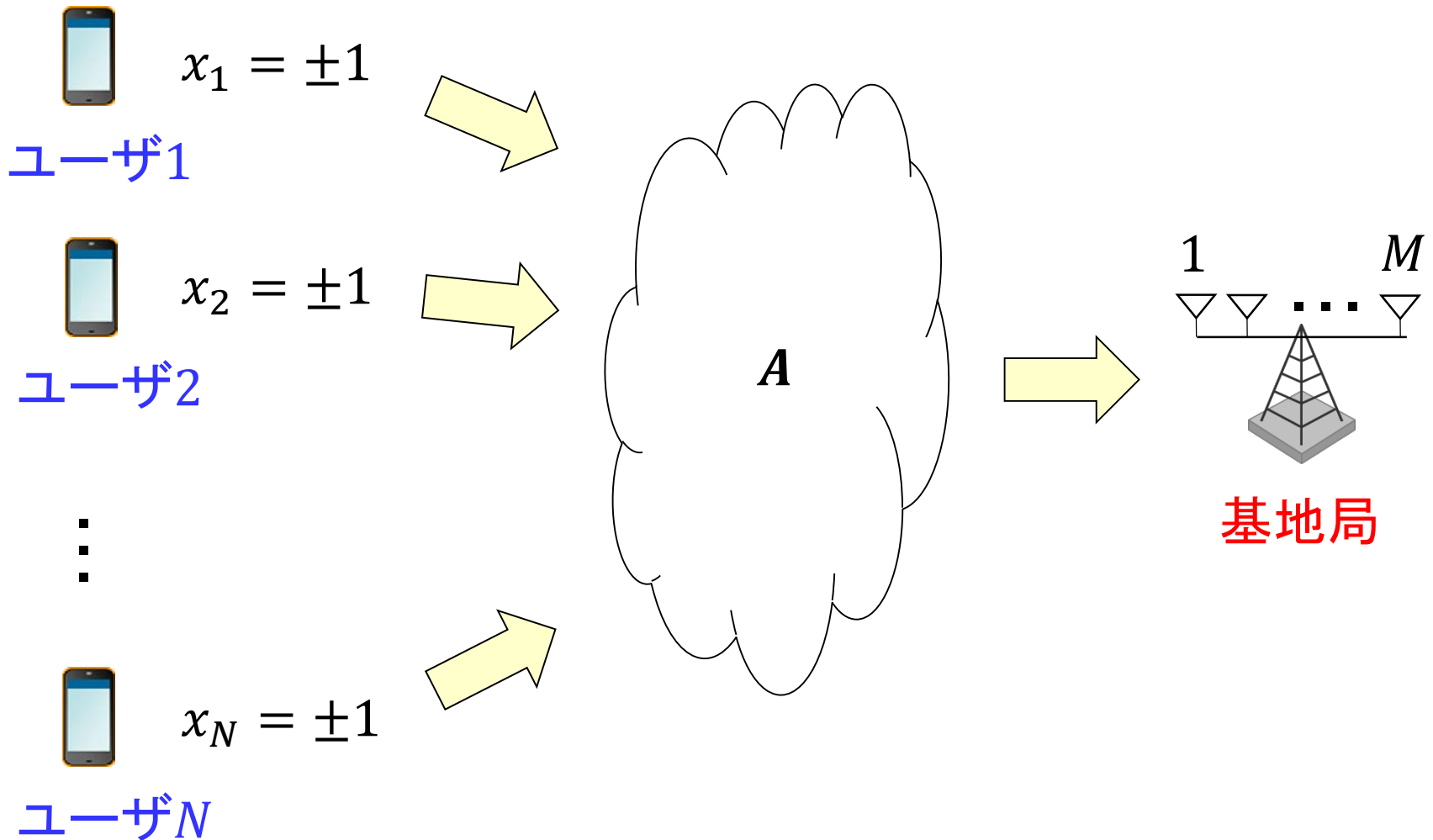
観測行列

A は平均0分散 M^{-1} のガウス分布に従う独立な成分からなる。

電力の規格化

$$\frac{1}{N} \mathbb{E}[\|\mathbf{Ax}\|^2] = \frac{1}{N} \mathbb{E}[\text{Tr}(\mathbf{AA}^T)] = 1.$$

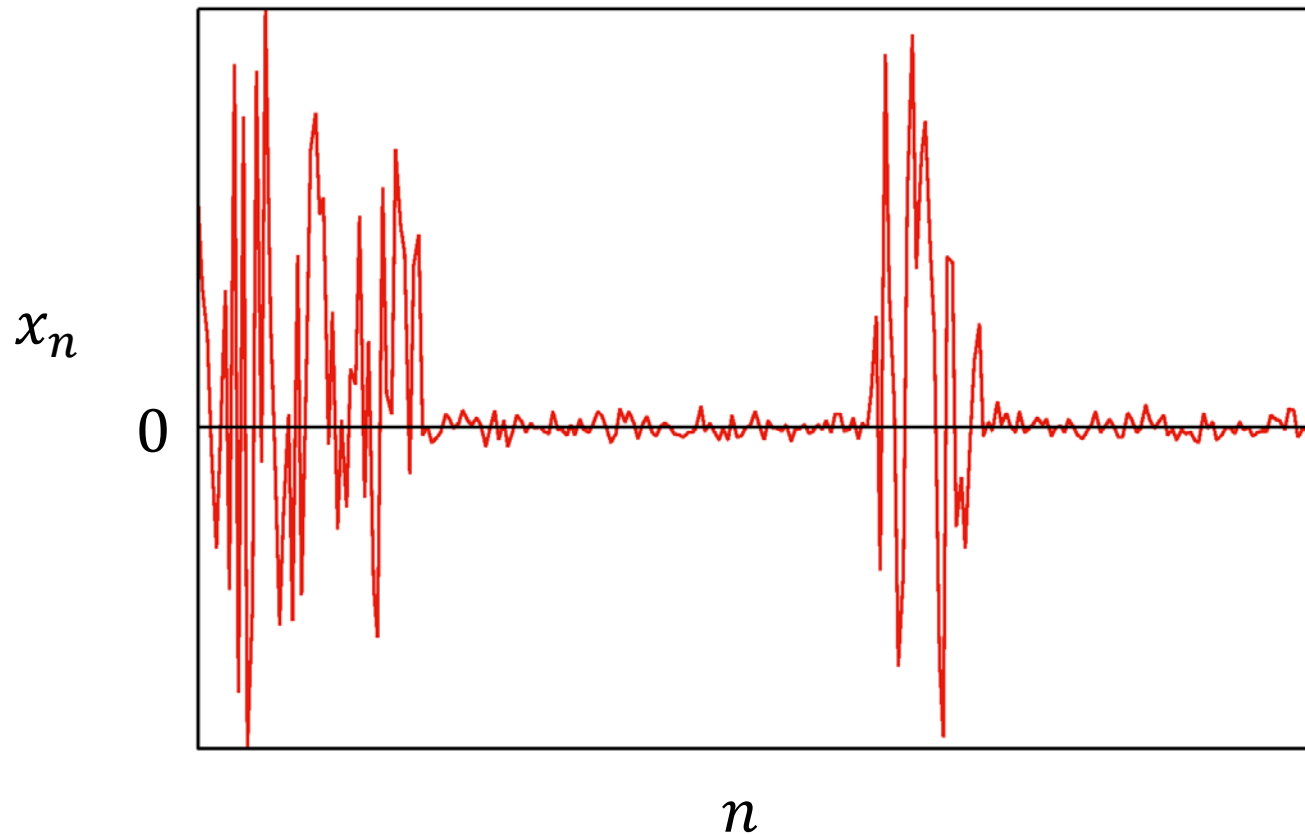
Multiple-input multiple-output (MIMO)



圧縮センシング

K -スパース信号

信号ベクトル x の非零要素数は高々 K である。



スパース重ね合わせ符号[1-1]

セクション数 L ブロックサイズ B の符号語

$$\mathbf{c} = \mathbf{A}\mathbf{x} = [\mathbf{A}[1] \cdots \mathbf{A}[L]] \begin{bmatrix} \mathbf{x}[1] \\ \vdots \\ \mathbf{x}[L] \end{bmatrix}$$

$\mathbf{x}[l] \in \mathbb{R}^B$: 非零要素が1の B 次元1-スパース情報ベクトル

符号化率

$$R = \frac{L}{M} \log B$$

[1-1] A. Joseph and A. R. Barron, "Least squares superposition codes of moderate dictionary size are reliable at rates up to capacity," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2541-2557, May 2012.

情報理論的スパース性[1-2]

ノイズなしの場合

ノイズがない場合を考え、信号成分 $\{x_n\}$ は特異成分を持たない独立同一分布(i.i.d.)に従うと仮定する。

圧縮率 $\delta = M/N$ を固定して次元 M と N を無限大にした**大システム極限**において、任意の $\epsilon \in (0,1)$ に対して信号再構成に失敗する確率が ϵ 以下となるような観測行列 A と再構成方法が存在するための**必要十分条件**は、圧縮率 δ が**レニ一情報次元** $d(x_1)$ 以上であることである。

[1-2] Y. Wu and S. Verdú, "Rényi information dimension: Fundamental limits of almost lossless analog compression," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3721-3748, Aug. 2010.

レニー情報次元[1-3]

以下、 $H(|X|) < \infty$ を満たすとする。

確率変数 X のレニー情報次元 d

レニーエントロピー

$$H(n^{-1}[nX]) = d \log n + h + o(1) \text{ as } n \rightarrow \infty.$$

離散確率変数 X の場合

$$H(n^{-1}[nX]) = 0 \log n + H(X) + o(1).$$

絶対連続な確率変数 X の場合

$$H(n^{-1}[nX]) = 1 \log n + h(X) + o(1).$$

混合分布 P に従う確率変数 X の場合

$$H(n^{-1}[nX]) = \rho \log n + O(1).$$

$$P = (1 - \rho)P_d + \rho P_c \text{ (} P_d \text{: 離散分布、} P_c \text{: 絶対連続分布)}$$

[1-3] A. Rényi, "On the dimension and entropy of probability distributions,"
Acta Mathematica Hungarica, vol. 10, no. 1-2, Mar. 1959.

詳細な解析[1-4]

ノイズがない場合を考え、信号成分 $\{x_n\}$ は ± 1 を独立に等確率で取るものと仮定する。

圧縮率 $\delta = M/N$ が以下の不等式を満たすならば、

$$\frac{M}{N} > \frac{2}{\log_3 N}$$

次元 M と N を無限大した極限において、ビット誤り確率が0に収束するような観測行列 A と信号再構成方法が存在する。

注意

A をi.i.d.ガウス行列に取れば良い。

[1-4] M. A. Sedaghat, R. R. Müller, and F. Marvasti, "On Optimum Asymptotic Multiuser Efficiency of Randomly Spread CDMA," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6635-6642, Dec. 2015.

ノイズがある場合

ノイズがある場合の性能限界の評価は困難である。

レプリカ法[1-5, 1-6]

- 統計力学で開発された計算手法である。
- 他の厳密な手法では評価が困難な問題を解ける。
- いくつかの未証明な仮定を使う。

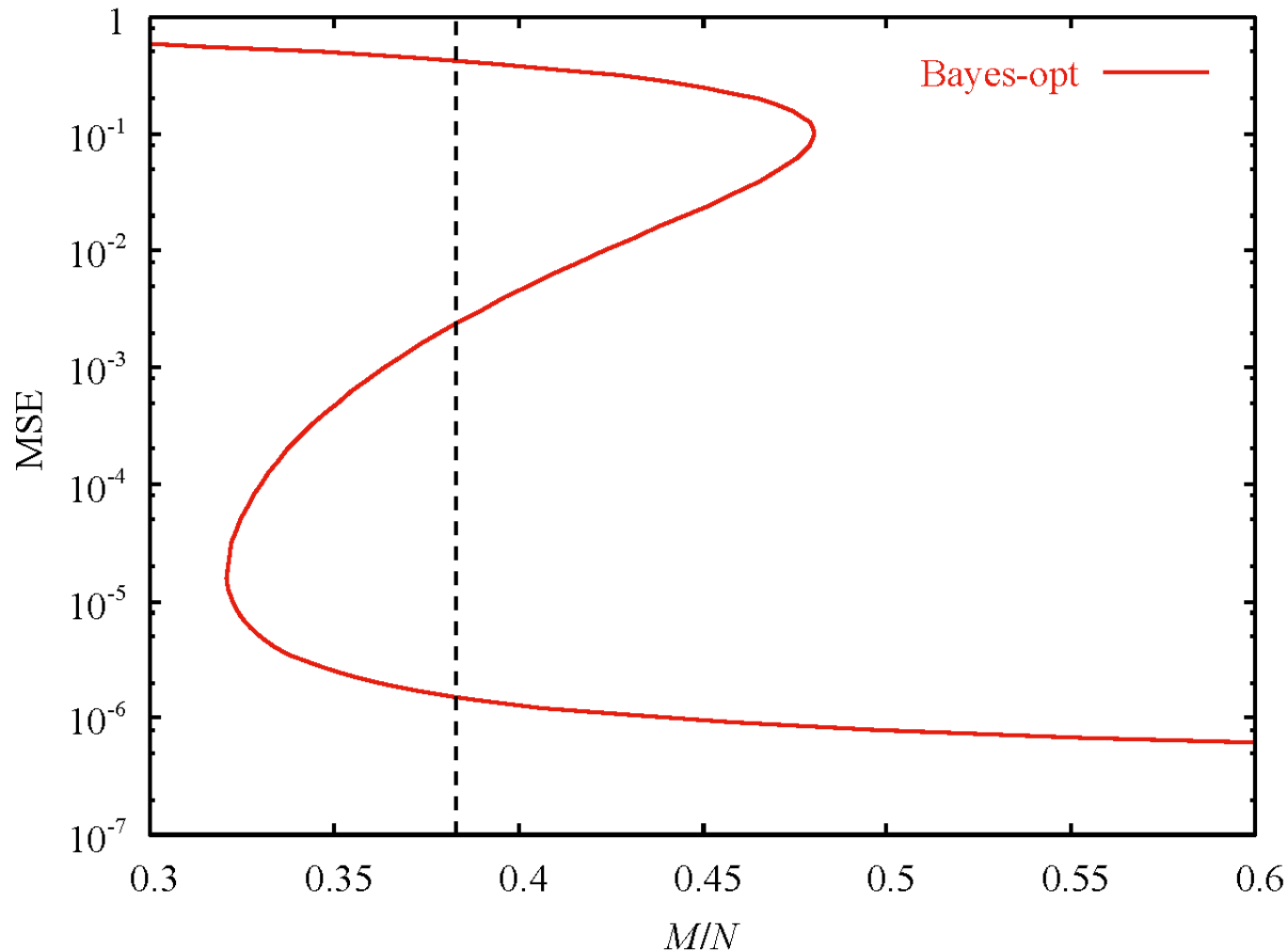
仮定の一例

モーメント系列 $\{E[X^n]\}_{n=1}^{\infty}$ から、期待値 $E[X^\alpha]$ ($\alpha < 1$)を外挿できる。

[1-5] T. Tanaka, "A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors," *IEEE Trans. Inf. Theory*, vol. 48, no. 11, pp. 2888–2910, Nov. 2002.

[1-6] D. Guo and S. Verdú, "Randomly spread CDMA: Asymptotics via statistical physics," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1983–2010, Jun. 2005.

ノイズがある場合[1-5, 1-6]



信号密度 $\rho = 0.3$ のBG事前分布、SNR: $1/\sigma^2 = 60$ dB.

厳密性に関するコメント

文献[1-7, 1-8, 1-9]の結果を組み合わせると、前ページの結果は正しいと結論づけられる。

相互情報量 $I(x; y|A)$ に関する厳密解

[1-7] G. Reeves and H. D. Pfister, “The replica-symmetric prediction for compressed sensing with Gaussian matrices is exact,” in *Proc. 2016 IEEE Int. Symp. Inf. Theory*, Barcelona, Spain, Jul. 2016, pp. 665–669.

[1-8] J. Barbier, M. Dia, N. Macris, and F. Krzakala, “The mutual information in random linear estimation,” in *Proc. 54th Annual Allerton Conf. Commun. Control & Computing*, Urbana-Champaign, IL, USA, Sep. 2016, pp. 625–632.

相互情報量とMMSEの厳密な関係式

[1-9] D. Guo, S. Shamai (Shitz), and S. Verdú, “Mutual information and minimum mean-square error in Gaussian channels,” *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261–1282, Apr. 2005.

ベイズ最適な推定

$\hat{\boldsymbol{x}} \in \mathbb{R}^N$ を信号ベクトル \boldsymbol{x} の推定量とする。

最適性の基準・・・平均二乗誤差(MSE)

$$\hat{\boldsymbol{x}}_{\text{opt}} = \underset{\hat{\boldsymbol{x}} \in \mathbb{R}^N}{\text{argmin}} \mathbb{E}[\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2 | A]$$

事後平均推定量 $\hat{\boldsymbol{x}}_{\text{PME}} = \mathbb{E}[\boldsymbol{x} | \boldsymbol{y}, A]$ は最適である。

証明

$$\begin{aligned} \mathbb{E}[\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2] &= \mathbb{E}_{\boldsymbol{y}}[\mathbb{E}_{\boldsymbol{x}}[\|\boldsymbol{x} - \hat{\boldsymbol{x}}_{\text{PME}} + \hat{\boldsymbol{x}}_{\text{PME}} - \hat{\boldsymbol{x}}\|^2 | \boldsymbol{y}]] \\ &= \mathbb{E}[\|\boldsymbol{x} - \hat{\boldsymbol{x}}_{\text{PME}}\|^2] + \mathbb{E}[\|\hat{\boldsymbol{x}}_{\text{PME}} - \hat{\boldsymbol{x}}\|^2] \geq \mathbb{E}[\|\boldsymbol{x} - \hat{\boldsymbol{x}}_{\text{PME}}\|^2] \end{aligned}$$

A に関する条件付けを省略した。 ■

事後平均推定量

事後平均推定量の定義

$$\hat{x}_{\text{PME}} = \int x p(x|y, A) dx$$

ベイズの公式

$$p(x|y, A) = \frac{p(y|A, x)p(x)}{p(y|A)}, \quad p(y|A) = \int p(y|A, x)p(x) dx$$

事後平均推定量

$$\hat{x}_{\text{PME}} = \frac{\int x p(y|A, x)p(x) dx}{\int p(y|A, x)p(x) dx}$$

事後平均推定量の問題点

計算量

x が連続の場合・・・ N 重積分が必要

$x \in \{\pm 1\}^N$ の場合・・・少なくとも 2^N 回の足し算が必要

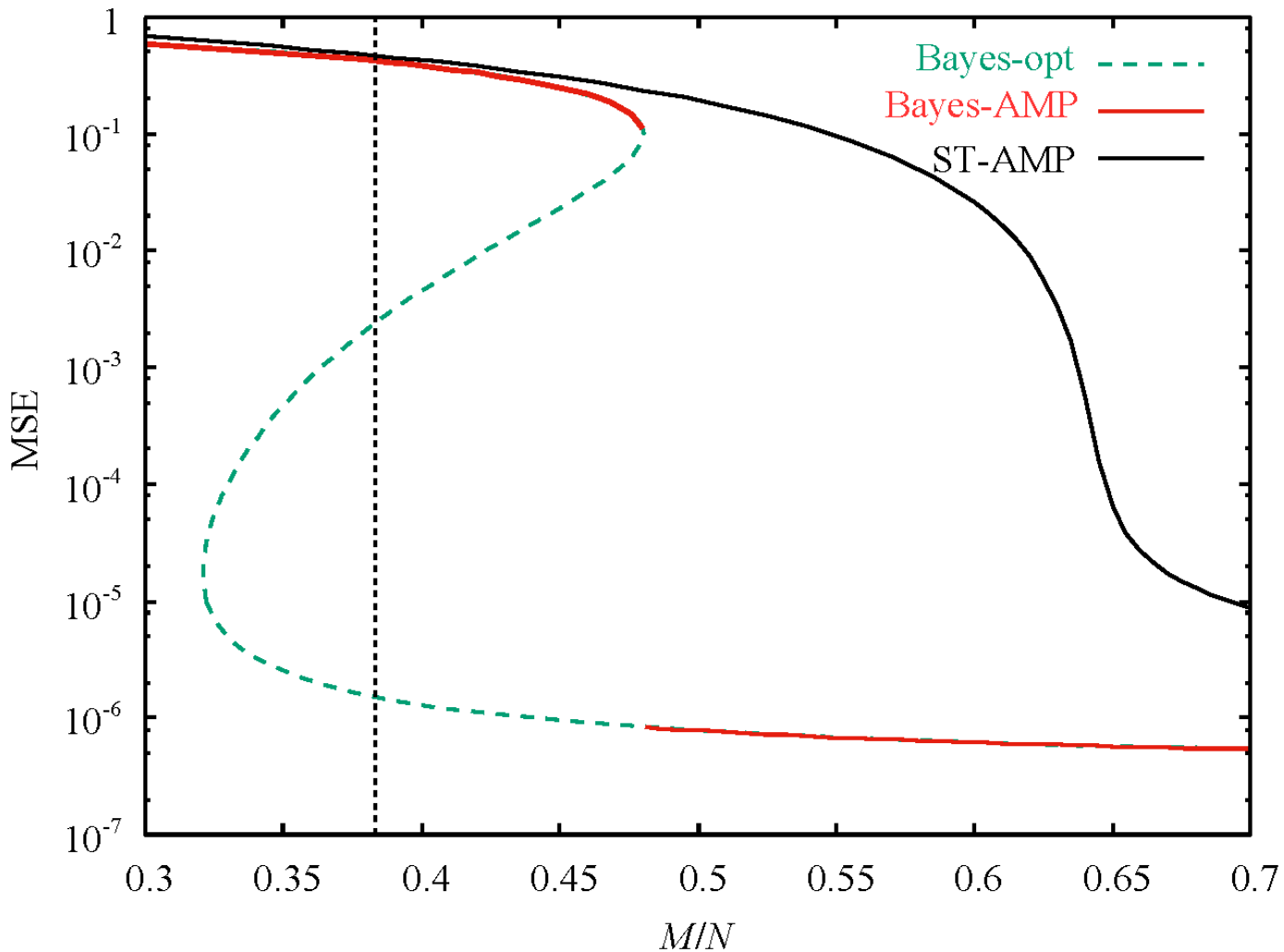
最小計算量の定義

$O(IMN)$ の計算量で評価できる x の推定量 \hat{x} は、
最小計算量で評価可能と言う。

I : M や N に直接依存しないパラメータ(反復回数)

本講義で扱うAMPは、最小計算量である。

数値的比較



信号密度 $\rho = 0.3$ のBG事前分布、SNR: $1/\sigma^2 = 60$ dB.