Density Evolution of Sparse Source Signals

Erich Zöchmann
Institute of Telecommunications
Vienna University of Technology
1040 Vienna, Austria
ezoechma@nt.tuwien.ac.at

Peter Gerstoft
Scripps Institution of Oceanography
La Jolla, CA 92093-0238, USA
gerstoft@ucsd.edu

Christoph F. Mecklenbräuker
Christian Doppler Lab Wireless
Technologies for Sustainable Mobility
Vienna University of Technology
1040 Vienna, Austria
cfm@nt.tuwien.ac.at

***Abstract*—A sequential Bayesian approach to density evolution for sparse source reconstruction is proposed and analysed which alternatingly solves a generalized LASSO problem and its dual. Waves are observed by a sensor array. The waves are emitted by a spatially-sparse set of sources. A weighted Laplace-like prior is assumed for the sources such that the maximum a posteriori source estimate at the current time step is the solution to a generalized LASSO problem. The posterior Laplace-like density at step k is approximated by the corresponding dual solution. The posterior density at step k leads to the prior density at $k+1$ by applying a motion model. Thus, a sequence of generalized LASSO problems is solved for estimating the temporal evolution of a sparse source field.**

***Index Terms*—sequential estimation, Bayesian estimation, sparsity, generalized LASSO.**

I. Introduction

In this contribution, the online estimation of sparse signals is solved from noisy data samples that become available sequentially in time [2], [3], [4]. The proposed online estimator alternatingly solves a generalized LASSO problem and its dual. Besides the actual reconstruction, we are also interested in the probability density of the source amplitudes and their temporal evolution.

Previously, the Bayesian approach [5], [6], [7] was extended to *sequential* Maximum A Posteriori (MAP) estimation for sparse signals [3], [1]. A sequential MAP filter which preserves sparsity was approximated based on [9]. It uses a single new measurement snapshot in each step.

The theory is formulated so that it is applicable to sparse source estimation in higher spatial dimensions.

II. Dual problem to the generalized LASSO

The generalized LASSO as introduced in [8] penalizes a weighted sum of the optimization variables

$$\min_{\boldsymbol{x}} \left(\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \mu \|\boldsymbol{D}\boldsymbol{x}\|_1 \right) \ , \tag{1}$$

where $\boldsymbol{A}$ is the complex-valued dictionary, $\boldsymbol{x}, \boldsymbol{y}$ are complex-valued vectors and $\mu > 0$. We seek solutions $\boldsymbol{x}$ with given sparsity degree $s \in \mathbb{N}$,

$$\|\boldsymbol{x}\|_0 = s \ . \tag{2}$$

The financial support by the Austrian Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development is gratefully acknowledged.

The regularization parameter μ is chosen to satisfy (2).

Following [11], the generalized complex-valued LASSO problem is re-written as

$$\min_{\boldsymbol{x},\boldsymbol{z}} \left(\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \mu \|\boldsymbol{z}\|_1 \right) \quad \text{subject to} \quad \boldsymbol{z} = \boldsymbol{D}\boldsymbol{x} \ , \tag{3}$$

but now we restrict $\boldsymbol{D}$ to be diagonal with real, positive entries. This substitution provides a Lagrangian multiplier for each element in $\boldsymbol{x}$, and in fact these Lagrangian multipliers will update the corresponding hyperparameters later on.

The dual problem to the generalized LASSO (3) is [8], [10],

$$\max_{\boldsymbol{u} \in \mathbb{C}^M} \boldsymbol{y}^H \boldsymbol{y} - \|\boldsymbol{A}\boldsymbol{A}^+ \boldsymbol{y} - \tfrac{1}{2} \left(\boldsymbol{D}\boldsymbol{A}^+ \right)^H \boldsymbol{u}\|_2^2 \tag{4a}$$

$$\text{subject to} \qquad \|\boldsymbol{u}\|_\infty \leq \mu, \tag{4b}$$

$$(\boldsymbol{D}\boldsymbol{U})^H \boldsymbol{u} = \boldsymbol{0} \ , \tag{4c}$$

where $\boldsymbol{A}^+$ denotes the Moore-Penrose inverse of the dictionary $\boldsymbol{A}$. Having solved the primal problem, the corresponding dual solution is easily computed with the help of the following theorem [10].

Theorem 1. *The dual vector $\boldsymbol{u}$ is the output of a weighted classical beamformer (weighted matched filter) acting on the vector of residuals, i.e.,*

$$\boldsymbol{u} = 2(\boldsymbol{D}^H)^{-1} \boldsymbol{A}^H (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_{\ell_1}) \ , \tag{5}$$

where $\boldsymbol{x}_{\ell_1}$ - the solution to the primal problem (17) - is such, that the box constraint

$$\|\boldsymbol{u}\|_\infty \leq \mu \tag{6}$$

is fulfilled.

Built upon Theorem 1 the following corollary was proven in [10].

Corollary 1. *If the mth primal coordinate is active, i.e. $x_{\ell_1,m} \neq 0$ then the box constraint (6) is tight in the mth dual coordinate. Formally, for any choice $\delta > 0$,*

$$|x_{\ell_1,m}| > \delta \quad \Rightarrow \quad |u_m| = \mu, \qquad (m = 1, \ldots, M) \ . \tag{7}$$

Informally, we say that the mth dual coordinate hits the boundary when the mth primal coordinate becomes active.

We define the active set $\mathcal{M}$ as the set of all indices m with $|x_{\ell_1,m}| > \delta$,

$$\mathcal{M} = \{m \mid |x_{\ell_1,m}| > \delta\}. \tag{8}$$

III. Direction of Arrival Estimation

For the numerical examples, we model a uniform linear array (ULA) which is described with its steering vectors representing the incident wave for each array element.

Let $\boldsymbol{x} = (x_1, \dots, x_M)^{\mathrm{T}}$ be a vector of complex-valued source amplitudes. We observe time-sampled waveforms on an array of N sensors which are stacked in the vector $\boldsymbol{y}$. The following linear model for the narrowband sensor array data $\boldsymbol{y}$ at frequency ω is assumed,

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n} \ . \tag{9}$$

The mth column of the transfer matrix $\boldsymbol{A}$ is the array steering vector $\boldsymbol{a}_m$ for hypothetical waves from direction of arrival (DOA) θ_m. All columns are normalized such that their ℓ_2 norm is one. The transfer matrix $\boldsymbol{A}$ is constructed by sampling all possible directions or arrival, but only very few of these correspond to real sources. Therefore, the dimension of $\boldsymbol{A}$ is $N \times M$ with $N \ll M$ and $\boldsymbol{x}$ is sparse. The linear model equations (9) are under-determined.
The nmth element of $\boldsymbol{A}$ is modelled by

$$A_{nm} = \frac{1}{\sqrt{N}} \exp\left[-\mathrm{j}(n-1)\pi \sin\theta_m\right]. \tag{10}$$

Here $\theta_m = \frac{\pi(m-1)}{M} - \pi/2$ is the DOA of the mth hypothetical DOA to the nth sensor element of the sensor array.

The additive noise vector $\boldsymbol{n}$ is assumed to be spatially uncorrelated and follows the zero-mean complex normal distribution with diagonal covariance matrix $\sigma^2\boldsymbol{I}$.

For the observation $\boldsymbol{y}$ according to the linear model (9), the conditional probability density given the source vector $\boldsymbol{x}$ is

$$p(\boldsymbol{y}|\boldsymbol{x}) = \frac{\exp\left(-\frac{1}{\sigma^2}\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}\|_2^2\right)}{(\pi\sigma^2)^N} \ . \tag{11}$$

For the source vector $\boldsymbol{x}$, a prior probability density is assumed in form of a multivariate complex Laplace-like density [12],

$$p(\boldsymbol{x}) = \prod_{m=1}^{M} p_m(x_m), \text{ with } p_m(x) = \frac{(\lambda_m)^2}{2\pi}\mathrm{e}^{-\lambda_m|x_m|} \ , \tag{12}$$

with associated hyperparameters $\lambda_m > 0$ modelling the source signal strength at location θ_m. $x_m = |x_m|\mathrm{e}^{j\varphi_m}$ is the complex source signal at hypothetical source location θ_m. Note that (12) defines the joint distribution for $|x_m| = r_m$ and φ_m with the phases uniformly distributed on $[0, 2\pi)$, for $m = 1, \dots, M$. The prior mean and variances are

$$\mathbb{E}\{\boldsymbol{x}\} = \boldsymbol{0}, \qquad \mathbb{E}\{\boldsymbol{x}\boldsymbol{x}^H\} = 6\,\mathrm{diag}\left(\lambda_1^{-2}, \dots, \lambda_M^{-2}\right) \ . \tag{13}$$

Taking the logarithm of (12) gives

$$-\ln p(\boldsymbol{x}) = \sum_{m=1}^{M} \lambda_m|x_m| - 2\sum_{m=1}^{M} \ln\lambda_m + M\ln 2\pi \ . \tag{14}$$

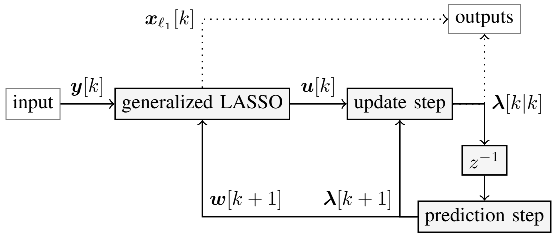

Fig. 1. Signal flow diagram for sequential Bayesian estimation at step k

For the posterior probability density function (pdf) $p(\boldsymbol{x}|\boldsymbol{y})$, Bayes' rule is used for obtaining the generalized LASSO Lagrangian [8], [1]

$$\frac{1}{\sigma^2}\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}\|_2^2 + \mu\,\|\boldsymbol{W}\boldsymbol{x}\|_1 \tag{15}$$

with bounded weights $\|\boldsymbol{w}\|_\infty = 1$

$$\boldsymbol{W} = \mathrm{diag}(\boldsymbol{w}) = \frac{1}{\mu}\,\mathrm{diag}(\boldsymbol{\lambda}) \ . \tag{16}$$

Equivalently to (15), this is reformulated as

$$\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}\|_2^2 + \mu\,\|\boldsymbol{D}\boldsymbol{x}\|_1 \ , \tag{17}$$

with

$$\boldsymbol{D} = \sigma^2\boldsymbol{W} \ . \tag{18}$$

The minimization of (17) constitutes a strictly convex optimization problem. Minimizing the generalized LASSO Lagrangian (17) with respect to $\boldsymbol{x}$ for given μ, and $\boldsymbol{w} = (w_1, \dots, w_M)^T$, $\boldsymbol{\lambda} = \mu\boldsymbol{w}$, gives a sparse MAP source estimate $\boldsymbol{x}_{\ell_1}$. This minimization problem promotes sparse solutions in which the ℓ_1 constraint is weighted by giving every source amplitude its own hyperparameter w_m.

IV. Sequential Bayesian Estimation

In [1], a sequential Bayesian sparse source reconstruction was proposed and analyzed which is now interpreted as solving both the generalized LASSO problem (3) and its dual (4a)–(4c) at step k. In the following, the dependency of time is denoted explicitly in all relevant variables, e.g. $\boldsymbol{y}[k]$ to denote the data at step k.

First, the history of all previous array observations is summarized in $\boldsymbol{Y}[k-1] = (\boldsymbol{y}[1], \dots, \boldsymbol{y}[k-1])$. Given the history $\boldsymbol{Y}[k-1]$ and the new data $\boldsymbol{y}[k]$, we seek the maximum a posteriori (MAP) source estimate $\boldsymbol{x}_{\ell_1}[k]$ for the linear model

$$\boldsymbol{y}[k] = \boldsymbol{A}\boldsymbol{x}[k] + \boldsymbol{n}[k] \ , \tag{19}$$

at step k under the ℓ_1–constraint. The additive noise $\boldsymbol{n}[k]$ is assumed to be both spatially and temporally white,

$$\mathrm{E}(\boldsymbol{n}[k]\boldsymbol{n}^H[k+l]) = \begin{cases} \sigma^2\boldsymbol{I}, & \text{for } l = 0, \\ \boldsymbol{0} & \text{otherwise.} \end{cases} \tag{20}$$

The algorithm in [1] is reformulated in terms of the vector of dual variables in Table I and shown schematically in Figure 1. It is actually the vector of dual variables which carries the sequential information from each step, and not the primal variables as customary in sequential filtering [14].

A. Update Step

In [1] two approximations were introduced in order to relate the posterior weight vector $\boldsymbol{\lambda}[k|k]$ to the prior weight vector $\boldsymbol{\lambda}[k]$ in the form of (12). By means of Theorem 1 and Corollary 1, both approximations for the posterior weight vector are expressible by the dual solution. In the sequel we express the superior approximation, the mean fit, by the dual vector.
In the complement of the active set, the relation between posterior and prior weight vector is given as

$$\lambda_m[k|k] = \lambda_m[k]\left(1 - \frac{\left|\boldsymbol{a}_m^H \frac{2}{\sigma^2}(\boldsymbol{y}[k] - \boldsymbol{A}\boldsymbol{x}_{\ell_1})\right|^2}{\lambda_m^2[k]}\right) \quad \forall m \notin \mathcal{M}[k], \tag{21}$$

and in the active set the posterior weight vector must zero.

$$\lambda_m[k|k] = 0, \quad \forall m \in \mathcal{M}[k] \tag{22}$$

By Theorem 1 we express the numerator of (21) by the dual vector $\boldsymbol{u}$ and the weights $\boldsymbol{w}$. Corollary 1 links Equations (21) and (22), as (21) is zero for $|u_m| = \mu$.

Theorem 2. *With the mean fit approximation, the posterior weight vector $\boldsymbol{\lambda}[k|k]$ is related to the prior weight vector $\boldsymbol{\lambda}[k]$ by the dual solution $\boldsymbol{u}[k]$ at step k,*

$$\lambda_m[k|k] = \lambda_m[k]\left(1 - \frac{|u_m[k]|^2}{\mu^2[k]}\right) . \tag{23}$$

Due to Theorem 1 and Corollary 1, μ is equal to the max-norm of $\boldsymbol{u}$ and Theorem 2 is expressible solely by the dual vector $\boldsymbol{u}$

$$\lambda_m[k|k] = \lambda_m[k]\left(1 - \frac{|u_m[k]|^2}{\|\boldsymbol{u}[k]\|_\infty^2}\right) . \tag{24}$$

Equation (24) shows that the dual coordinate equals μ and the posterior weights become zero at source positions $m \in \mathcal{M}$. Outside the active set, the probability of finding a source depends on the relative sidelobe power level of the beamformer of the LASSO residuals, cf. Theorem 1.

B. Prediction Step

In sequential estimation, typically the prior for the upcoming step $k+1$ is calculated from the current posterior and a state-transition probability density function ("motion model"). In a Markovian stochastic framework this is based on the Chapman-Kolmogorov equation [14]. For Brownian motion the state-transition probability density satisfies the diffusion equation. Our prediction step is therefore based on a diffusion model. Where diffusion occurs just in the neighbourhood of active sources.

1) Neighborhood of an active source: The index neighborhood of m is denoted as $\mathcal{N}_m = \{j \mid m-l, \dots, m+l\}$. If any $\lambda_j[k|k] \in \mathcal{N}_m$ is less than the threshold λ_0 then a source is active in the neighborhood of m with high probability.

The motion model is defined via the complementary cumulative distribution function (ccdf) of the neighborhood magnitudes,

$$\mathrm{P}\left\{|x_m[k+1]| > \delta \middle| \boldsymbol{x}[k]\right\} = \sum_{j=-l}^{l} \alpha_j \,\mathrm{P}\left\{|x_{m+j}[k]| > \delta\right\} , \tag{25}$$

with non-negative coefficients α_j and $\sum_j \alpha_j = 1$. The ccdf, after a polar coordinate transformation $x_m = r_m e^{i\varphi_m}$, evaluates to

$$\begin{aligned}
&\mathrm{P}\left\{|x_m[k+1]| > \delta \middle| \boldsymbol{x}[k]\right\} = \qquad (26)\\
&= \sum_{j=-l}^{l} \alpha_j \int_{-\pi}^{\pi} \frac{\mathrm{d}\varphi_{m+j}}{2\pi} \int_{\delta}^{\infty} \lambda_{m+j}^2[k|k] e^{-\lambda_{m+j}[k|k] r_{m+j}}\; r_{m+j}\, \mathrm{d}r_{m+j}\\
&= \sum_{j=-l}^{l} \alpha_j \left(\lambda_{m+j}[k|k]\,\delta + 1\right) e^{-\lambda_{m+j}[k|k]\,\delta} ,
\end{aligned}$$

and by taking the negative derivative w.r.t. δ, we obtain the magnitude's probability density

$$\begin{aligned}
&-\frac{\partial}{\partial\delta}\mathrm{P}\left\{|x_m[k+1]| > \delta \middle| \boldsymbol{x}[k]\right\} =\\
&= 2\pi\delta \sum_{j=-l}^{l} \alpha_j \frac{(\lambda_{m+j}[k|k])^2}{2\pi}\; e^{-\lambda_{m+j}[k|k]\,\delta}
\end{aligned} \tag{27}$$

which is a mixture of Erlang-2 distributions with variances $2(\lambda_{m+j}[k|k])^{-2}$, cf. Eq. (12).

We approximate the mixture by a Laplace-like density of the form (12). We choose to fit the variance of the Laplace-like density such that

$$\frac{1}{(\lambda_m[k+1])^2} = \sum_{j=-l}^{l} \alpha_j \frac{1}{(\lambda_{m+j}[k|k])^2} . \tag{28}$$

We note that (28) is ill-behaved whenever a posterior weight $\lambda_{m+j}[k|k] = 0$. In this case, a small offset $\varepsilon > 0$ is added to stabilize (28) numerically. The predicted $\boldsymbol{\lambda}[k+1]$ is the product of the regularization parameter $\mu[k+1]$ and the weights $\boldsymbol{w}[k+1]$. As $\mu[k+1]$ is not yet known at step k, we need to assume that the regularization parameter remains constant between k and $k+1$, i.e.,

$$\frac{1}{(\lambda_m[k+1])^2} = \frac{1}{(\mu[k+1] w_m[k+1])^2} \approx \frac{1}{(\mu[k] w_m[k+1])^2} . \tag{29}$$

The predicted weights $w_m[k+1]$ are then calculated from the weighted harmonic mean, i.e.,

$$(w_m[k+1])^2 = \left(\sum_{j=-l}^{l} \frac{\alpha_j}{(w_{m+j}[k])^2}\right)^{-1} . \tag{30}$$

The weighted harmonic mean is a special instance of the weighted Hölder mean [13], see Sec. V. To express the uncertainty of the prediction, the weights are increased by an offset $w_0 > 0$, similar to process noise in Kalman filtering.

	Implementation of density evolution procedure: Given constants: $\boldsymbol{A} \in \mathbb{C}^{N \times M}$, $\boldsymbol{w}[1] \in [0,1]^M$, $s \in \mathbb{N}$
1:	for $k = 1, 2, 3, \ldots$
2:	Input: $\boldsymbol{y}[k] \in \mathbb{C}^N$
	$\boldsymbol{w}[k] = \boldsymbol{w}[k] / \|\boldsymbol{w}[k]\|_\infty$
3:	$\boldsymbol{D}[k] = \sigma^2 \operatorname{diag}(\boldsymbol{w}[k])$
4:	$\boldsymbol{x}_{\ell_1}[k] =$ s-sparse solution to generalized LASSO (3) at k
5:	$\boldsymbol{u}[k] =$ corresponding dual solution via Theorem 1
6:	$\mu[k] = \|\boldsymbol{u}[k]\|_\infty$
7:	Update $\boldsymbol{\lambda}[k\|k]$ via Theorem 2
8:	$\boldsymbol{w}[k+1] =$ motion model prediction$(\boldsymbol{\lambda}[k\|k])$
9:	Output: $\boldsymbol{x}_{\ell_1}[k] \in \mathbb{C}^M$, $\boldsymbol{\lambda}[k\|k] \in \mathbb{C}^M$
10:	end

TABLE I
PRIMAL/DUAL FORMULATION OF SEQUENTIAL BAYESIAN SPARSE SIGNAL RECONSTRUCTION

2) Not in the neighborhood of an active source: The posterior $\lambda_j[k|k]$ exceeds the threshold λ_0 which indicates that it is improbable for a source to be near DOA θ_j. At step $k+1$, we penalize the DOA j by adding a multiple of weight uncertainty w_0, i.e., $w_m[k+1] = w_m[k] + c\,w_0$ with $c > 1$. In the simulations, $w_0 = 0.01$ and $c = 10$.

To guarantee that the weights remain upper bounded by 1, the weighting vector is normalized to $\|\boldsymbol{w}\|_\infty = 1$.
The Bayesian procedure is formalized in Table I as a loop over time step k which processes the single snapshot array observation $\boldsymbol{y}[k]$ when it becomes available. In line 3, the weighting coefficients for the generalized LASSO problem (3) are defined for the current step k. The s-sparse solution in line 4 is implemented via the LASSO path [1], [10]. Next, the corresponding dual solution is evaluated by weighted beamforming of the residuals. Finally, the posterior weighting coefficients are evaluated in line 7 which are needed for the prediction step in line 8.

V. CONSERVATIVE CHOICE OF THE WEIGHTS

The weighted harmonic mean (30) is a pessimistic mean as low values have stronger impact on the mean. Generally, it tends to broaden the low weight region. This broad low weight region leads to a jitter of the DOA estimate. To mitigate this undesirable effect, we investigate alternative rules for the predicted weights.
A weighted Hölder mean is defined as [13]

$$M_p(w_1^2, \ldots, w_n^2) = \left(\sum_{j=l}^{l} \alpha_j \left(w_j^2\right)^p \right)^{\frac{1}{p}}, \qquad \sum_{j=l}^{l} \alpha_j \quad = 1 \ . \tag{31}$$

For the choice of power $p = -1$, the weighted Hölder mean coincides with the weighted harmonic mean (30). The following inequality holds for weighted Hölder means,

$$M_p < M_q, \quad \text{for} \quad p < q \ . \tag{32}$$

Any Hölder mean with $p > 0$ will not be dominated by lower weights and the arithmetic mean ($p = 1$) is the tightest conservative choice of weighting coefficients for Laplace-like prior. [9] has used a max-log approximation instead of (28) which amounts to picking $M_{+\infty}$, the least tight bound.

VI. SIMULATIONS

A. Weight Evolution

We investigate the weight evolution from step $k = 1$ to $k = 100$, where the generalized LASSO of Table I is solved by CVX [15] at each step. The ULA is equipped with $N = 30$ sensors and the angular space is sampled equidistantly with half degree spacing between $-90°$ and $90°$.
In Figure 2 the weight evolution of sources with trivial motion model, $l = 0$ and $\alpha_0 = 1$ is shown. In Figure 3 movement is modelled with a uniform motion model ($l = 2$, $\alpha_j = \frac{1}{2l+1} = 0.2$). Observe the trade off between having precise estimates for the static sources and a good quality estimate of the moving source.

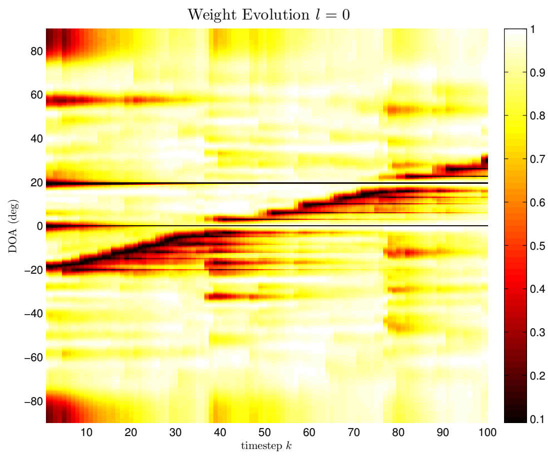

Fig. 2. Weight evolution for 3 sources at DOA $20°, 0°, -20°$, the third source moves with $0.5°$ per time step; $w_0 = 0.01$, $c = 1$, SNR $= 20$dB

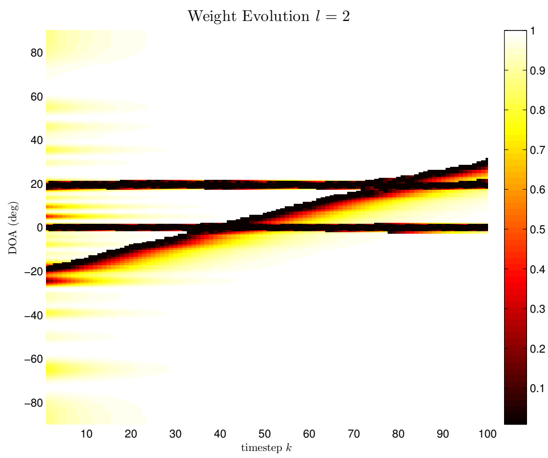

Fig. 3. Weight evolution for 3 sources at DOA $20°, 0°, -20°$, the third source moves with $0.5°$ per time step; $w_0 = 0.01$, $c = 10$, SNR $= 20$dB

A reasonable compromise of capturing the motion of a source while still improving the estimate of the static sources is to use an $l > 0$ and a conservative choice of the weights. Figure 4 uses the same motion model as in Fig. 3, but the weighted arithmetic mean is used, i.e. Equation (31) for power parameter $p = 1$. For the arithmetic mean, the low weight region of the static sources is narrower than for the harmonic mean. This comes at the expense of the traceability of the moving source.

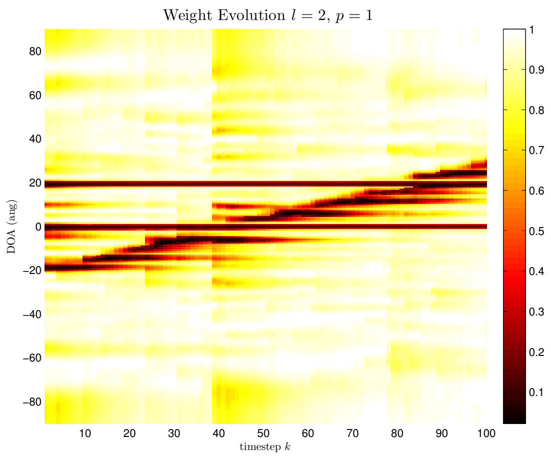

Fig. 4. Conservative weight evolution for 3 sources at DOA $20°; 0°; -20°$, the third source moves with $0.5°$ per time step; $w_0 = 0.01$, $c = 10$, SNR $=$ 20dB

B. *Comparison of the Tracking Results*

The proposed DOA tracking procedure from Table I is compared to "Compressive Sensing on Kalman filtered residuals (KF-CS)" [17] in Figure 5. For KF-CS μ is chosen non-adaptively analogous to the value given in [18]:Algorithm 1. The density evolution approach with $p = -1$ mean recovers the static sources worse than the Kalman filter and the conservative ($p = 1$) approach, but in return the moving source is traced well.

VII. Conclusion

A sequential reconstruction procedure was proposed which uses both the primal and the dual solution to the generalized LASSO. The dual variable is propagated to the update step, which approximates the posterior distribution with a Laplace-like distribution (see Fig. 1). From the approximated posterior and a motion model, the prior for the next step is derived and the procedure is ready for the next step. Without the prediction step, the proposed procedure is fully equivalent to the procedure in [1]. By including the motion model and prediction step, we show superior performance by means of a synthetic example.

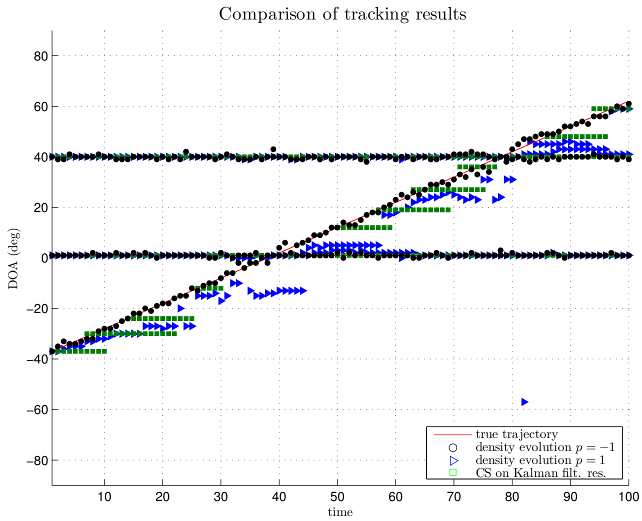

Fig. 5. Active set $\mathcal{M}$ of different tracking algorithms, SNR $=$ 20dB.

References

[1] C. F. Mecklenbräuker, P. Gerstoft, A. Panahi, M. Viberg: Sequential Bayesian Sparse Signal Reconstruction Using Array Data, IEEE Trans. Signal Process., vol. 61, no. 24, pp. 6344–6354, Dec. 2013.

[2] D. M. Malioutov, S. Sanghavi, A. S. Willsky: Sequential Compressed Sensing, IEEE J. Sel. Top. Sig. Proc, vol. 4, pp. 435–444, 2010.

[3] D. Angelosante, J. A. Bazerque, G. B. Giannakis: Online adaptive estimation of sparse signals: Where RLS meets the ℓ_1-norm, IEEE Trans. Signal Process., vol. 58, no. 7, pp. 3436–3447, 2010.

[4] D. Eiwen, G. Tauböck, F. Hlawatsch, H. G. Feichtinger: Compressive tracking of doubly selective channels in multicarrier systems based on sequential delay-Doppler sparsity, in Proc. ICASSP 2011, pp. 2928–2931, Prague, Czech Republic, May 2011.

[5] D. P. Wipf, B. D. Rao: An Empirical Bayesian Strategy for Solving the Simultaneous Sparse Approximation Problem, IEEE Trans. Signal Process., vol. 55, no. 7, pp. 3704–3716, Jul. 2007.

[6] S. Ji, Y. Xue, L. Carin: Bayesian compressive sensing, IEEE Trans. Signal Process., vol. 56, no. 6, pp. 2346–2356, Jun. 2008.

[7] S. D. Babacan, R. Molina, A. K. Katsaggelos: Bayesian compressive sensing using Laplace priors, IEEE Trans. Image Proc., vol. 19, no. 1, pp. 53–63, Jan. 2010.

[8] R. J. Tibshirani and J. Taylor: The Solution Path of the Generalized Lasso, Ann. Statist., vol. 39, no. 3, pp. 1335–1371, Jun. 2011.

[9] A. Panahi, M. Viberg: Fast LASSO based DOA tracking, in Proc. Fourth International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (IEEE CAMSAP 2011), San Juan, Puerto Rico, Dec. 2011.

[10] C. F. Mecklenbräuker, P. Gerstoft, E. Zöchmann: Beamforming of the Residuals is the LASSO's Dual, submitted to IEEE Trans. Signal Process., Dec. 2014, online: `http://arxiv.org/abs/1502.04643`

[11] S. J. Kim, K. Koh, S. Boyd, D. Gorinevsky, l_1 trend filtering. SIAM Rev., vol. 51, pp. 339–360. MR2505584.

[12] Z. He, S. Xie, S. Ding, A. Cichocki: Convolutive blind source separation in the frequency domain based on sparse representation, IEEE Trans. Audio, Speech, and Language Proc., vol. 15, no. 5, Jul. 2007.

[13] P.S. Bullen: *Handbook of means and their inequalities*, Springer, 2003.

[14] B. Ristic, S. Arulampalam, N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Boston, MA: Artech House, 2004.

[15] M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, version 1.21, http://cvxr.com/cvx, Apr. 2011.

[16] A. Xenaki, P. Gerstoft and K. Mosegaard: Compressive beamforming, The Journal of the Acoustical Society of America, vol. 136, no. 1, pp. 260–271, 2014

[17] N. Vaswani. "KF-CS: compressive sensing on Kalman filtered residual." arXiv:0912.1628, 2009.

[18] C. Qiu, W. Lu, and N. Vaswani: "Real-time dynamic MR image reconstruction using Kalman filtered compressed sensing." IEEE International Conference on Acoustics, Speech and Signal Processing, 2009.