THALES

Werkgemeenschap voor
Informatie- en
Communicatietheorie

WIC 2019: TU/e, 25 januari 2019

Information-based Processing in Radar and Communications

Compressive Sensing and Information Geometry in Radar

Radmila Pribić
Radar Researcher Signal Processing
Sensors Advanced Developments
Thales Nederland BV
Delft

www.thalesgroup.com

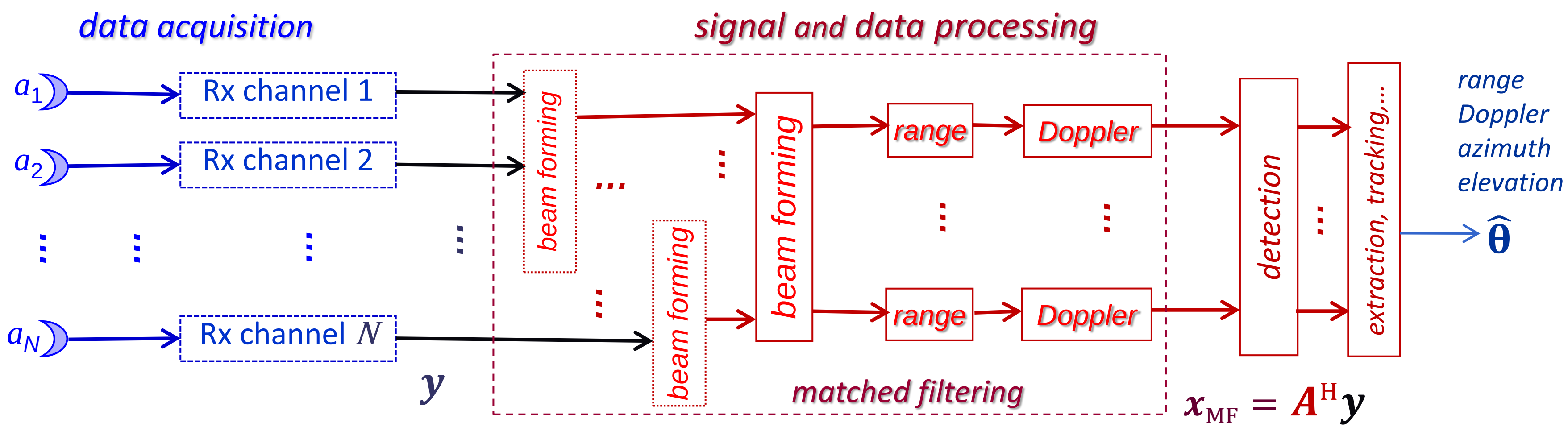
Contents *and* *Collaborations*

- ❑ Motivation: Information-based Radar Processing
- ❑ Compressive Sensing (CS)
- ❑ Information Geometry (IG)
- ❑ Information-based Performance Analysis
- ❑ Future Work and Conclusions

- Interns/PhD: UT, TUD, RUG and French schools (Paris, Nantes, Strasbourg, Toulouse): **2009**-
- Prof. Geert Leus, TU Delft: **2012**-
- Prof. Ioannis Kyriakides, University of Nicosia, Cyprus: 2012-2014
- Han Lun YAP, DSO National Labs, Singapore: 2013-2017
- Prof. Giampiero Gerini, TU/e and TNO Delft: **2018**-

THALES

Motivation: Information-based processing (CS and IG) in Radar



measurements $\boldsymbol{y}$ = sensing-model $\boldsymbol{A}(\boldsymbol{\theta})$ profile $\boldsymbol{x}$ + receiver-noise $\boldsymbol{z}$, $z \sim CN\,(0, \gamma I)$, i.e. complex Gaussian $p(\boldsymbol{y}|\boldsymbol{\theta})$

Data sizes **are growing**, e.g. with *higher* resolution in range, Doppler and angles

large data size & low information density! ⇒ ***CS*** *and* ***IG*** *in radar*

Compressive Sensing: Radar Performance

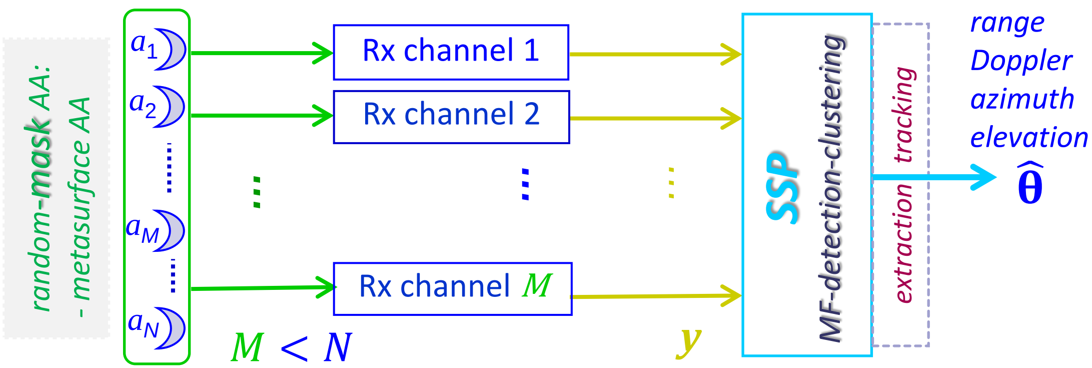
CS foundations $By = BAx$

$B \in \mathbb{C}^{M\text{x}N}$, $M < N$ $y = Ax$

✓ incoherence of A : $\mu(A)$, RIP or NSP, e.g. low $\mu(A)$, $\mu(A) = \max_{i,j,i \neq j} |a_i^H a_j|$

✓ sparsity of x, $K = \dim(T)$, $K < M \leq N$, T ... true support set

compressive-data acquisition (CDA) *sparse-signal processing* (SSP)



radar profile x

l1-norm optimization

$$x_{\text{SSP}} = \arg\min_{x} \|y - Ax\|^2 + h\|x\|_1$$

measurements y = sensing-model $A(\theta)$ profile x + receiver-noise z , $z \sim CN(0, \gamma I)$

CS/SSP benefits: higher resolution, multi-target analysis, fewer data/channels, ...

CS/SSP performance in *processing gain, detection, resolution and accuracy*?

CS: Reduced **complexity** while improving, or at least maintaining, **performance**

THALES

Information Geometry : Radar Performance

Information Geometry (IG) is stochastic SP where the intrinsic geometrical structure of a data model is characterized locally by the **Fisher information matrix** (FIM).

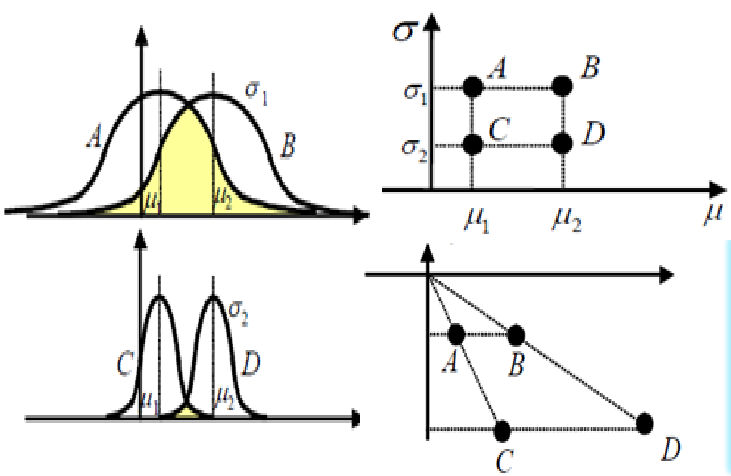
Fisher-Rao metric ds: $ds^2 = d\boldsymbol{\theta}^{\mathrm{H}} \boldsymbol{J}_{\theta} d\boldsymbol{\theta}$, where $\boldsymbol{J}_{\theta}$ is FIM and $d\boldsymbol{\theta}$ infinitesimal change of $\boldsymbol{\theta}$

Information Distance (ID) between pdf-s:

resolution: $p(\boldsymbol{y}|\boldsymbol{\theta})$ and $p(\boldsymbol{y}|\boldsymbol{\theta}+\delta\boldsymbol{\theta})$ whose $\boldsymbol{\theta}$ differ by a small $\delta\boldsymbol{\theta}$ (of close targets)

detection: $p(\boldsymbol{y}|\mathbf{0})$ and $p(\boldsymbol{y}|\boldsymbol{\mu}(\boldsymbol{\theta}))$ whose $\boldsymbol{y}$ is without or with signals $\boldsymbol{\mu}(\boldsymbol{\theta})$, $\boldsymbol{\mu}(\boldsymbol{\theta})$=$E[\boldsymbol{y}]$

Bayesian variational inferences: true $p(\boldsymbol{\theta}|\boldsymbol{y})$ and latent$\boldsymbol{\epsilon}$-variational $q(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{\epsilon})$ posteriors



$\mathrm{MSE}(\hat{\theta}) \geq \boldsymbol{J}_{\theta}{}^{-1}$

accuracy

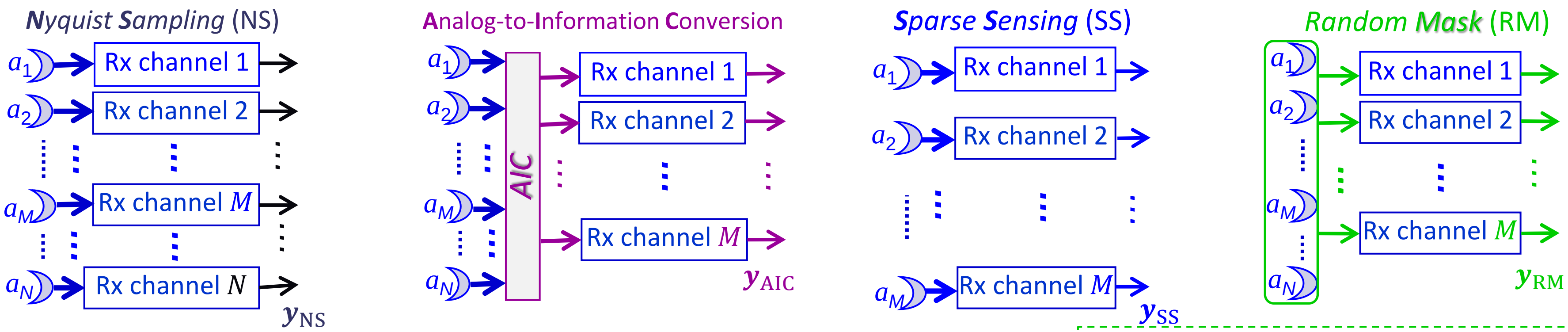
FIM $\boldsymbol{J}_{\theta} \equiv -\mathrm{E}\left[\frac{\partial^2 \ln p(\boldsymbol{y}|\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{H}}\right]$

IG: information distances
resolution
detection
Bayesian variational inferences

machine/deep learning
to be combined with **CS-IG:**

- natural gradient most efficient
- Fisher-Rao metric ⇒ *IDs*
- Bayesian variational inferences

Performance Analysis: *Compressive Data Acquisition*



MSc project (proposed by R.Pribić and G.Gerini, 2018)
Metasurfaces for CS on Radar Array Antenna Systems

$$\boldsymbol{y}_{\mathrm{NS}} = \boldsymbol{a}\alpha + \boldsymbol{z}$$

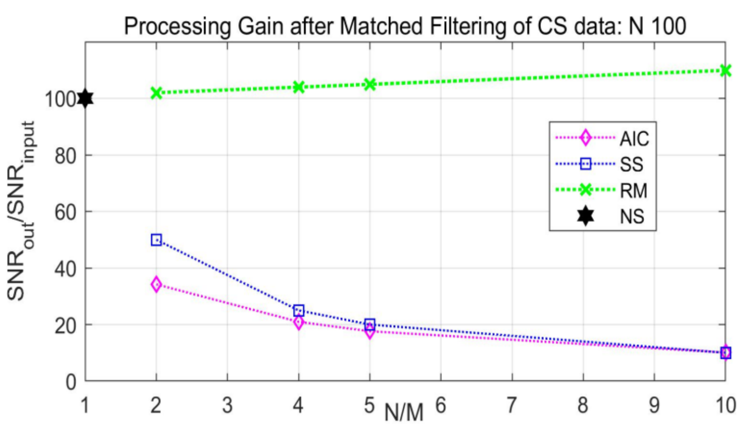
$$\boldsymbol{y}_{\mathrm{AIC}} = \boldsymbol{B}(\boldsymbol{a}\alpha + \boldsymbol{z})$$

$$\boldsymbol{y}_{\mathrm{SS}} = \boldsymbol{a}_{\mathrm{SS}}\alpha + \boldsymbol{z}_{\mathrm{SS}}$$

$$\boldsymbol{y}_{\mathrm{RM}} = \boldsymbol{B}\boldsymbol{a}\alpha + \boldsymbol{z}_{\mathrm{RM}}$$

SNR after MF, $\boldsymbol{B}$ = $\boldsymbol{B}$:

$$\alpha_{*,\mathrm{MF}} = \boldsymbol{a}_*^{\mathrm{H}}\boldsymbol{y}_*$$



CS: Reduced complexity while improving, or at least maintaining, performance

$$\mathrm{PG}_{\mathrm{MF,NS}} = N$$

$$\mathrm{PG}_{\mathrm{MF,RM}} = N\left(1 + \frac{N-1}{MN}\right) \approx N$$

$$\mathrm{PG}_{\mathrm{MF,SS}} = M$$

$$\mathrm{PG}_{\mathrm{MF,AIC}} = M\left(\frac{1}{M} + \frac{N-1}{M+N-1}\right) \leq M$$

R. Pribić, G. Leus and C. Tzotzadinis, "*SNR Analysis of Compressive Data Acquisition*", IEEE Workshop SSP 2018.

Performance Analysis : *Sparse-Signal Processing (SSP)*

Existing (radar) processing

Existing SSP theory: $\boldsymbol{x}_{\mathrm{SSP}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + h\|\boldsymbol{x}\|_1$

point-spread function (PSF) ⟹ ✓ sensing coherence: low $\mu(\boldsymbol{A}) = \max_{i,j,i\neq j} |\boldsymbol{a}_i^{\mathrm{H}}\boldsymbol{a}_j|$, $\|\boldsymbol{a}_n\| = 1$

only a few targets in a scene ⟹ ✓ sparsity of $\boldsymbol{x}$, $K = \dim(\boldsymbol{T}), K < M \leq N$, $\boldsymbol{T}$ … true support set

matched-filtered residuals ⟹ ✓ *SSP feasibility for $\boldsymbol{x}_{\mathrm{SSP}}$ in an estimated set $\boldsymbol{S}$:* $|u_{\mathrm{SSP},n}| \leq 1$, $n \in \boldsymbol{S}$

subgradient $u_{\mathrm{SSP},n}$ indicates nonzeros: $u_{\mathrm{SSP},n} = \boldsymbol{a}_n^{\mathrm{H}}(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_{\mathrm{SSP}})/h$

Neyman-Pearson detection ⟹ ✓ $h \equiv \eta_{\mathrm{GLRT}}$ *if noise only* (no targets or $\mu(\boldsymbol{A}) = 0$)

$\boxed{\textit{SSP is sparse model-based detection-driven refinement of MF}}$ ⟹ *masking*, *resolution*, *multi-target*

Stochastic *(Bayesian) SSP*

radar processing is ***stochastic***, e.g.

Woodward, P. M. (1953) *Probability and information theory, with applications to radar, Pergamon.*

$\boldsymbol{y} = \boldsymbol{A}\underline{\boldsymbol{x}} + \boldsymbol{z}$ $\xleftarrow{\underline{x}\ \sim\ \mathit{sparse}()}$ $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{z}$

Hubert J. Flisijn, *Implementation of (Bayesian) CS in Radar Systems.* MSc thesis at Thales NL Delft, University of Twente, 2011.

I.Kyriakides and R.Pribić , *Bayesian CS using Monte Carlo Methods*. Eusipco 2013.

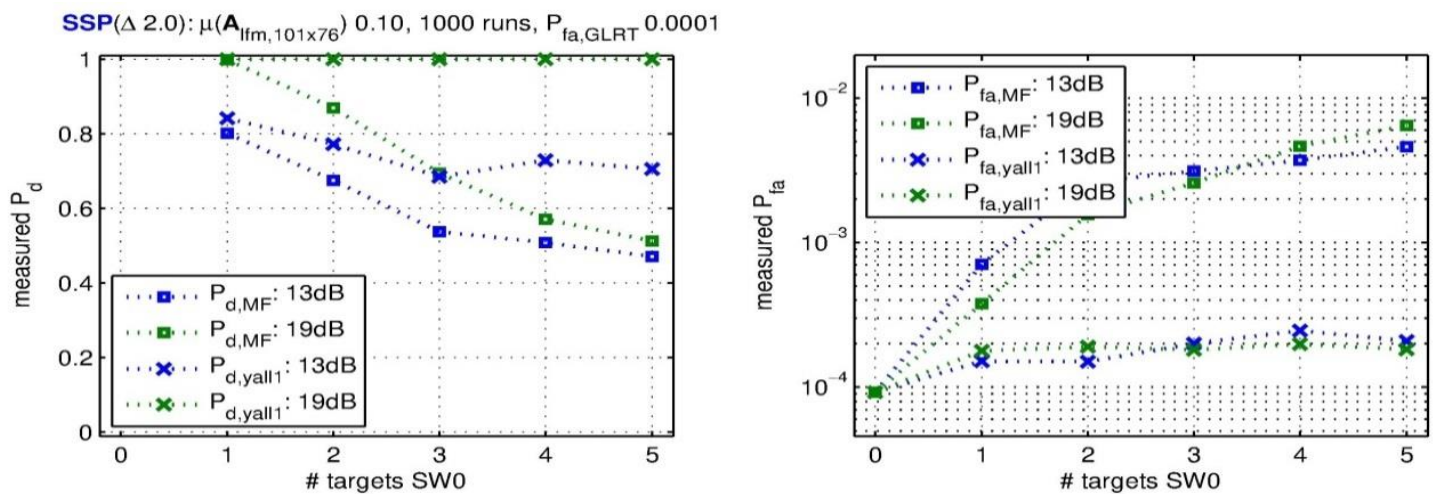
THALES

Multi-target Performance Analysis: SSP *Metrics*

detection: $P_{\mathrm{fa,SSP}}$ and $P_{\mathrm{d,SSP}}$ at h**?**

$P_{\mathrm{fa,SSP}} = \mathrm{P}\{|x_{\mathrm{SSP},l}| \neq 0\},\ l \notin \boldsymbol{T}$ i.e. nonzero in a cell with no target

$P_{\mathrm{d,SSP}} = \mathrm{P}\{|x_{\mathrm{SSP},k}| \neq 0\},\ k \in \boldsymbol{T}$ i.e. nonzero in a cell with a target



R Pribić and HL Yap, "False Alarms in Radar Detection within Sparse-signal Processing", IEEE Workshop CoSeRa 2016.

resolution: $P_{\mathrm{res,SSP}}$ at $\Delta\theta$, SNR and h**?**

$$P_{\mathrm{res,SSP}} = \mathrm{P}\{(x_{\mathrm{SSP},i} \neq 0) \wedge (x_{\mathrm{SSP},j} \neq 0)\},\ i \neq j, i, j \in \boldsymbol{T}$$

accuracy: bounds (CRLB) and mean squared error (MSE) at target separation $\Delta\theta$, K, SNR, and h**?**

Performance Analysis: CDA and Multi-target *Accuracy*

data $\boldsymbol{y} \sim CN(\alpha \boldsymbol{a}(\theta), \gamma \boldsymbol{I})$: $\quad \boldsymbol{y} = \alpha e^{j\boldsymbol{\beta}\theta} + \boldsymbol{z}, \qquad D_\beta = \max_n \beta_n - \min_n \beta_n$

$$J_{\theta,\mathrm{NS}} = \frac{2|\alpha|^2}{\gamma} \left\| \frac{\partial \boldsymbol{a}(\theta)}{\partial \theta} \right\|^2 = \frac{2|\alpha|^2}{\gamma} \|\boldsymbol{\beta}\|^2 \to \frac{|\alpha|^2}{\gamma} \frac{D_\beta{}^2}{6}$$

$$J_{\theta,\mathrm{RM}} = \mathrm{E}\left[\frac{2|\alpha|^2}{\gamma} \|\boldsymbol{B}_{\mathrm{RM}}\boldsymbol{\beta}\|^2\right] = J_{\theta,\mathrm{NS}}$$

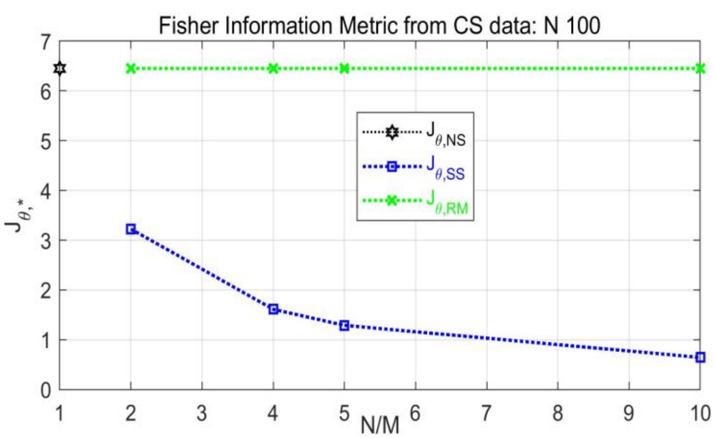
$$J_{\theta,\mathrm{SS}} = \mathrm{E}\left[\frac{2|\alpha|^2}{\gamma} \|\boldsymbol{B}_{\mathrm{SS}}\boldsymbol{\beta}\|^2\right] = \frac{M}{N} J_{\theta,\mathrm{NS}}$$

Multi-target $\boldsymbol{y} = \alpha e^{j\boldsymbol{\beta}\theta_1} + \alpha e^{j\boldsymbol{\beta}\theta_2} + \boldsymbol{z}, \; \delta\theta = \theta_2 - \theta_1$

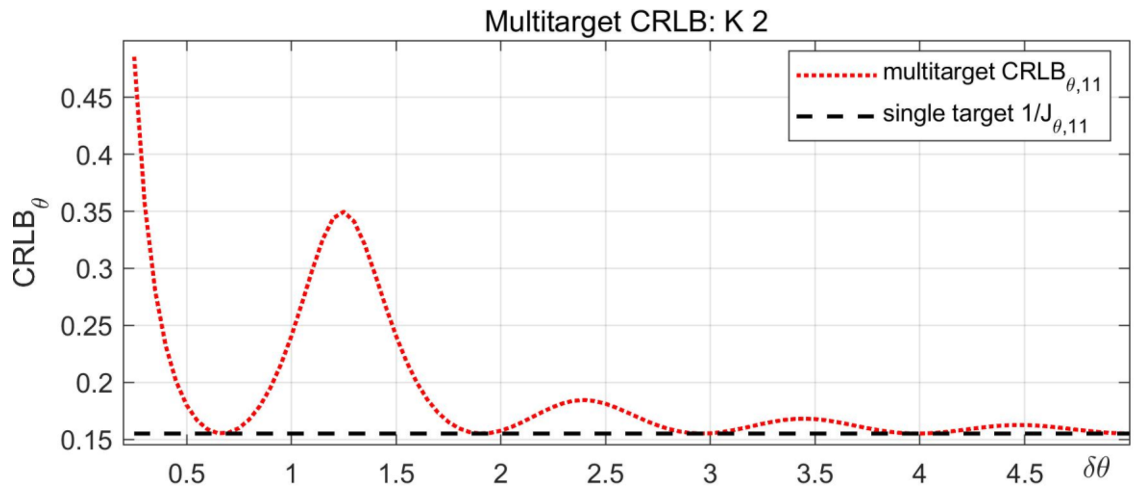
$$\mathrm{CRLB}\left(\begin{bmatrix}\theta_1 \\ \theta_2\end{bmatrix}\right) = \boldsymbol{J_\theta}^{-1} = \frac{\gamma}{|\alpha|^2}\begin{bmatrix} J_{1,1} & J_{2,1} \\ J_{1,2} & J_{2,2}\end{bmatrix}^{-1} = \frac{\gamma/|\alpha|^2}{J_{1,1}J_{2,2} - J_{1,2}J_{2,1}}\begin{bmatrix} J_{2,2} & -J_{2,1} \\ -J_{1,2} & J_{1,1}\end{bmatrix}$$

$$J_{2,1} = 2\mathrm{Re}\left\{\frac{\partial \boldsymbol{a}^H(\boldsymbol{\theta})}{\partial \theta_2}\frac{\partial \boldsymbol{a}(\boldsymbol{\theta})}{\partial \theta_1}\right\} = J_{1,2} \to \frac{2}{D_\beta \delta\theta}\left[\left(\frac{D_\beta{}^2}{2} - \frac{4}{\delta\theta^2}\right)\sin\frac{D_\beta \delta\theta}{2} + \frac{2D_\beta}{\delta\theta}\cos\frac{D_\beta\delta\theta}{2}\right]$$

$$J_{1,1} = J_{2,2} = D_\beta{}^2/6$$



$$\mathrm{MSE}(\hat{\theta}) \geq \mathrm{CRLB}(\theta) = J_\theta{}^{-1}$$



Multi-target CRLB realistic at smaller separations $\delta\theta$

R. Pribić, "Information-based Analysis of Compressive Data Acquisition", (accepted) IEEE Radar 2019

Performance Analysis: *Resolution Metrics*

data $\boldsymbol{y} \sim CN(\boldsymbol{\mu}(\theta), \gamma \boldsymbol{I})$: $\boldsymbol{y} = \alpha e^{j\boldsymbol{\beta}\theta} + \boldsymbol{z}$, $\mathrm{D}_\beta = \max_n \beta_n - \min_n \beta_n$, $\mathrm{SNR} = |\alpha|^2/\gamma$

- *Deterministic:* ***Rayleigh distance*** $\theta_{res} \propto \frac{1}{\mathrm{D}_\beta}$ (only by *array sensing* bandwidth D_β)
- ***Stochastic:*** targets also involved, i.e. their **SNR** and **separation** $\delta\theta$
  - ***Estimation approach:*** $\theta_{res} \propto \frac{1}{\mathrm{D}_\beta \sqrt{\mathrm{SNR}}}$, i.e. best accuracy (CRLB) of estimated *single-target* parameter θ with given SNR and *array configuration*
  - ***Detection approach:*** $\mathrm{P}_{res} \propto f\left(\mathrm{D}_\beta, \mathrm{SNR}, \delta\theta\right) \equiv$ probability of resolving targets at given **SNR** and separation $\delta\theta$ with given *array configuration*
    - ✓ Novel: **Information-geometry (IG) distances for a complete** P_{res}

IG: ***Information distances*** between pdf-s: $CN(\boldsymbol{\mu}(\theta), \gamma \boldsymbol{I})$ and $CN(\boldsymbol{\mu}(\theta + \delta\theta), \gamma \boldsymbol{I})$

R. Pribić and G. Leus, “Information Distances in Radar Resolution Analysis”, IEEE Workshop CAMSAP 2017
R. Pribić, “Information Distances in Resolution Analysis”, SEE Workshop GSI 2017

Performance Analysis: *Proposed* ***Stochastic Resolution Analysis*** (GSI'17 and CAMSAP'17)

Resolution test: $H_0\!: \; \boldsymbol{\delta\theta} = \mathbf{0}$ and $H_1\!: \; \boldsymbol{\delta\theta} \neq \mathbf{0}$ (Rao, 1945)

Novel: via LR (likelihood ratio) and information distances $d_{\boldsymbol{\mu}(\theta)}$, $d_{\boldsymbol{\mu}(\theta)} = \sqrt{\delta\boldsymbol{\mu}^H \boldsymbol{J}_{\boldsymbol{\mu}} \delta\boldsymbol{\mu}}$ between $\mathcal{CN}\,(\boldsymbol{\mu}(\theta), \gamma\boldsymbol{I})$ and $\mathcal{CN}\,(\boldsymbol{\mu}(\theta + \delta\theta), \gamma\boldsymbol{I})$ which θ differs by $\delta\theta$ (two close targets)

$$H_0\!: \; \boldsymbol{y} = \boldsymbol{\mu}(\theta) + \boldsymbol{\mu}(\theta) + \boldsymbol{z} = \; \boldsymbol{y}_0$$
$$H_1\!: \; \boldsymbol{y} = \boldsymbol{\mu}(\theta) + \boldsymbol{\mu}(\theta + \delta\theta) + \boldsymbol{z} = \; \boldsymbol{y}_0 + \; \delta\boldsymbol{\mu}$$

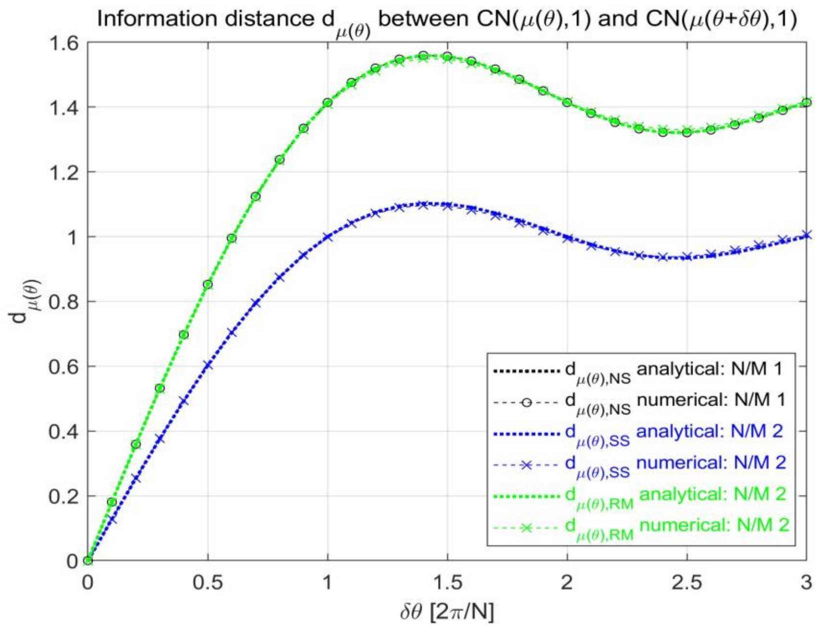
$$\gamma \ln \mathrm{LR} = 2Re\{[\boldsymbol{y} - 2\boldsymbol{\mu}(\theta)]^H \delta\boldsymbol{\mu}\} - \|\delta\boldsymbol{\mu}\|^2 \quad \Rightarrow \xi_{\mathrm{LR},d_{\boldsymbol{\mu}(\theta)}} \sim N(d_{\boldsymbol{\mu}(\theta)}, 1)$$

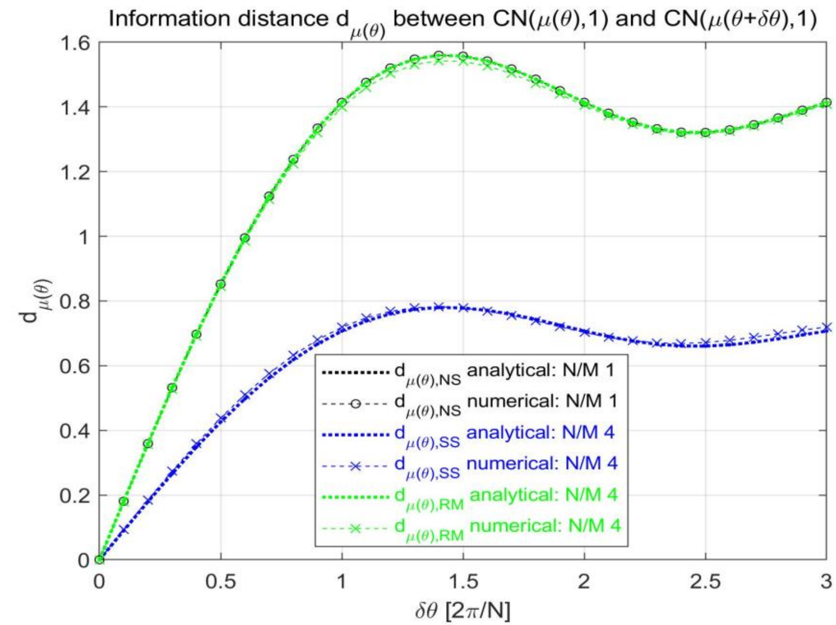
✓ **Novel**: LR distribution found and *linked to information distance* $d_{\boldsymbol{\mu}(\theta)}$

$$P_{\mathrm{res},d_{\boldsymbol{\mu}(\theta)}} = \mathrm{P}\left\{\xi_{\mathrm{LR},d_{\boldsymbol{\mu}(\theta)}} > \rho \mid H_1\right\}, \quad \text{where} \quad \rho = N^{-1}(0, 1, P_{\mathrm{fa}})$$

Stochastic Resolution Analysis: CDA *Information Distances*

$$d_{\boldsymbol{\mu}(\theta),\mathrm{NS}} = \|\delta\boldsymbol{\mu}\|/\sqrt{\gamma} \rightarrow \sqrt{\mathrm{D}_\beta \mathrm{SNR}\left(1-\frac{\sin \mathrm{D}_\beta\delta\theta/2}{\mathrm{D}_\beta\delta\theta/2}\right)} \qquad d_{\boldsymbol{\mu}(\theta),\mathrm{SS}} = \sqrt{M/N}\, d_{\boldsymbol{\mu}(\theta),\mathrm{NS}} \qquad d_{\boldsymbol{\mu}(\theta),\mathrm{RM}} = d_{\boldsymbol{\mu}(\theta),\mathrm{NS}}$$



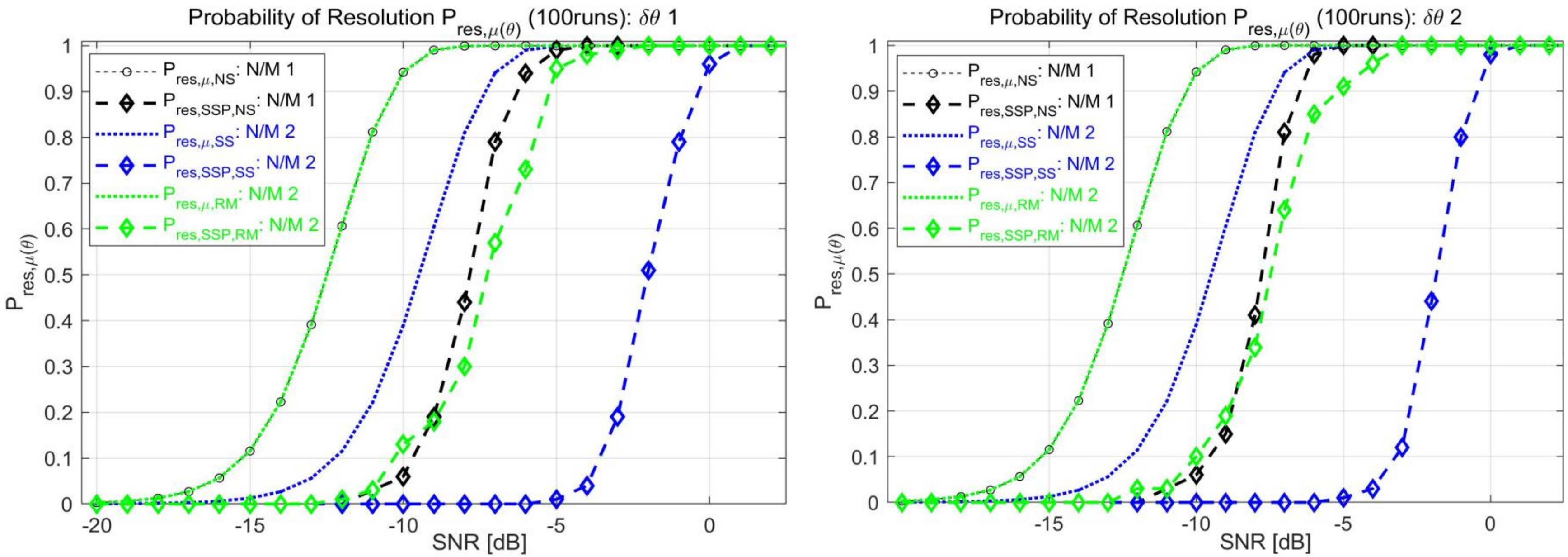


- ✓ Compression at reception (signal only, e.g. random masking, RM) preserves $d_{\boldsymbol{\mu}(\theta),\mathrm{NS}}$
- ✓ Otherwise, e.g. compression before reception (sparse sensing, SS) harms information distances

R. Pribić, “Resolution Analysis of Compressive Data Acquisition”, Workshop CoSeRa 2018

Stochastic Resolution Analysis: CDA *Resolution Probability*

bounds $P_{\mathrm{res},\boldsymbol{\mu},*}$ and $P_{\mathrm{res,SSP}} = \mathrm{P}\{(x_{\mathrm{SSP},i} \neq 0) \wedge (x_{\mathrm{SSP},j} \neq 0) | H_1\}$, $i \neq j$, in two target cells i and j



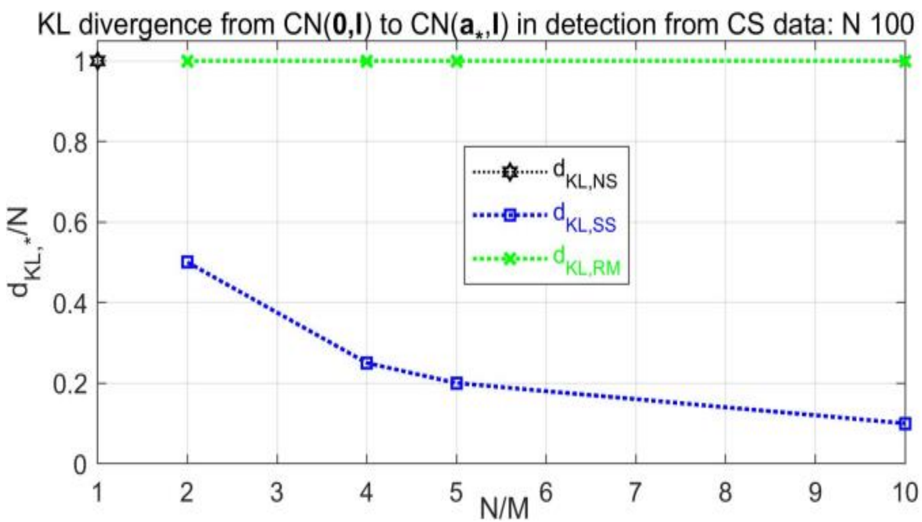
- ✓ Lower $P_{\mathrm{res},*,\mathrm{SS}}$ while $P_{\mathrm{res},*,\mathrm{RM}}$ comparable with $P_{\mathrm{res},*,\mathrm{NS}}$ when $M < N$
- ✓ SSP resolution $P_{\mathrm{res,SSP},*}$ far (4dB or more with SS) from the bounds given by the IG-based probability $P_{\mathrm{res},\mu,*}$
- ✓ $P_{\mathrm{res},\mu}$ and $P_{\mathrm{res,SSP}}$ remain stable and realistic at larger separation (as $d_{\boldsymbol{\mu}(\theta)}$)

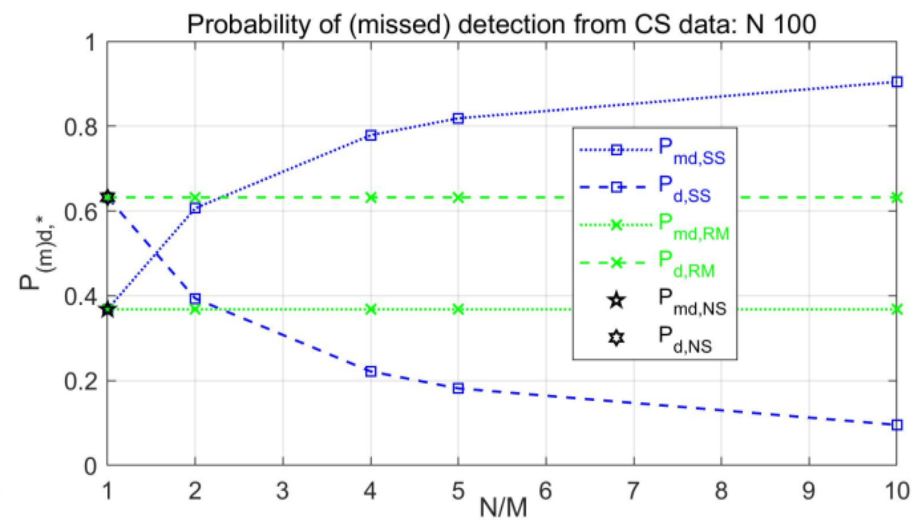
R. Pribić, "Resolution Analysis of Compressive Data Acquisition", Workshop CoSeRa 2018

Performance Analysis CDA: *Detection*

Chernoff-Stein lemma: $P_{\mathrm{md},*} \propto \exp\left(-d_{\mathrm{KL},*}\right), P_{\mathrm{md},*} + P_{\mathrm{d},*} = 1$

d_{KL} Kullback-Leibler divergence from $p(\boldsymbol{y}|\mathbf{0})$ to $p(\boldsymbol{y}|\boldsymbol{\mu}(\boldsymbol{\theta}))$





Smaller $d_{\mathrm{KL,ss}}$ while $d_{\mathrm{KL,RM}}$ preserves $d_{\mathrm{KL,NS}}$ if $M < N$

Lower $P_{\mathrm{d,ss}}$ while $P_{\mathrm{d,RM}}$ preserves $P_{\mathrm{d,NS}}$ if $M < N$

Higher $P_{\mathrm{md,ss}}$ while $P_{\mathrm{md,RM}}$ preserves $P_{\mathrm{md,NS}}$

R. Pribić, "Information-based Analysis of Compressive Data Acquisition", (accepted) IEEE Radar 2019

Further work: SSP Performance

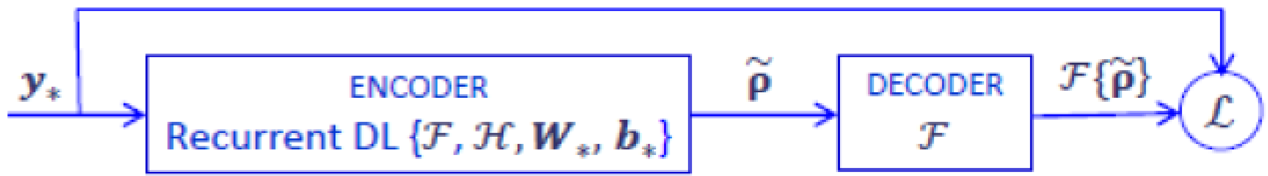
CS/SSP: sensing model $\boldsymbol{A}(\boldsymbol{\theta})$ essential in deconvolution but often not fully known in practice, …

Stochastic Deep Learning

complex-valued *Deep Learning* combined with the *stochastic* approach: IG and Bayesian methods ⇐

refining $\boldsymbol{A}(\boldsymbol{\theta})$ from data with known contents
estimating unknowns $[\alpha\ \theta]$ with refined $\boldsymbol{A}(\boldsymbol{\theta})$

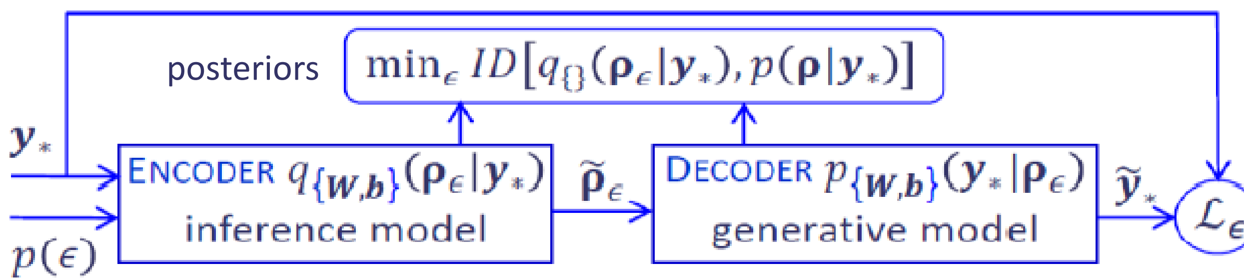
Deep Learning:
autoencoder



$$\boldsymbol{y}_l = \psi\left(\boldsymbol{W}_{l-1}^H \boldsymbol{y}_{l-1} + \boldsymbol{b}_{l-1}\right)$$

$$\tilde{\boldsymbol{\rho}} = \mathcal{H}\{\boldsymbol{y}_*\} \qquad \boldsymbol{y}_* = \mathcal{F}\{\boldsymbol{\rho}\} + \boldsymbol{z}_* \qquad \boldsymbol{\rho} = [\alpha\ \theta]$$

Stochastic Deep Learning:
(Wasserstein) variational
autoencoder



R. Pribić, “Stochastic Deep Learning in CS Radar ”, submitted to SEE Radar 2019
Thales NL Internship (T. Magalas, INP-ENSEEIHT Toulouse) “ Machine Deep Learning (MDL) with CS and IG”, March-September 2019.
Thales NL Internship (L. Isselin, Univ. Strasbourg) “ Links of MDL with CS and IG”, June-August 2018.

THALES

Conclusions and Future Work

CS and IG in information-based (Radar) Processing of complex-Gaussian measurements

- CS/SSP performance illustrated by processing gain , accuracy, resolution and detection, and
- assessed with IG tools for fewer measurements: compression after (AIC), before (SS) and at (RM) reception

The proposed information-based performance analysis of CS: *CDA* and *SSP*, shows:

- ✓ completeness of the (radar-) essential performance metrics
- ✓ close links between CS and IG due to the emphasis on information content in data
- ✓ clear preference to compression at reception (signal only!), e.g. with *random masking* (RM)
- ✓ otherwise, with SS or AIC, radar performance heavily sacrificed
- ✓ close ties between detection, accuracy and resolution (at small separations)

Further work

- ➢ demonstrator of RM in a metasurface antenna array (together with prof. G. Gerini at TU/e and TNO)
- ➢ information-based analysis of CS/SSP performance with *multiple parameters and multiple targets*
- ➢ analysis in continuous domain to determine the reference before any discretisation
- ➢ stochastic deep learning for more accurate knowledge of sensing models
- ➢ implementation of CS in an actual radar system: *CDA* together with *SSP in a ThalesNL radar system!*

Questions?

THALES