Super-resolution Sparse Channel Estimation for Localization, Prospects and Issues

……………………………………………………..

Bernard H. Fleury [1]

[1] *Department of Electronic Systems, Aalborg University,*
fleury@es.aau.dk

Mots clés (en français et en anglais) : Apprentissage Bayésien de Signaux Epars, Localisation dans les Réseaux de Senseurs Sans Fil,
Sparse Bayesian Learning, Localization in Wireless Sensor Networks

Présentation en session plénière / Plenary session communication

Résumé / Abstract

In this talk we consider a class of sparse Bayesian algorithms for the estimation of wireless channels in the context of their application to localization. In a nutshell, these algorithms aim at detecting and estimating dominant "specular-like multipath components" in the channel response. Specifically, the number of said components and their parameters, such as their relative delay and complex amplitude, are estimated. This information can be exploited for localization purpose, e.g. by finger-printing or by reconstructing the corresponding physical propagation paths between transmitter and receiver, as done in SLAM.

We discuss the key properties of these algorithms, such as their ability to detect components and to resolve them in the dispersion domain (e.g. with respect to their relative delay). We also shed some light on the correct interpretation of "components" extracted by such algorithms (and actually by any parametric algorithm). We discuss the implications of these properties on localization schemes based on multipath reconstruction.

Super-resolution Sparse Channel Estimation for Localization - Prospects and Issues

Bernard H. Fleury

Aalborg University (AAU), Denmark

Contributors: Thomas L. Hansen (AAU), Erik Leitinger (TU Graz), Stefan Grebien (TU Graz)

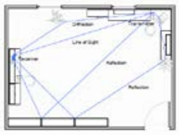
Joint Work with TU Graz

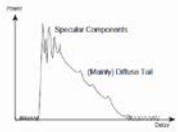
29 March 2018

JS18 URSI France

Context / Motivation I

Multipath propagation:



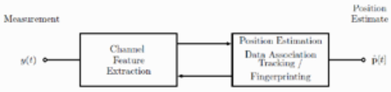


Localization exploiting multipath propagation:

- Multipath-propagarion-based finger printing
- Multipath-propagation-based simultaneous localization and mapping (SLAM)

Context / Motivation II

Generic architecture of a (two-stage) localization algorithm:



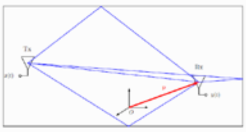
Relevant issues regarding the channel estimator:

- Detection of artefacts,
- Miss of components
- Superresolution capability
- Physical reality of estimated multipath components

Table of Contents

Signal Model - Time Dispersion

The channel as a linear time-invariant system:



Frequency domain characterization:

- $X(f)$: transmitted signal with bandwidth B
- $Y(f) \triangleq Y(f;\boldsymbol{p})$: received signal
- $W(f)$: White Gaussian noise
- $H(f) \triangleq H(f;\boldsymbol{p})$: Channel (frequency) transfer function

Input-output relationship:

$$Y(f) = H(f)X(f) + W(f)$$

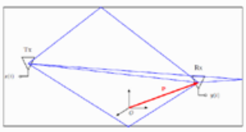
Traditional assumption:

$$H(f) = \sum_{\tilde{\ell}}^{\tilde{L}} \tilde{\alpha}_{\tilde{\ell}} \exp\{-j2\pi f \tilde{\tau}_{\tilde{\ell}}\}$$

$\tilde{L}$ small: specular channel; $\tilde{L}$ large: dense (diffuse) channel

Signal Model - Time Dispersion

The channel as a linear time-invariant system:



Channel sounding (OFDM):

- $x(t)$: transmitted signal
- $y(t) \triangleq y(t;\boldsymbol{p})$: received signal
- $w(t)$: white Gaussian noise
- $H(f) \triangleq H(f;\boldsymbol{p})$: channel (frequency) transfer function

Traditional assumption:

$$H(f) = \sum_{\tilde{\ell}}^{\tilde{L}} \tilde{\alpha}_{\tilde{\ell}} \exp\{-j2\pi f \tilde{\tau}_{\tilde{\ell}}\}$$

$\tilde{L}$ small: specular channel; $\tilde{L}$ large: dense (diffuse) channel

Discrete-time Model

$$\boldsymbol{y} = \boldsymbol{h} + \boldsymbol{w}$$

where

- $\boldsymbol{y} \in \mathbb{C}^N$: received symbol
- $\boldsymbol{w} \in \mathbb{C}^N$: channel noise
- We assume an all-one OFDM symbol $[1,\dots,1]^{\mathsf{T}}$.

and

$$\boldsymbol{h} \triangleq [h_1,\dots,h_N]^{\mathsf{T}} \triangleq [H(n\Delta f) : n = 0,\dots,N-1]^{\mathsf{T}}$$

with

- Δf: carrier spacing
- N: number of carrier

Under the traditional assumption:

$$\boldsymbol{h} = \boldsymbol{\Psi}(\tilde{\boldsymbol{\theta}})\,\tilde{\boldsymbol{\alpha}}$$

$$\boldsymbol{\Psi}(\tilde{\boldsymbol{\theta}}) \triangleq [\psi(\tilde{\theta}_{\tilde{\ell}}) : \tilde{\ell} = 1,\dots,\tilde{L}]$$

$$\psi(\theta) \triangleq [\exp\{-j2\pi n\theta\} : n = 0,\dots,N-1]^{\mathsf{T}}$$

Frequency variable:
$\theta \in [-1/2, +1/2]$

$\tilde{\theta}_{\tilde{\ell}} \triangleq \Delta f \tilde{\tau}_{\tilde{\ell}}$

First- and Second-order Characterization

Expectation:

$$\mathbb{E}[\boldsymbol{h}] = \mathbb{E}[\boldsymbol{w}] = \mathbb{E}[\boldsymbol{y}] = 0$$

Covariance matrices:

$$\boldsymbol{\Sigma}_{\boldsymbol{hh}} \triangleq \mathbb{E}[\boldsymbol{h}\boldsymbol{h}^{\mathsf{H}}] \quad \text{Channel covariance matrix}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{ww}} = \beta^{-1}\boldsymbol{I}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{yy}} = \boldsymbol{\Sigma}_{\boldsymbol{hh}} + \beta^{-1}\boldsymbol{I}$$

Randon vectors are assumed circularly symmetric.

Rank of the channel covariance matrix:

$$L \triangleq \operatorname{rank}(\boldsymbol{\Sigma}_{\boldsymbol{hh}}) \le N$$

The rank is system-dependent: it depends on N for fixed Δf.

Carathéodory Representation

If $\mathbf{\Sigma}_{hh}$ is Toeplitz, it can be decomposed as

$$\mathbf{\Sigma}_{hh} = \sum_{\ell=1}^{L} \gamma_\ell \psi(\theta_\ell)\psi(\theta_\ell)^{\mathsf{H}} = \mathbf{\Psi}(\boldsymbol{\theta})\mathbf{\Gamma}\mathbf{\Psi}(\boldsymbol{\theta})^{\mathsf{H}}$$

$$L = \mathrm{rank}(\mathbf{\Sigma}_{hh}) \in \{1,\dots,N\}$$
$$\boldsymbol{\gamma} \triangleq [\gamma_\ell : \ell = 1,\dots,L]^{\mathsf{T}} \in (0,\infty)^L$$
$$\boldsymbol{\theta} \triangleq [\theta_\ell : \ell = 1,\dots,L]^{\mathsf{T}} \in [-1/2,+1/2)^L$$
$$\mathbf{\Gamma} \triangleq \mathrm{diag}(\boldsymbol{\gamma})$$

If $L < N$, the representation is unique.

It follows that

$$\boldsymbol{h} = \mathbf{\Psi}(\boldsymbol{\theta})\boldsymbol{\alpha} \qquad \boldsymbol{\alpha} \triangleq [\alpha_\ell : \ell = 1,\dots,L]^{\mathsf{T}} \in \mathbb{C}^L$$

Under the reasonable assumption that $\boldsymbol{h}$ is wide-sense stationary, $\mathbf{\Sigma}_{hh}$ is Toeplitz.

We assume that $L < N$, see numerical results.

Line Spectral Estimation

Generic model:

$$\boldsymbol{y} = \mathbf{\Psi}(\boldsymbol{\theta})\boldsymbol{\alpha} + \boldsymbol{w} = \sum_{\ell=1}^{L} \alpha_\ell \psi(\theta_\ell) + \boldsymbol{w}$$

Inference problem:

Estimate L, $(\alpha_\ell, \theta_\ell); \ell = 1,\dots,L$, and noise precision β!

The fact that L is unknown makes the problem more difficult ... and scientifically challenging.

Back to the traditional assumption:

Traditional assumption leads to the same representation of $\boldsymbol{h}$:

$$\boldsymbol{h} = \mathbf{\Psi}(\boldsymbol{\theta})\boldsymbol{\alpha} = \mathbf{\Psi}(\tilde{\boldsymbol{\theta}})\tilde{\boldsymbol{\alpha}}$$

While the components in $\mathbf{\Psi}(\tilde{\boldsymbol{\theta}})\tilde{\boldsymbol{\alpha}}$ are intended to have a physical meaning, those in $\mathbf{\Psi}(\boldsymbol{\theta})\boldsymbol{\alpha}$ do not. The latter are virtual components. They might coincides with a phisical component, but not always.
A sparse channel estimator exploits the low-rank structure of $\mathbf{\Sigma}_{hh}$. Thus it estimates the virtual components.

Signal Model for Inference

We fix the number of components:

$$\boldsymbol{y} = \sum_{m=1}^{M} \alpha_m \psi(\theta_m) + \boldsymbol{\alpha} \qquad M \geq N > L$$
$$= \mathbf{\Psi}(\boldsymbol{\theta})\boldsymbol{\alpha} + \boldsymbol{\alpha}$$

$\mathbf{\Psi}(\boldsymbol{\theta}) = [\psi(\theta_1),\dots,\psi(\theta_M)]$ Dictionary matrix (frequency dependent)

Rationale:

We use a sparse estimator that will set the estimates of the weight of surnumerous components to zero.

Our Choice: Sparse Bayesian learning

We use a Bayesian framework.

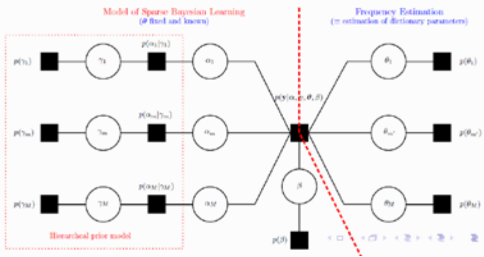
We use a sparsity inducing hierarchical prior for $\boldsymbol{\alpha}$, specifically a Gamma-Gaussian prior.

Sparse Bayesian Learning with Parametric Dictionary Estimation

Joint probability density function (pdf):

$$p(\boldsymbol{y},\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\theta},\beta) = \underbrace{p(\boldsymbol{y}|\boldsymbol{\alpha},\boldsymbol{\theta},\beta)}_{p(\boldsymbol{y}|\boldsymbol{\alpha},\mathbf{\Psi}(\boldsymbol{\theta}),\beta)} \underbrace{\prod_{m=1}^{M} p(\alpha_m|\gamma_m) \prod_{m'=1}^{M} p(\gamma_{m'})}_{\text{Hierarchical prior model}} p(\beta) \prod_{m''=1}^{M} p(\theta_{m''})$$

Factor graph:



Probabilistic Model in the Sparse Bayesian Framework

Joint pdf:

$$p(\boldsymbol{y},\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\theta},\beta) = p(\boldsymbol{y}|\boldsymbol{\alpha},\boldsymbol{\theta},\beta) \prod_{m=1}^{M} p(\alpha_m|\gamma_m) \prod_{m'=1}^{M} p(\gamma_{m'}) \prod_{m''=1}^{M} p(\theta_{m''})\, p(\beta)$$

where

$$p(\boldsymbol{y}|\boldsymbol{\alpha},\boldsymbol{\theta},\beta) = p_{\mathrm{CN}}(\boldsymbol{y};\mathbf{\Psi}(\boldsymbol{\theta})\boldsymbol{\alpha},\beta^{-1}\mathbf{I})$$
i.e. $\mathbf{n}$ is a white Gaussian noise vector
$$p(\alpha_m|\gamma_m) = p_{\mathrm{CN}}(\alpha_m;0,\gamma_m^{-1}), \quad m = 1,\dots,M$$
$$p(\gamma_m) = p_{\mathrm{G}}(\gamma;c,d), \quad m = 1,\dots,M$$
$$p(\theta_m) = p_{\mathrm{VM}}(\theta_m;\mu,\kappa), \quad m = 1,\dots,M$$
$$p(\beta) = p_{\mathrm{G}}(\beta;a,b)$$

The distribution we will consider subsequently, specifically their pdfs:

- $p_{\mathrm{CN}}(\mathbf{x};\mathbf{m},\mathbf{V}) = \frac{1}{\pi^M|\mathbf{V}|}\exp\{-(\mathbf{x}-\mathbf{m})^{\mathrm{H}}\mathbf{V}^{-1}(\mathbf{x}-\mathbf{m})\}$ complex Gaussian
- $p_{\mathrm{VM}}(\theta;\mu,\kappa) = \frac{1}{2\pi I_0(\kappa)}\exp\{\kappa\cos(\theta-\mu)\}, \quad \theta \in [-\pi,+\pi)$ von Mises
- $p_{\mathrm{G}}(x;u,v) = \frac{b^a}{\Gamma(a)}x^{-u-1}\exp\{-v/x\}, \quad x > 0$ Gamma

Mean-Field Approximation

Posterior pdf:

$$p(\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\theta},\beta|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{\alpha},\boldsymbol{\theta},\beta) \prod_{m} p(\alpha_m|\gamma_m) \prod_{m'} p(\gamma_{m'}) \prod_{m''} p(\theta_{m''})\, p(\beta)$$

Family of proxy pdfs:

We consider a family $\mathcal{Q}$ of pdfs with the "simpler" factorization

$$q(\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\theta},\beta) = q(\boldsymbol{\alpha}) \prod_{m} q(\gamma_m) \prod_{m'} q(\theta_{m'})\, q(\beta)$$

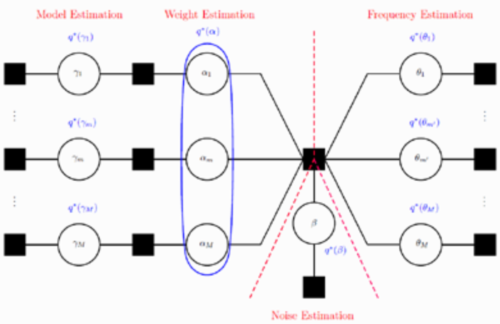
We select as an approximation of $p(\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\theta},\beta|\boldsymbol{y})$ the one element in $\mathcal{Q}$ that is the closest:

$$q^*(\mathbf{z}) = \arg\min_{q(\mathbf{z})\in\mathcal{Q}} \mathrm{KL}(q(\mathbf{z})\|p(\mathbf{z}|\mathbf{y})) \qquad \mathbf{z} \triangleq (\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\theta},\beta).$$

Computation of $q^*(\mathbf{z})$:

Different iterative approaches exist to compute (approximate) the solution to the arg min problem (e.g. variational EM).

Mean-Field Approximation - The Four Estimation Tasks



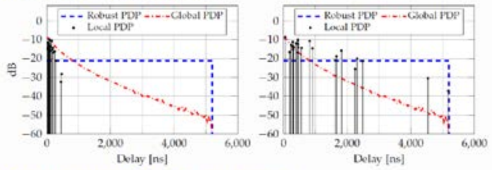
Facts about and Properties of SBL Estimators

- Various hierarchical models for α can be used that promote sparsity: Gamma-Gaussian, Bernouilli-Gaussian or Bernouilli-Gamma-Gaussian.
- SBL are Type II Bayesian estimators, while conventional methods, such as basis pursuit denoising (LASSO), atomic norm denoising, can be interpreted as Type I Bayesian estimator.
- Type II Bayesian estimators promote greater sparsity than Type I Bayesian estimators.
- SBL can also be interpreted as stochastic maximum-likelihood with model order estimation.
- SBL inherently integrates the order estimation process. No hypothesis testing problem needs to be solved, like in classical model order estimation based on information theoretic criteria.

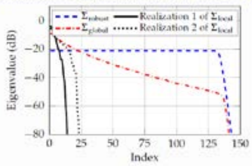
Numerical Studies - Rank of Radio Channels

A: Specular Channel - ITU-R M.2135 UMa NLOS Channel Model

Two generated impulse responses:



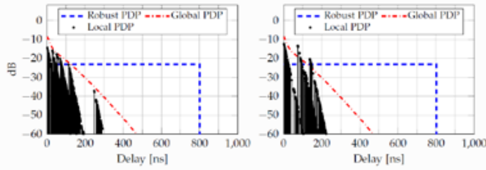
Covariance eigenvalues of the two responses:



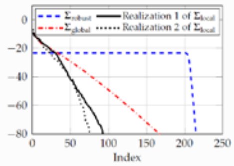
- Bandwidth: 25.6 MHz
- Sampling period: 25 KHz
- Dimension of $\mathbf{y}$: 1024

B: Dense Channel - IEEE 802.15.a Outdoor NLOS Channel Model

Two generated impulse responses:

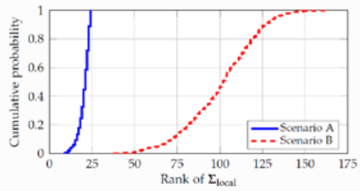


Covariance eigenvalues of the two responses:



- Bandwidth: 256 MHz
- Sampling period: 250 KHz
- Dimension of $\mathbf{y}$: 1024

Effective Rank of the Synthetic Channels



Experimental Results

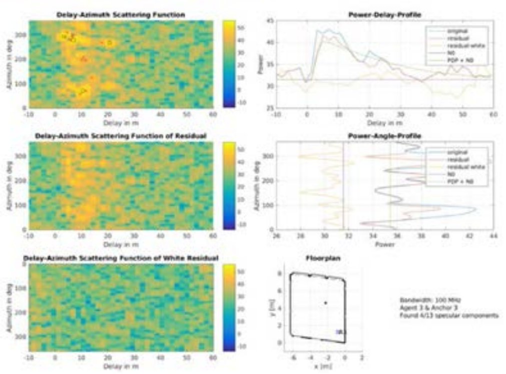
Description of the Measurement Experiment

Investigated indoor environment:

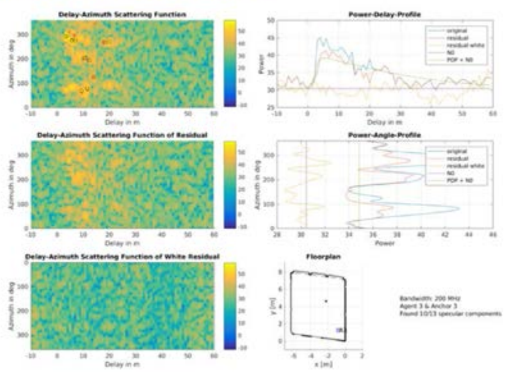
Equipment:

- Center frequency: 7 GHz
- Agents equipped with a 5x5 (virtual) array; array spacing: 2 cm
- Anchors equipped with a single antenna
- Bandwidth is varied: 100, 200, 500, 1000 MHz

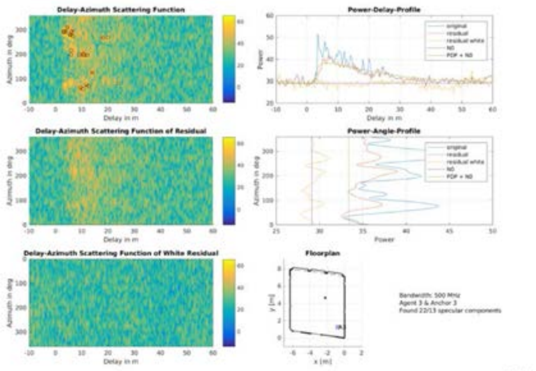
Bandwidth=100MHz, Agent 3, Anchor 3



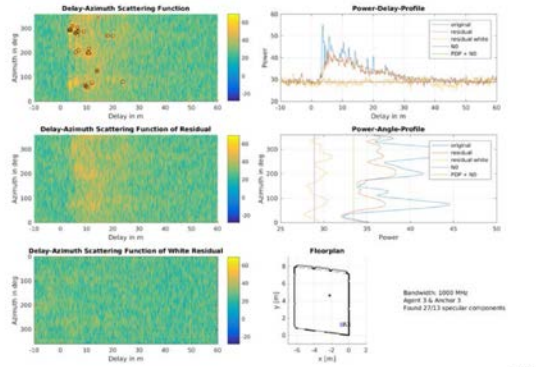
Bandwidth=200MHz, Agent 3, Anchor 3



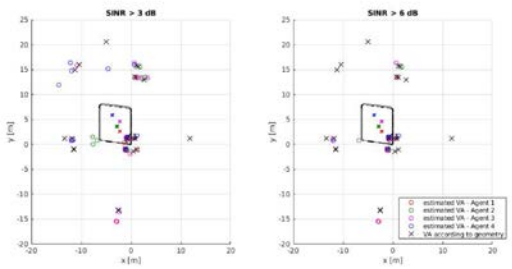
Bandwidth=500MHz, Agent 3, Anchor 3



Bandwidth=1000MHz, Agent 3, Anchor 3



Reconstructed and Estimated Mirror Sources



Conclusion and Outlook

Conclusion:

- The inherent threshold of SBL needs to be increased in order to decrease the number of artefacts (VALSE is an alternative).
- Superresolution demonstrated in synthetic channels (half the Nyquist period)
- SBL shows a sensible behaviour in real conditions.
- Caution is needed in the interpretation of estimated components as physical ones.
- Behaviour is strongly dependent on the selected iterative implementation.
- SBL works for both specular-like and diffuse channels

Outlook:

- Implementation in a localization estimation and tracking system

References

[BHF17] M. A. Badiu, T. Hansen, and B. Fleury, Variational bayesian inference of line spectra, IEEE Transactions on Signal Processing **PP** (2017), no. 99, 1–1.

[HFR18] T. L. Hansen, B. H. Fleury, and B. D. Rao, Superfast line spectral estimation, IEEE Transactions on Signal Processing **PP** (2018), no. 99, 1–1.