A Unified Algorithmic Framework for Dynamic Compressive Sensing

Xiaozhi Liu
Joint work with Prof. Yong Xia

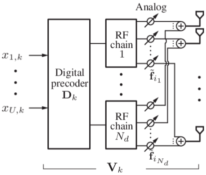
School of Mathematical Sciences, Beihang University
xzliu@buaa.edu.cn

The 21st Annual Conference of CSIAM, Kunming
October 14, 2023

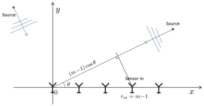
Presentation Overview

Outline

- A narrowband massive MIMO system with HBF structure



(a) HBF structure[1]



(b) ULA structure[2]

- The observed precoded channel $h_t \in \mathbb{C}^{N_{\mathrm{RF}}}$ at time slot t

$$h_t = W_t H_t + n_t, n_t \sim \mathcal{CN}(0, \sigma_m^2 \mathbf{I}_{N_{\mathrm{RF}}}), \tag{1}$$

[1] Song Noh, Michael D Zoltowski, and David J Love. “Training sequence design for feedback assisted hybrid beamforming in massive MIMO systems”. In: *IEEE Transactions on Communications* 64.1 (2015), pp. 187–200.

[2] Zai Yang et al. “Sparse methods for direction-of-arrival estimation”. In: *Academic Press Library in Signal Processing, Volume 7*. Elsevier, 2018, pp. 509–581.

Dynamic channel tracking

- A half-wave spaced ULA at the receiver

$$H_t = \sum_{i=1}^{N_L} \alpha_i \mathrm{a}(\theta_i^t), \quad (2)$$

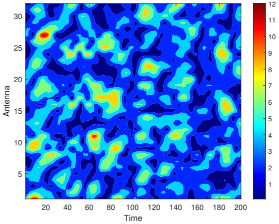
where α_i denoted the complex gain of the ith path and $\mathrm{a}(\theta_i^t) = \frac{1}{\sqrt{N_r}}[1 \;\; e^{j\pi \sin \theta_i^t} \;\; \ldots \;\; e^{j\pi (N_r - 1) \sin \theta_i^t}]^T$

- H_t can be transformed into the sparse angle domain

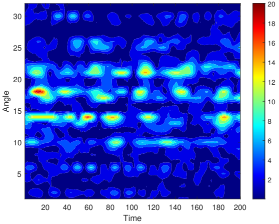
$$H_t = \mathrm{D}\tilde{H}_t, \quad (3)$$

where D is the transform dictionary determined by the geometrical structure of the antenna array.

Temporal structured sparsity



(c)



(d)

Figure: (a) Antenna-Time domain. (b) Angle-Time domain.

- we can rewrite (1) as

$$h_t = A_t \tilde{H}_t + n_t, \tag{4}$$

Outline

Dynamic Compressive sensing (DCS)

- The main goal of DCS problem is to recursively reconstruct $\{x_t\}$ from $\{y_t\}$(i.e., $M \ll N$)

$$y_t = A_t x_t + n_t, n_t \sim \mathcal{CN}(0, R_t), \tag{5}$$

- Focusing on the dynamic filtering model

$$x_t = f_t(x_{t-1}) + \nu_t, \tag{6}$$

where ν_t is the filtering noise (innovation).

Existing DCS algorithms

- Recursive algorithms
  1. Exploiting the slow support changing feature
     - LS-CS[3], Modified-CS[4], Weighted-ℓ_1[5].
  2. Exploiting the slow support and value changing feature.
     - KF-CS[6], RegModCS[7]
- Question: *Do these existing recursive DCS algorithms have some intrinsic correlations or not?*

---

[3] Namrata Vaswani. "LS-CS-residual (LS-CS): Compressive sensing on least squares residual". In: *IEEE Transactions on Signal Processing* 58.8 (2010), pp. 4108–4120.

[4] Namrata Vaswani and Wei Lu. "Modified-CS: Modifying compressive sensing for problems with partially known support". In: *IEEE Transactions on Signal Processing* 58.9 (2010), pp. 4595–4607.

[5] Adam S Charles and Christopher J Rozell. "Dynamic filtering of sparse signals using reweighted ℓ_1". In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE. 2013, pp. 6451–6455.

[6] Namrata Vaswani. "KF-CS: Compressive sensing on Kalman filtered residual". In: *arXiv preprint arXiv:0912.1628* (2009).

[7] Wei Lu and Namrata Vaswani. "Regularized modified BPDN for noisy sparse reconstruction with partial erroneous support and signal value knowledge". In: *IEEE Transactions on Signal Processing* 60.1 (2011), pp. 182–196.

Outline

Kalman Filter[8]

- Predict:

$$\hat{x}_{t|t-1} = F_t \hat{x}_{t-1}, \tag{7}$$
$$P_{t|t-1} = F_t P_{t-1} F_t^H + Q_t, \tag{8}$$

- Update:

$$K = P_{t|t-1} A_t^H (A_t P_{t|t-1} A_t^H + R_t)^{-1}, \tag{9}$$
$$\hat{x}_t = \hat{x}_{t|t-1} + K(y_t - y_{t|t-1}), y_{t|t-1} = A_t \hat{x}_{t|t-1}, \tag{10}$$
$$P_t = (I - K A_t) P_{t|t-1}, \tag{11}$$

[8] Rudolph Emil Kalman. “A new approach to linear filtering and prediction problems”. In: (1960).

Equivalent MAP estimate

- MMSE = MAP

$$\hat{x}_t = \arg\min_x \{\|y_t - A_t x\|^2_{R_t^{-1}} + \|x - \hat{x}_{t|t-1}\|^2_{P_{t|t-1}^{-1}}\}, \qquad (12)$$

where the matrix weighted norm is defined as $\|z\|_R^2 = z^H R z$

Partial-Laplacian filtering sparsity model

- To capture the dynamic sparsity outside the support, we propose the **Partial-Laplacian filtering sparsity model** sparsity model

$$
\begin{aligned}
(x_t)_{T_{t-1}} &= (F_t x_{t-1})_{T_{t-1}} + (\nu_t)_{T_{t-1}}, \\
(x_t)_{T_{t-1}^c} &= (F_t x_{t-1})_{T_{t-1}^c} + (\nu_t)_{T_{t-1}^c},
\end{aligned}
\quad (13)
$$

where T_t denotes the the support set of x_t. We assume that $(\nu_t)_{T_{t-1}} \sim \mathcal{CN}(0, Q_t^1)$ and $(\nu_t)_{T_{t-1}^c}$ have independent Laplacian but non-identical distributions with inverse scale w_i, i.e. $p((\nu_t)_i) = \frac{w_i}{2} e^{-w_i |(\nu_t)_i|}, i \in T_{t-1}^c$.

Equivalent MAP estimate

- Based on the dynamic model (13), the MAP estimate is

$$\hat{x}_t = \arg\min_x \{ \|y_t - A_t x\|^2_{R_t^{-1}} + \gamma \|(x)_T - (\hat{x}_{t|t-1})_T\|^2_{(P_{t|t-1})_1^{-1}} + \|W_t((x)_{T^c} - (\hat{x}_{t|t-1})_{T^c})\|_1 \}, \quad (14)$$

where $(P_{t|t-1})_1 = \gamma(P_{t|t-1})_{T,T}$ and $W_t = diag(w_{i_1}, w_{i_2}, ..., w_{i_{N-L}})$.

Algorithmic framework

Algorithm Partial-Laplacian Dynamic CS (PLAY-CS)

Input: $\{y_1, y_2, ..., y_T\}$, $A_t, \forall t$, σ_m^2, σ_m^2, α, a, b, $F_t, \forall t$, $W_t, \forall t$

Initialize: $Q_t = \sigma_m^2 I, R_t = \sigma_f^2 I, \forall t$, $P_0 = I, \hat{x}_0 = 0, T_0 = \emptyset$

for all $t = 1, 2, ..., T$ **do**

Prediction

$\hat{x}_{t|t-1} = F_t \hat{x}_{t-1},$

$P_{t|t-1} = F_t P_{t-1} F_t^H + Q_t,$

$T = T_{t-1}.$

Update

$K = P_{t|t-1} A_t^H (A_t P_{t|t-1} A_t^H + R_t)^{-1}.$

Estimate $\hat{x}_t$ using (14).

$P_t = (I - KA_t) P_{t|t-1}.$

Support estimation: $T_t = \{i : |(\hat{x}_t)_i| > \alpha\}$.

end for

Output: $\{\hat{x}_1, \hat{x}_2, ..., \hat{x}_T\}$

Outline

Unified DCS algorithmic framework

Table: Summary of the connection between PLAY-CS and the existing DCS algorithms

Algorithm	Connections with PLAY-CS
KF-CS[9]	make $(x_t)_T$ and $(x_t)_{T^c}$ independent
Modified-CS[10]	$W_t = I, \hat{x}_{t\|t-1} = 0, \gamma = 0$
RegModCS[11]	$W_t = I, (\hat{x}_{t\|t-1})_{T^c} = 0$
Weighted-ℓ_1[12]	$\hat{x}_{t\|t-1} = 0, T = \emptyset$

[9] Namrata Vaswani. "KF-CS: Compressive sensing on Kalman filtered residual". In: *arXiv preprint arXiv:0912.1628* (2009).

[10] Namrata Vaswani and Wei Lu. "Modified-CS: Modifying compressive sensing for problems with partially known support". In: *IEEE Transactions on Signal Processing* 58.9 (2010), pp. 4595–4607.

[11] Wei Lu and Namrata Vaswani. "Regularized modified BPDN for noisy sparse reconstruction with partial erroneous support and signal value knowledge". In: *IEEE Transactions on Signal Processing* 60.1 (2011), pp. 182–196.

[12] Adam S Charles and Christopher J Rozell. "Dynamic filtering of sparse signals using reweighted ℓ_1". In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE. 2013, pp. 6451–6455.

- KF-CS[13]
  1. Running a reduced KF on $(x_t)_T$.
  2. Estimating the additions on $(x_t)_{T^c}$ through the Dantzig selector.

Connections

KF-CS can be viewed as a special case of the PLAY-CS when (ν_t) can be divided into two independent parts: $(\nu_t)_T$ and $(\nu_t)_{T^c}$ and the W_t is I.

[13] Namrata Vaswani. "KF-CS: Compressive sensing on Kalman filtered residual". In: *arXiv preprint arXiv:0912.1628* (2009).

Connections with Modified-CS

- Modified-CS[14]

$$\hat{x}_t = \arg\min_{x} \|y_t - A_t x\|_2^2 + \|(x)_{T^c}\|_1. \quad (15)$$

Connections

When $\gamma = 0, \hat{x}_{t|t-1} = 0$ and $W_t = I$ in (14), the Modified-CS can be viewed as a special case of the PLAY-CS.

[14] Namrata Vaswani and Wei Lu. "Modified-CS: Modifying compressive sensing for problems with partially known support". In: *IEEE Transactions on Signal Processing* 58.9 (2010), pp. 4595–4607.

Connections with RegModCS

- RegModCS[15]

$$\hat{x}_t = \arg\min_x \|y_t - A_t x\|_2^2 + \gamma \|(x)_T - (\hat{x}_{t|t-1})_T\|_2^2 + \|(x)_{T^c}\|_1, \quad (16)$$

Connections

RegModCS can be derived based on the PLAY-CS when we set $W_t = I$ and $(\hat{x}_{t|t-1})_{T^c} = 0$ in (14).

[15] Wei Lu and Namrata Vaswani. "Regularized modified BPDN for noisy sparse reconstruction with partial erroneous support and signal value knowledge". In: *IEEE Transactions on Signal Processing* 60.1 (2011), pp. 182–196.

Connections with Weighted-ℓ_1

- RWL1-DF[16]

$$\hat{x}_t = \arg\min_x \|y_t - A_t x\|_2^2 + \|W_t x\|_1, \tag{17}$$

Connections

Weighted-ℓ_1 can be considered a simplified variant of the PLAY-CS when we set $T = \emptyset$ and $\hat{x}_{t|t-1} = 0$.

[16] Adam S Charles and Christopher J Rozell. "Dynamic filtering of sparse signals using reweighted ℓ_1". In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE. 2013, pp. 6451–6455.

Outline

Partial-LSM sparsity model

- Motivation: Optimal W_t is difficult to determine.
- To address the above issues, the **Partial Laplacian scale mixture** (**Partial-LSM**) filtering sparsity model shown as Fig. 2 is proposed in this study.

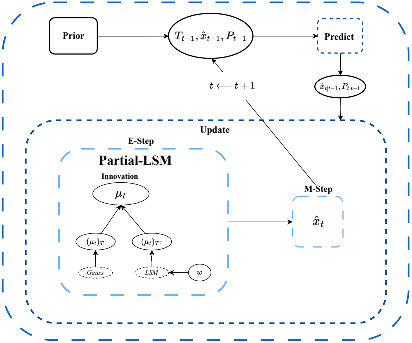


Figure: The hierarchical structure of the Partial-LSM model

Partial-LSM sparsity model

- In model (13), we futher model the inverse scale parameter w_i of $(\nu_t)_i, i \in T_{t-1}^c$ as independent Gamma distributions

$$\begin{aligned}(x_t)_{T_{t-1}} &= (F_t x_{t-1})_{T_{t-1}} + (\nu_t)_{T_{t-1}},\\(x_t)_{T_{t-1}^c} &= (F_t x_{t-1})_{T_{t-1}^c} + (\nu_t)_{T_{t-1}^c},\end{aligned} \tag{18}$$

where we assume

- $(\nu_t)_{T_{t-1}} \sim \mathcal{CN}(0, Q_t^1)$.
- $p((\nu_t)_i) = \frac{w_i}{2} e^{-w_i|(\nu_t)_i|}, i \in T_{t-1}^c$ and

$$p(w_i) = \frac{b^a}{\Gamma(a)} w_i^{a-1} e^{-bw_i}, i \in T_{t-1}^c. \tag{19}$$

Partial-LSM sparsity model

- Using Bayes's rule, the MAP estimate is given by

$$\begin{aligned}\hat{x}_t &= \arg\min_x \{-\log p(x \mid y_t)\} \\ &= \arg\min_x \{-\log p(y_t \mid x) - \log p(x)\} \\ &= \arg\min_x \{-\log p(y_t \mid x) - \log p((x)_T) - \log p((x)_{T^c})\}\end{aligned} \tag{20}$$

- In general, we do not necessarily have an analytical expression for the $\log p((x)_{T^c})$. The typical approach when dealing with such a problem is the EM algorithm.

EM algorithm

- Using Jensen's inequality, we obtain the following upper bound

$$
\begin{aligned}
-\log p(x \mid y_t) \leq & -\log p(y_t \mid x) - \log p((x)_T) \\
& - \int_w q(w) \log \frac{p((x)_{T^c}, w)}{q(w)} dw := \mathcal{L}(q, x)
\end{aligned} \tag{21}
$$

- Based on EM algorithm, we can perform coordinate descent in $\mathcal{L}(q, x)$

$$
\text{E Step} \quad q^{(k+1)} = \arg\min_q \mathcal{L}\left(q, x^{(k)}\right) \tag{22}
$$

$$
\text{M Step} \quad x^{(k+1)} = \arg\min_x \mathcal{L}\left(q^{(k+1)}, x\right) \tag{23}
$$

EM algorithm

- Let $\langle . \rangle_q$ denote the expectation with respect to $q(w)$. The M Step (23) simplifies to

$$\hat{x}_t = \arg\min_{x} \{ \|y_t - A_t x\|^2_{R_t^{-1}} + \gamma \|(x)_T - (\hat{x}_{t|t-1})_T\|^2_{(P_{t|t-1})_1^{-1}} + \|W_t((x)_{T^c} - (\hat{x}_{t|t-1})_{T^c})\|_1 \}, \quad (24)$$

where $(W_t)^k = diag(\langle w_{i_1} \rangle_{q^k}, \langle w_{i_2} \rangle_{q^k}, ..., \langle w_{i_{N-L}} \rangle_{q^k})$.

- Let $\langle . \rangle_q$ denote the expectation with respect to $q(w)$. The M Step (23) simplifies to

$$\hat{x}_t = \arg\min_x \{ \|y_t - A_t x\|^2_{R_t^{-1}} + \gamma \|(x)_T - (\hat{x}_{t|t-1})_T\|^2_{(P_{t|t-1})_1^{-1}} + \|W_t((x)_{T^c} - (\hat{x}_{t|t-1})_{T^c})\|_1 \}, \quad (24)$$

where $(W_t)^k = diag(\langle w_{i_1} \rangle_{q^k}, \langle w_{i_2} \rangle_{q^k}, ..., \langle w_{i_{N-L}} \rangle_{q^k})$.

- We have tight equality in the (21) if $q(w) = p(w \mid x)$, which implies that the E step (22) reduces to

$$q^{(k+1)}(w) = p(w \mid x^k). \quad (25)$$

- Note that in the M step we only need to compute

$$\langle w_i \rangle_{p(w|x^k)} = \frac{a+1}{b+|(x^k)_i|}, \tag{26}$$

which is based on the assumption of the Partial-LSM model.

Algorithmic framework

Algorithm PLAY-CS with LSM (PLAY$^+$-CS)

Input: $\{y_1, y_2, ..., y_T\}$, $A_t, \forall t$, σ_m^2, σ_m^2, α, a, b, $F_t, \forall t$

Initialize: $Q_t = \sigma_m^2 I$, $R_t = \sigma_f^2 I, \forall t$, $P_0 = I$, $\hat{x}_0 = 0$, $T_0 = \emptyset$

for all $t = 1, 2, ..., T$ **do**

Prediction

$$\hat{x}_{t|t-1} = F_t \hat{x}_{t-1},$$

$$P_{t|t-1} = F_t P_{t-1} F_t^H + Q_t,$$

$$T = T_{t-1}.$$

Update

$$K = P_{t|t-1} A_t^H (A_t P_{t|t-1} A_t^H + R_t)^{-1}.$$

E-Step

Set diagonal matrix W_t using (26)

M-Step

Estimate $\hat{x}_t$ using (14).

$$P_t = (I - KA_t) P_{t|t-1}.$$

Support estimation: $T_t = \{i : |(\hat{x}_t)_i| > \alpha\}$.

end for

Output: $\{\hat{x}_1, \hat{x}_2, ..., \hat{x}_T\}$

Outline

Experimental setup

- Datasets: CDL-B channel[17]
  1. ULA, $N_r = 32$.
  2. $N_L = 23$, $T = 200$.
- Comparison Methods: Regular-CS[18], KF-CS, Modified-CS, RegModCS, Weighted-ℓ_1
- Evaluation Measures:
  - **NMSE**

$$\text{NMSE} := \frac{\|\hat{x}_t - x_t\|^2}{\|x_t\|^2}, \tag{27}$$

  - **Corr**

$$\text{Corr} := \frac{\hat{x}_t^H x_t}{\|x_t\| \, \|\hat{x}_t\|}, \tag{28}$$

  - **TNMSE/TCorr**

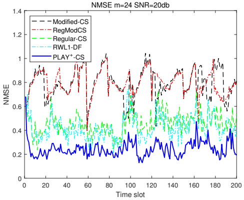
$$\text{TNMSE} := \frac{1}{T}\sum_{t=1}^{T} \frac{\|\hat{x}_t - x_t\|^2}{\|x_t\|^2}, \quad \text{TCorr} := \frac{1}{T}\sum_{t=1}^{T} \frac{\hat{x}_t^H x_t}{\|x_t\| \, \|\hat{x}_t\|}. \tag{29}$$

[17] *Study on Channel Model for Frequencies From 0.5 to 100 GHz*. document TR 38.901. V 15.0.0. 3GPP, June 2018.

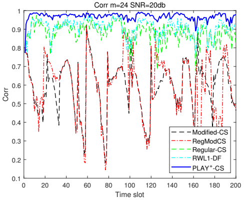
[18] Scott Shaobing Chen, David L Donoho, and Michael A Saunders. "Atomic decomposition by basis pursuit". In: *SIAM review* 43.1 (2001), pp. 129–159.

Performance Comparison

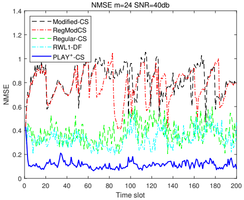
- The NMSE curves and Corr curves of different methods when $m = 24$. (a), (c) TNMSE curves. (b), (d) TCorr curves.



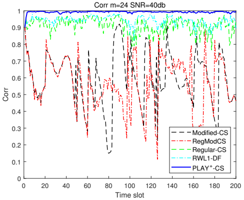
(a) SNR=20db



(b) SNR=20db

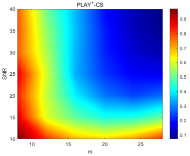


(c) SNR=40db

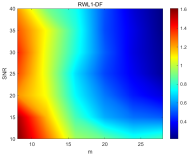


(d) SNR=40db

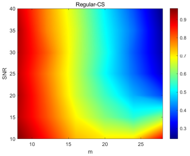
- The TNMSE and TCorr performance of various algorithms under diffirent SNR and CR levels. (i), (j), (k) The TNMSE performance. (l), (m), (n) The TCorr performance.



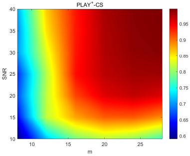
(e) **PLAY^{+}-CS**



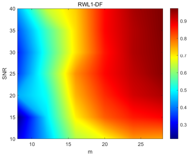
(f) RWL1-DF



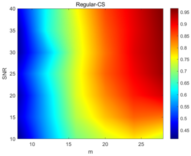
(g) Regular-CS



(h) **PLAY^{+}-CS**



(i) RWL1-DF



(j) Regular-CS

Impact of SNR

Table: The TNMSE of different methods when $m = 24$

	SNR=15	SNR=20	SNR=25	SNR=30	SNR=35	SNR=40
Modified-CS	0.5379	0.4562	0.4100	0.3937	0.4100	0.3877
RegModCS	0.8152	0.7954	0.8032	0.8250	0.7795	0.8255
Regular-CS	0.7977	0.7977	0.7724	0.8260	0.7870	0.7759
RWL1-DF	0.5414	0.4186	0.3678	0.3456	0.3536	0.3307
PLAY^{+}-CS	**0.4366**	**0.2432**	**0.1790**	**0.1438**	**0.1259**	**0.1150**

Table: The TCorr of different methods when $m = 24$

	SNR=15	SNR=20	SNR=25	SNR=30	SNR=35	SNR=40
Modified-CS	0.8426	0.8830	0.9019	0.9119	0.9080	0.9124
RegModCS	0.5816	0.5908	0.6048	0.5421	0.6046	0.5953
Regular-CS	0.5831	0.5953	0.6053	0.5432	0.6244	0.5621
RWL1-DF	0.8595	0.9093	0.9253	0.9358	0.9344	0.9419
PLAY^{+}-CS	**0.9098**	**0.9689**	**0.9823**	**0.9885**	**0.9914**	**0.9925**

Outline

Conclusions

- We propose the Partial-Laplacian filtering sparsity model to model the structured dynamic sparsity of the realistic channel.
- We establish a unified DCS framework (PLAY-CS) that exhibits versatility by encompassing various existing DCS algorithms.
- We develop a variant of the DCS algorithm, leveraging the Partial-LSM filtering sparsity model we introduced. We call the new DCS algorithm PLAY$^+$-CS.
- We show the enhanced performance of the PLAY$^+$-CS algorithm compared to existing DCS algorithms through the realistic channel tracking testing.

Thanks for Your Attention!

Email: xzliu@buaa.edu.cn

https://arxiv.org/abs/2310.07202