メッセージ伝播法入門
後半第7回講義資料
メッセージ伝播法の発展的内容

広島市立大学
令和４年９月２６日～９月２８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

AMPの利点と欠点

AMPの利点

- 計算量は、干渉除去に整合フィルタを使うため、最小限である。

反復当たりの計算量: $\mathcal{O}(MN)$

- 観測行列の成分が独立で$\mathcal{N}(0,1/M)$に従う場合、

状態発展方程式の不動点が唯一ならば、AMPは大システム極限でベイズ最適な性能を達成する。[3−4, 1−7, 1−8]

AMPの欠点

悪条件[7−1]や非ゼロ平均[7−2]など、観測行列が非i.i.d.行列の場合、AMPは収束しない。

[7-1] S. Rangan, P. Schniter, A. Fletcher, and S. Sarkar, “On the convergence of approximate message passing with arbitrary matrices,” *IEEE Trans. Inf. Theory*, vol. 65, no. 9, pp. 5339–5351, Sep. 2019.

[7-2] F. Caltagirone, L. Zdeborová, and F. Krzakala, “On convergence of approximate message passing,” in *Proc. 2014 IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jul. 2014, pp. 1812–1816.

Orthogonal/Vector AMP[7-3, 7-4]

モジュールA

線形フィルタ

$$\boldsymbol{x}_{\mathrm{A},t} = \boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t} + \boldsymbol{W}_t^{\mathrm{T}}\left(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t}\right),$$

$$\xi_{\mathrm{A},t} = \frac{1}{N}\mathrm{Tr}\left(\boldsymbol{I} - \boldsymbol{W}_t^{\mathrm{T}}\boldsymbol{A}\right)$$

オンサーガ補正

$$\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t} = \frac{\boldsymbol{x}_{\mathrm{A},t} - \xi_{\mathrm{A},t}\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t}}{1 - \xi_{\mathrm{A},t}},$$

$$v_{\mathrm{A}\to\mathrm{B},t} = \frac{\xi_{\mathrm{A},t} v_{\mathrm{B}\to\mathrm{A},t}}{1 - \xi_{\mathrm{A},t}}$$

モジュールB

Denoiser

$$\boldsymbol{x}_{\mathrm{B},t+1} = \eta_t(\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t}),$$

$$\xi_{\mathrm{B},t} = \langle \eta'_t(\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t}) \rangle$$

オンサーガ補正

$$\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t+1} = \frac{\boldsymbol{x}_{\mathrm{B},t+1} - \xi_{\mathrm{B},t}\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t}}{1 - \xi_{\mathrm{B},t}},$$

$$v_{\mathrm{B}\to\mathrm{A},t+1} = \frac{\xi_{\mathrm{B},t} v_{\mathrm{A}\to\mathrm{B},t}}{1 - \xi_{\mathrm{B},t}}$$

[7-3] J. Ma and L. Ping, “Orthogonal AMP,” *IEEE Access*, vol. 5, pp. 2020–2033, Jan. 2017.

[7-4] S. Rangan, P. Schniter, and A. K. Fletcher, “Vector approximate message passing,” *IEEE Trans. Inf. Theory*, vol. 65, no. 10, pp. 6664–6684, Oct. 2019.

OAMP/VAMP

ベイズ最適なOAMP/VAMP

以下の線形MMSEフィルタとベイズ最適なdenoiserを使用

$$\boldsymbol{W}_t = v_{\mathrm{B}\to\mathrm{A},t}\left(\sigma^2\boldsymbol{I} + v_{\mathrm{B}\to\mathrm{A},t}\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\right)^{-1}\boldsymbol{A}$$

関連アルゴリズム

- ベイズ最適なOAMP/VAMPは、期待値伝播法[7–5]の大システム近似[7–6, 7–7]と等価である。
- OAMP/VAMPの特別な場合は、[7–8]で最初に提案された。

[7-5] T. P. Minka, “Expectation propagation for approximate Bayesian inference,” in *Proc. 17th Conf. Uncertainty Artif. Intell.*, Seattle, WA, USA, Aug. 2001, pp. 362–369.

[7-6] J. Céspedes, P. M. Olmos, M. Sánchez-Fernández, and F. Perez-Cruz, “Expectation propagation detection for high-order high-dimensional MIMO systems,” *IEEE Trans. Commun.*, vol. 62, no. 8, pp.2840–2849, Aug. 2014.

[7-7] K. Takeuchi, “Rigorous dynamics of expectation-propagation-based signal recovery from unitarily invariant measurements,” *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 368–386, Jan. 2020.

[7-8] M. Opper and O. Winther, “Expectation consistent approximate inference,” *J. Mach. Learn. Res.*, vol. 6, pp. 2177–2204, Dec. 2005.

行列の直交不変性

$\boldsymbol{X}$を任意のランダム行列とする。

左直交不変性

$\boldsymbol{\Phi X}$が定義可能で、$\boldsymbol{X}$と独立なすべての直交行列$\boldsymbol{\Phi}$に対して、以下の条件が成立するとき、$\boldsymbol{X}$は左直交不変であると言う。

$$\boldsymbol{X} \sim \boldsymbol{\Phi X}.$$（両辺の分布が同一）

右直交不変性

$\boldsymbol{X\Psi}$が定義可能で、$\boldsymbol{X}$と独立なすべての直交行列$\boldsymbol{\Psi}$に対して、以下の条件が成立するとき、$\boldsymbol{X}$は右直交不変であると言う。

$$\boldsymbol{X} \sim \boldsymbol{X\Psi}.$$

状態発展法による解析

状態発展方程式[7-4, 7-7]

モジュールA

$$\bar{\xi}_{\mathrm{A},t} = \lim_{M=\delta N\to\infty} \frac{1}{N}\mathrm{Tr}\left(\boldsymbol{I} - \boldsymbol{W}_t^{\mathrm{T}}\boldsymbol{A}\right), \qquad \bar{v}_{\mathrm{A}\to\mathrm{B},t} = \frac{\bar{\xi}_{\mathrm{A},t}\bar{v}_{\mathrm{B}\to\mathrm{A},t}}{1-\bar{\xi}_{\mathrm{A},t}}$$

モジュールB

$$\bar{\xi}_{\mathrm{B},t} = \mathbb{E}[\eta'_t(x_1+\omega_1)],$$

$$\omega_1 \sim \mathcal{N}\left(0, \bar{v}_{\mathrm{A}\to\mathrm{B},t}\right).$$

$$\bar{v}_{\mathrm{B}\to\mathrm{A},t+1} = \frac{\bar{\xi}_{\mathrm{B},t}\bar{v}_{\mathrm{A}\to\mathrm{B},t}}{1-\bar{\xi}_{\mathrm{B},t}}$$

右直交不変な観測行列の場合

大システム極限で、

$$\frac{1}{N}\left\|\boldsymbol{x} - \boldsymbol{x}_{\mathrm{B},t+1}\right\|^2 \to \mathbb{E}[\{x_1 - \eta_t(x_1+\omega_1)\}^2] \text{ almost surely.}$$

OAMP/VAMPの利点と欠点

OAMP/VAMPの利点

- 観測行列が右直交不変の場合、

状態発展方程式の不動点が唯一ならば、OAMP/VAMPは大システム極限で**ベイズ最適な性能**を達成する。[7-4, 7-7, 7-9, 7-10, 7-11]

OAMP/VAMPの欠点

- 線形MMSEフィルタの実装に、**$\boldsymbol{A}$の特異値分解**が必要である。

  特異値分解の計算量 : $\mathcal{O}(M^2N + MN^2 + N^3)$

- 信号事前分布として**ガウス分布を仮定できない**。

[7-9] K. Takeda, S. Uda, and Y. Kabashima, “Analysis of CDMA systems that are characterized by eigenvalue spectrum,” *Europhys. Lett.*, vol. 76, no. 6, pp. 1193–1199, 2006.

[7-10] A. M. Tulino, G. Caire, S. Verdú, and S. Shamai (Shitz), “Support recovery with sparsely sampled free random matrices,” *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4243–4271, Jul. 2013.

[7-11] J. Barbier, N. Macris, A. Maillard, and F. Krzakala, “The mutual information in random linear estimation beyond i.i.d. matrices,” in *Proc. 2018 IEEE Int. Symp. Inf. Theory*, Vail, CO, USA, Jun. 2018, pp. 1390–1394.

長期記憶メッセージ伝播法

一般形　$\boldsymbol{x}_{\mathrm{A},t} = \phi_t\left(\boldsymbol{x}_{\mathrm{B},0}, \dots, \boldsymbol{x}_{\mathrm{B},t}\right), \qquad \boldsymbol{x}_{\mathrm{B},t+1} = \psi_t(\boldsymbol{x}_{\mathrm{A},0}, \dots, \boldsymbol{x}_{\mathrm{A},t})$

メッセージを更新する際、過去のメッセージすべてを使う。

ϕ_tやψ_tの設計目標

最小限の計算量で、ベイズ最適な性能を達成する。

設計手順

1. 関数ϕ_tやψ_tの適切なモデルを考える。
2. 状態発展法の一般論[7-12]を利用して、誤差が漸近的にガウス分布するように、モデルパラメータの一部を設計する。
3. 状態発展法により、残りのパラメータを最適化する。

[7-12] K. Takeuchi, “Bayes-optimal convolutional AMP,” *IEEE Trans. Inf. Theory*, vol. 67, no. 7, pp. 4405–4428, Jul. 2021.

Convolutional AMP

Convolutional AMP (CAMP) [7-12, 7-13]

$$\boldsymbol{x}_{t+1} = \eta_t\left(\boldsymbol{x}_t + \boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}_t\right), \qquad \xi_t = \left\langle \eta_t'\left(\boldsymbol{x}_t + \boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}_t\right)\right\rangle$$

$$\boldsymbol{z}_t = \boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_t + \sum_{\tau=0}^{t-1}\left(\prod_{t'=\tau}^{t-1}\xi_{t'}\right)\left(\theta_{t-\tau}\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} - g_{t-\tau}\boldsymbol{I}_M\right)\boldsymbol{z}_\tau,$$

AMPのオンサーガ補正を過去の全メッセージの畳み込みに拡張

CAMPの利点

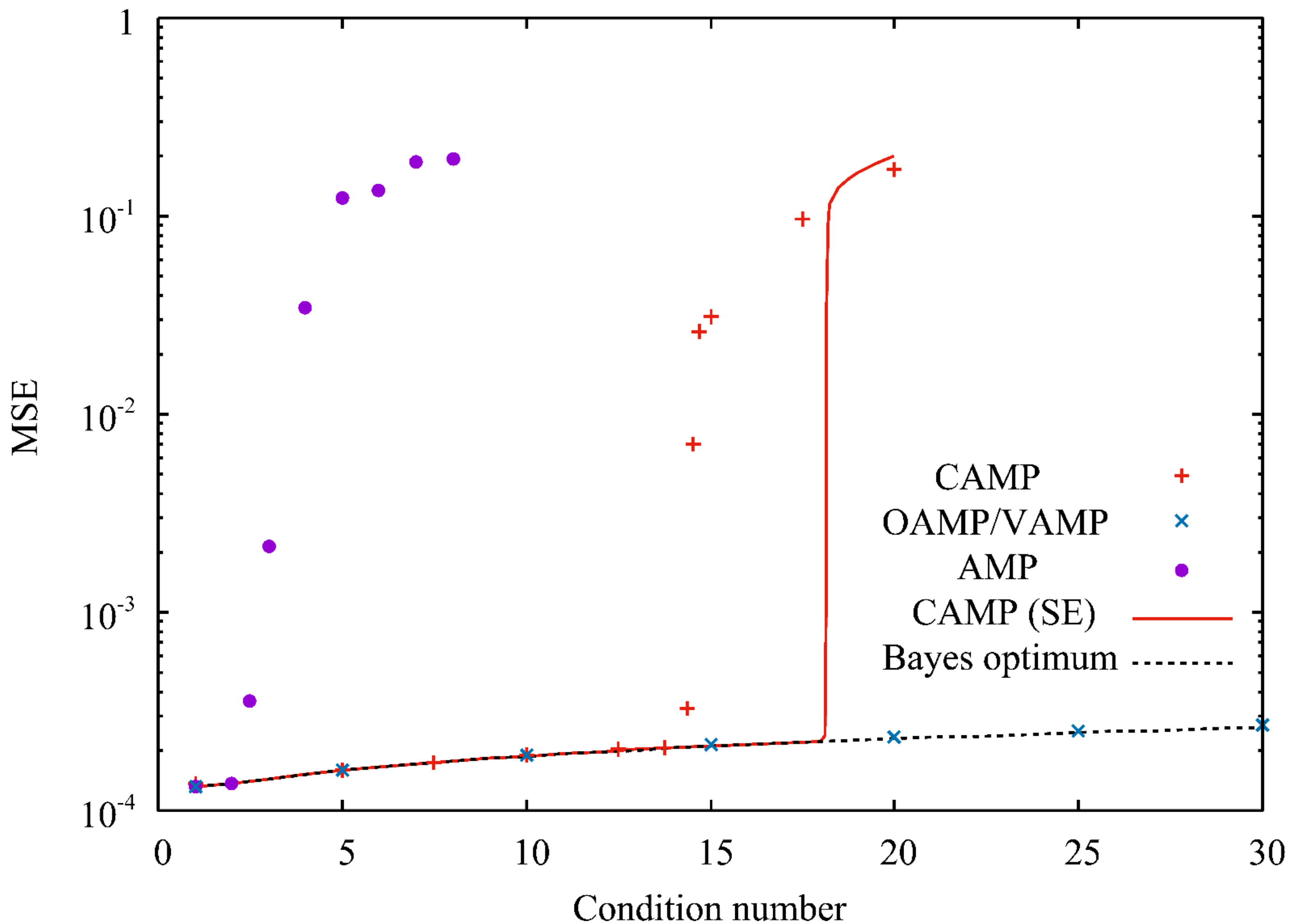
CAMPの計算量は、最小限である。

注意

$\boldsymbol{A}$の成分が独立で$\mathcal{N}(0,1/M)$に従うとき、CAMPはAMPと等価

[7-13] K. Takeuchi, "Convolutional approximate message-passing," *IEEE Signal Process. Lett.*, vol. 27, pp. 416–420, 2020.

数値実験（文献[7-12]のFig. 4）



$M = 2^{10}$、$N = 2^{11}$、信号密度$\rho = 0.1$のBG事前分布、$1/\sigma^2 = 30$ dB

Memory AMP [7-14]

モジュールA

$$\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t} = \frac{1}{\epsilon_t}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}_t - \sum_{\tau=0}^{t}\xi_{\mathrm{A},t,\tau}\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},\tau}\right),$$

$$\boldsymbol{z}_t = a_t\left(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t}\right) + b_t\left(\bar{\lambda} - \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\right)\boldsymbol{z}_{t-1},$$

$$\bar{\lambda} = \frac{\lambda_{\max}\left(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\right) + \lambda_{\min}\left(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\right)}{2}$$

モジュールB

$$\boldsymbol{x}_{\mathrm{B},t+1} = \eta_t\left(\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t}\right),$$

$$\xi_{\mathrm{B},t} = \left\langle \eta'_t\left(\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t}\right)\right\rangle,$$

$$\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t+1} = \sum_{\tau=0}^{t}\theta_{t,\tau}\frac{\boldsymbol{x}_{\mathrm{B},\tau+1} - \xi_{\mathrm{B},\tau}\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},\tau}}{1-\xi_{\mathrm{B},\tau}}$$

Memory AMPの計算量は、最小限である。

[7-14] L. Liu, S. Huang and B. M. Kurkoski, "Memory AMP," *IEEE Trans. Inf. Theory*, 2022, DOI: 10.1109/TIT.2022.3186166.

Memory AMPの特徴

モジュールA　以下の恒等式を使って、LMMSEフィルタを近似

$$(\boldsymbol{I}-\boldsymbol{G}_t)^{-1}=\sum_{\tau=0}^{\infty}\boldsymbol{G}_t^{\tau}, \qquad \boldsymbol{G}_t=\boldsymbol{I}-b_t\left(\frac{\sigma^2}{v_{\mathrm{B}\to\mathrm{A},t}}\boldsymbol{I}+\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\right).$$

$\{\xi_{\mathrm{A},t,\tau}\}$は誤差の漸近ガウス性が成立するように設計

モジュールB　長期記憶ダンピング[7−15]により、収束を保証[7−14]

Memory AMPの利点

- 観測行列が右直交不変の場合、

  状態発展方程式の不動点が唯一ならば、memory AMPは大システム極限でベイズ最適な性能を達成する。[7−14]

[7-15] K. Takeuchi, "On the convergence of orthogonal/vector AMP: Long-memory message-passing strategy," *IEEE Trans. Inf. Theory*, 2022, DOI: 10.1109/TIT.2022.3194855.

WS-CG-VAMP

Warm-Starting Conjugate Gradient VAMP (WS-CG-VAMP) [7-16]

モジュールA

$$x_{\mathrm{A}\to\mathrm{B},t} = \frac{1}{\epsilon_t}\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{z}_t^i - \sum_{\tau=0}^{t}\xi_{\mathrm{A},t,\tau}\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},\tau}\right),$$

$\boldsymbol{z}_t^i$は、線形方程式に対する解$\boldsymbol{z}_t$をi回のCG反復[7-17]で近似した解

$$\left((\sigma^2/v_{\mathrm{B}\to\mathrm{A},t})\boldsymbol{I} + \boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\right)\boldsymbol{z}_t = \boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t}$$

モジュールB

$$\boldsymbol{x}_{\mathrm{B},t+1} = \eta_t(\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t}), \qquad \boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t+1} = \frac{\boldsymbol{x}_{\mathrm{B},t+1} - \xi_{\mathrm{B},t}\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t}}{1-\xi_{\mathrm{B},t}},$$

$$\xi_{\mathrm{B},t} = \langle {\eta'}_t(\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t})\rangle \qquad v_{\mathrm{B}\to\mathrm{A},t+1} = \frac{\xi_{\mathrm{B},t}v_{\mathrm{A}\to\mathrm{B},t}}{1-\xi_{\mathrm{B},t}}$$

[7-16] N. Skuratovs and M. E. Davies, “Compressed sensing with upscaled vector approximate message passing,” *IEEE Trans. Inf. Theory*, vol. 68, no. 7, pp. 4818–4836, Jul. 2022.

[7-17] K. Takeuchi and C.-K. Wen, “Rigorous dynamics of expectation-propagation signal detection via the conjugate gradient method,” in *Proc. 18th IEEE Int. Workshop Sig. Process. Advances Wirel. Commun.*, Sapporo, Japan, Jul. 2017, pp. 88–92.

WS-CG-VAMPの特徴

WS−CG−VAMPの利点

- WS−CG−VAMPの計算量は、最小限である。
- 観測行列が右直交不変の場合、

  状態発展方程式の不動点が唯一ならば、WS−CG−VAMPは大システム極限でベイズ最適な性能を達成する。[7−16]

Memory AMPとWS−CG−VAMPの比較

Memory AMPは、CG反復1回のWS−CG−VAMPと同等性能

未解決課題

- Memory AMPとWS−CG−VAMP($i > 1$)の公平な比較
- システムが大きすぎる($N = 2^{14}$ [7−14]、$N = 2^{20}$ [7−16])

TOYOHASHI UNIVERSITY OF TECHNOLOGY

Rotationally Invariant AMP

Rotationally Invariant AMP (RI-AMP) [7-18]

モジュールA

$$\boldsymbol{x}_{\mathrm{A},t} = \phi_t\left(\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},1}, \dots, \boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t}, \boldsymbol{y}\right),$$

$$\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}_{\mathrm{A},t} - \sum_{\tau=1}^{t} \theta_{\mathrm{A},t,\tau}\boldsymbol{x}_{\mathrm{B},\tau}$$

モジュールB

$$\boldsymbol{x}_{\mathrm{B},t+1} = \psi_t\left(\boldsymbol{x}_{\mathrm{A}\to\mathrm{B},0}, \dots, \boldsymbol{x}_{\mathrm{A}\to\mathrm{B},t}\right),$$

$$\boldsymbol{x}_{\mathrm{B}\to\mathrm{A},t+1} = \boldsymbol{A}\boldsymbol{x}_{\mathrm{B},t+1} - \sum_{\tau=0}^{t} \theta_{\mathrm{B},t,\tau}\boldsymbol{x}_{\mathrm{A},\tau}$$

RI-AMPの計算量は、最小限である。

[7-18] Z. Fan, "Approximate message passing algorithms for rotationally invariant matrices," *Ann. Statist.*, vol. 50, no. 1, pp. 197–224, Feb. 2022.

RI-AMPの特徴

利点

- 状態発展法[7-18]により、誤差の漸近ガウス性が成立するように、$\{\theta_{\mathrm{A},t,\tau}, \theta_{\mathrm{B},t,\tau}\}$を設計できる。
- 状態発展法[7-18]は、Dynamical Functional Theory [7-19]を応用・厳密化した結果である。

未解決問題

- 状態発展法を用いてベイズ最適なϕ_tやψ_tを定義できるが、最適化問題が複雑すぎて、解が解けない。

[7-19] M. Opper, B. Çakmak, and O. Winther, “A theory of solving TAP equations for Ising models with general invariant random matrices,” *J. Phys. A: Math. Theor.*, vol. 49, no. 11, p. 114002, Feb. 2016.

まとめ

アルゴリズム	計算量	信号事前分布	観測行列	収束性	ベイズ最適
AMP [3-1]	最小限	任意	i.i.d.ガウス	証明済	証明済
OAMP/VAMP [7-3, 7-4]	高	非ガウス	右直交不変	証明済 [7-15]	証明済
CAMP [7-12]	最小限	任意	右直交不変	低条件数のみ	証明済（収束すれば）
Memory AMP [7-14]	最小限	非ガウス	右直交不変	証明済	証明済
WS-CG-VAMP [7-16]	最小限	非ガウス	右直交不変	証明済	証明済
RI-AMP [7-18]	最小限	任意	右直交不変	？	？

赤字は欠点/未解決問題を表す。

現状で、赤字の項目がないアルゴリズムは存在しない。