A General Method for Calculating Complex-Valued and Quaternionic Minimum Mean-Squared Error Estimators to be Used in (Vector) Approximate Message Passing

Robert F.H. Fischer **Elena Sterk**

NACHRICHTENTECHNIK Universität Ulm

universität ulm

Introduction

Problem Statement:

- observation model

measurement matrix $\in \mathbb{R}^{M \times N}$ — zero-mean Gaussian random vector $\in \mathbb{R}^{M}$, variance σ_n^2, independent from $\boldsymbol{x}$

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$$

measurement vector $\in \mathbb{R}^{M}$ — signal vector $\in \mathbb{R}^{N}$, i.i.d. elements, marginal pdf $f_{\mathsf{x}}(x)$

- task: given $\boldsymbol{y}$, find an estimate for $\boldsymbol{x}$
- problem: huge complexity if N and M are large

Solution:

- resort to suboptimal approaches — here: ***message passing principle***
- in particular, we are interested in ***approximate message passing (AMP)*** and ***vector approximate message passing (VAMP)***

Notation: random variables are written in sans-serif font, e.g., x, realizations in conventional italic font, e.g., x.

Introduction (II)

A-posteriori pdf: (inference, conditional pdf)

$$\mathsf{f}_{\boldsymbol{x}|\boldsymbol{y}}(\boldsymbol{x}) = \frac{1}{\mathsf{f}_{\boldsymbol{y}}(\boldsymbol{y})} \mathsf{f}_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{y}) \, \mathsf{f}_{\boldsymbol{x}}(\boldsymbol{x})$$

Bayes' rule

$$= \text{const} \cdot \mathsf{f}_{\boldsymbol{n}}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}) \, \mathsf{f}_{\boldsymbol{x}}(\boldsymbol{x})$$

additive noise,
independent of the data

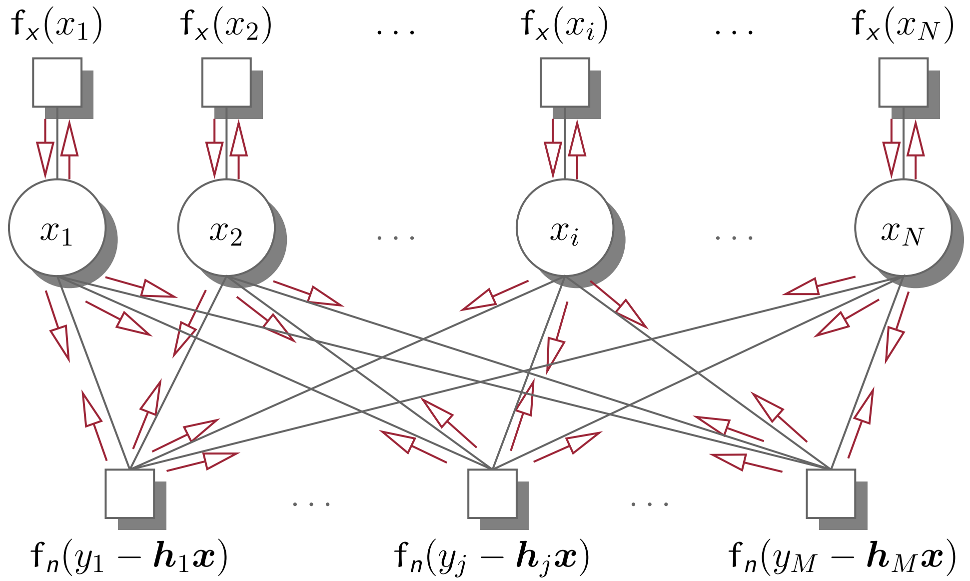
with $\boldsymbol{H} = \begin{bmatrix} \boldsymbol{h}_1^\mathsf{T} \\ \vdots \\ \boldsymbol{h}_M^\mathsf{T} \end{bmatrix}$

$$= \text{const} \cdot \prod_{j=1}^{M} \mathsf{f}_{n}(y_j - \boldsymbol{h}_j^\mathsf{T}\boldsymbol{x}) \cdot \prod_{i=1}^{N} \mathsf{f}_{x}(x_i)$$

i.i.d. data, i.i.d. noise

Factor Graph and Message Passing: [KFL'01], [LDH'07]

- *edge-dependent* messages are exchanged



Introduction (II)

A-posteriori pdf: (inference, conditional pdf)

$$\mathsf{f}_{\boldsymbol{x}|\boldsymbol{y}}(\boldsymbol{x}) = \frac{1}{\mathsf{f}_{\boldsymbol{y}}(\boldsymbol{y})} \mathsf{f}_{\boldsymbol{y}|\boldsymbol{x}}(\boldsymbol{y})\, \mathsf{f}_{\boldsymbol{x}}(\boldsymbol{x})$$

Bayes' rule

$$= \text{const} \cdot \mathsf{f}_{\boldsymbol{n}}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x})\, \mathsf{f}_{\boldsymbol{x}}(\boldsymbol{x})$$

additive noise, independent of the data

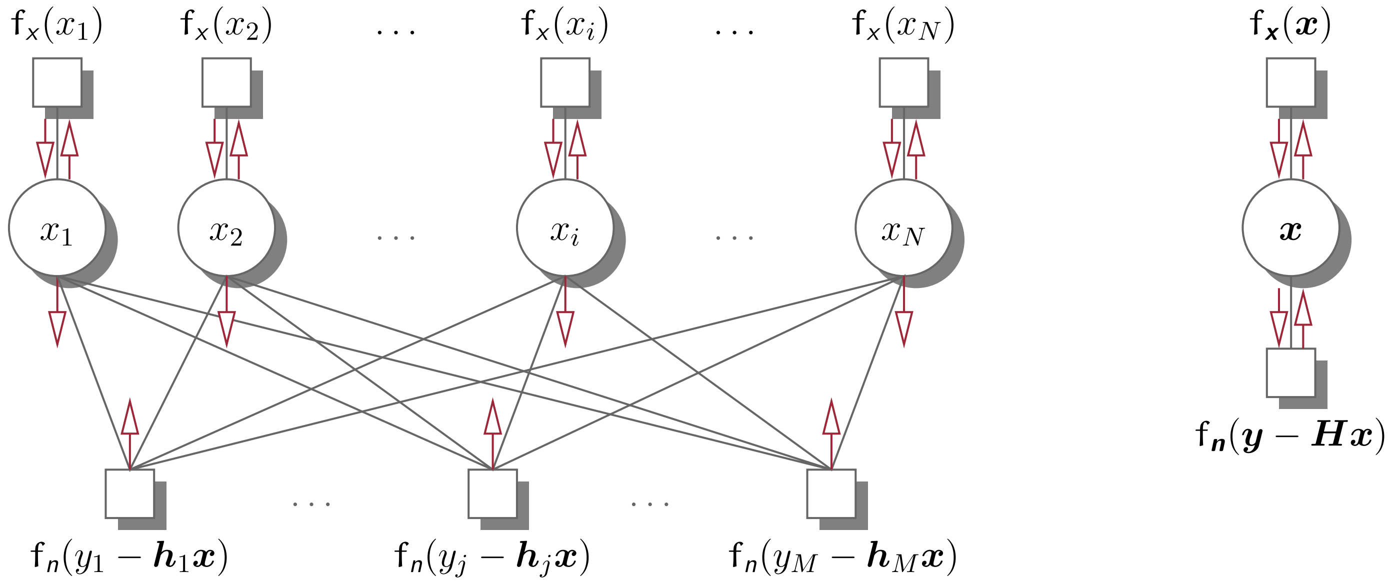
with $\boldsymbol{H} = \begin{bmatrix} \boldsymbol{h}_1^\mathsf{T} \\ \vdots \\ \boldsymbol{h}_M^\mathsf{T} \end{bmatrix}$

$$= \text{const} \cdot \prod_{j=1}^{M} \mathsf{f}_{n}(y_j - \boldsymbol{h}_j^\mathsf{T}\boldsymbol{x}) \cdot \prod_{i=1}^{N} \mathsf{f}_{x}(x_i)$$

i.i.d. data, i.i.d. noise

Factor Graph and Message Passing: [KFL'01], [LDH'07]

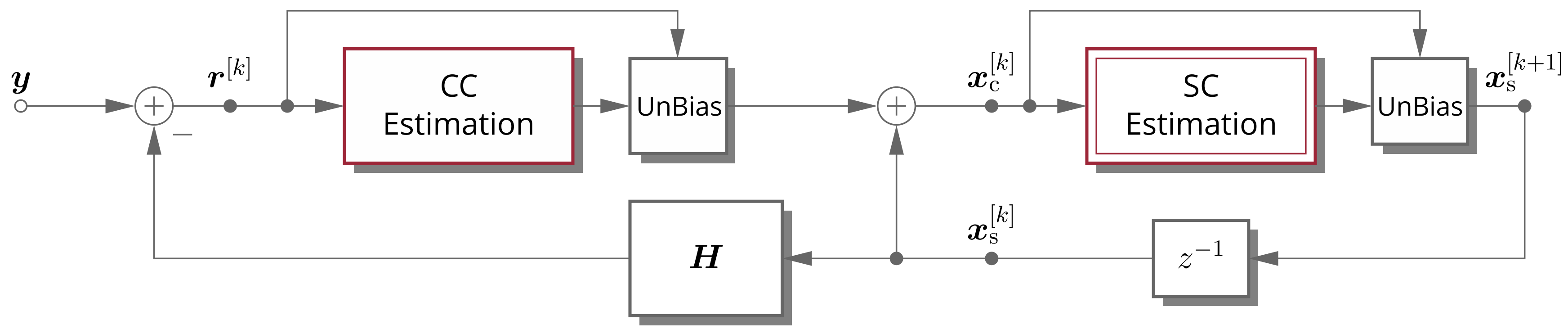
- *edge-dependent* messages are exchanged
- approximation: *state-dependent* messages are exchanged



Introduction (III)

Vector Approximate Message Passing (VAMP): (or Orthogonal AMP (OAMP)) [RSF'19], [MP'17]

- two estimation problems
  - "channel-constrained (CC)" — consider $\boldsymbol{H}$, ignore $\mathrm{f}_x(x)$
  - "signal-constrained (SC)" — consider $\mathrm{f}_x(x)$, ignore $\boldsymbol{H}$
- *bias* has to be removed / calculation of ***extrinsics***



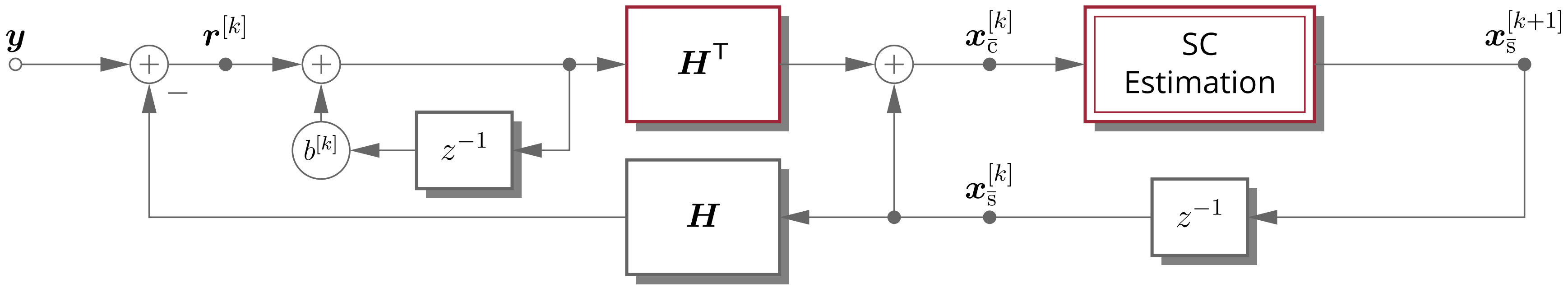
Minimum Mean-Squared Error (MMSE) Estimation:

- given the observation y, find the estimate $\hat{x}$ such that $\mathrm{E}\{(x-\hat{x})^2\} \to \min$
- solution: ***conditional mean estimator*** $\hat{x} = \mathrm{E}\{x \mid y\}$

Introduction (III)

Approximate Message Passing (AMP): [DMM'10]

- simplification
  - matched filter (MF) and
  - "signal-constrained (SC)" MMSE estimation (BAMP) or soft thresholding ("denoising")
- Onsager correction



Introduction (IV)

Usual Setting:

- the algorithms are designed for ***real-valued*** signals

Complex Signals:

- most signals in communications / radar signal processing are I/Q signals
- some extensions to ***complex-valued*** signals exist

[AMO'13], [SF'25]

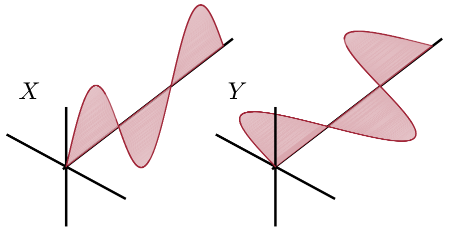
Hypercomplex Signal Processing:

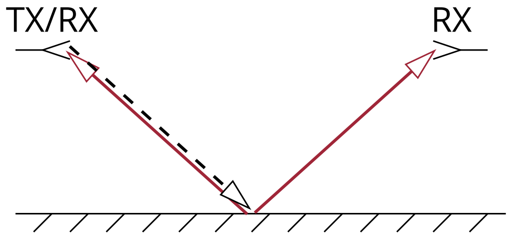
- has received increasing interest over the last years

  cf. the two special issues of the *IEEE Signal Processing Magazine*

[SPMag'24]

- four-dimensional signal processing employing ***quaternions*** is very popular whenever two complex signals are treated jointly
- most prominent examples:
  - fiber-optical communications
  - bistatic radar





[KA'16]

[GMW'23]

Introduction (V)

Calculating MMSE Solutions:

- let the a-posteriori pdf $\mathrm{f}_{\boldsymbol{x}|\boldsymbol{\lambda}}(\boldsymbol{x})$ be member of an ***exponential family***, i.e., [Efr'23]

natural parameters; $\boldsymbol{\lambda} = \mathrm{fct}(\boldsymbol{y})$ — sufficient statistics

$$\mathrm{f}_{\boldsymbol{x}|\boldsymbol{\lambda}}(\boldsymbol{x}) = h(\boldsymbol{x}) \cdot \mathrm{e}^{\boldsymbol{\lambda}^\mathsf{T} \boldsymbol{g}(\boldsymbol{x}) - A(\boldsymbol{\lambda})}$$

carrying density — log-partition function

- well-known result for ***real-valued signals*** $(\boldsymbol{e} \stackrel{\text{def}}{=} \boldsymbol{x} - \mathrm{E}\{\boldsymbol{x} \mid \boldsymbol{\lambda}\})$ [Efr'23]

$$\mathrm{E}\{\boldsymbol{x} \mid \boldsymbol{\lambda}\} = \frac{\partial}{\partial \boldsymbol{\lambda}} A(\boldsymbol{\lambda}) = \left[\frac{\partial}{\partial \lambda_i} A(\boldsymbol{\lambda}) \right]$$

$$\boldsymbol{\Phi}_{\boldsymbol{ee}} = \mathrm{E}\{\boldsymbol{e}\boldsymbol{e}^\mathsf{T}\} = \left[\frac{\partial}{\partial \lambda_i} \frac{\partial}{\partial \lambda_j} A(\boldsymbol{\lambda}) \right]$$

Generalization / Question:

Do similar expressions hold for complex-valued and quaternionic signals?

Complex Numbers

Complex Numbers:

- set $\mathbb{C}$ of numbers of the form

$$z = a + b\mathrm{i}$$

	1	i
1	1	i
i	i	-1

with
  - components $a, b \in \mathbb{R}$
  - *imaginary unit* i
  - basic relation $\mathrm{i}^2 = -1$

- most algebraic operations can be carried out like those for real numbers (e.g., addition is done component-wise)

- the *conjugate* is defined as

$$z^* \overset{\text{def}}{=} a - b\mathrm{i}$$

Quaternions:

[CS'03], [Zha'97]

- set $\mathbb{H}$ of numbers of the form

$$q = a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k}$$

	1	i	j	k
1	1	i	j	k
i	i	−1	k	−j
j	j	−k	−1	i
k	k	j	−i	−1

with
  - components $a, b, c, d \in \mathbb{R}$
  - three *imaginary units* i, j, and k
  - basic relations $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1$ and $\mathrm{ijk} = -1$

- most algebraic operations can be carried out like those for complex numbers (e.g., addition is done component-wise)
- however, the multiplication of two quaternions does not commute, i.e., $q_1 q_2 \neq q_2 q_1$ in general (this results in a number of peculiarities)
- $pq = qp$ holds for all $q \in \mathbb{H}$, if and only if $p \in \mathbb{R}$
- the *conjugate* is defined as

$$q^* \overset{\text{def}}{=} a - b\mathrm{i} - c\mathrm{j} - d\mathrm{k}$$

Functions and Derivatives

Real-Valued

- real-valued function of the real-valued variable x

$$f(x)$$

Complex-Valued

- complex-valued function of the complex-valued variable $z = a + b\mathrm{i}$

$$f(z) \;=\; f(a, b)$$

with a, $b \in \mathbb{R}$

Quaternionic

- quaternionic function of the quaternionic variable $q = a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k}$

$$f(q) \;=\; f(a, b, c, d)$$

with a, b, c, $d \in \mathbb{R}$

Functions and Derivatives (II)

Derivatives:

- in a number of situations ***derivatives*** are required
  e.g., for optimization
- the derivative of real-valued functions of real-valued variables is well known

Complex-Valued Functions:

- the derivative of complex-valued functions of complex-valued variables exists if the function is ***holomorphic***
- cost functions (mean-squared error) are real-valued;
  real-valued functions of complex-valued variables are ***not*** holomorphic
- one may resort to the ***Wirtinger derivatives*** [Fis'02]

$$\frac{\partial f}{\partial z} \overset{\text{def}}{=} \frac{1}{2}\left(\frac{\partial f}{\partial a} - \frac{\partial f}{\partial b}\mathrm{i}\right)$$

$$\frac{\partial f}{\partial z^*} \overset{\text{def}}{=} \frac{1}{2}\left(\frac{\partial f}{\partial a} + \frac{\partial f}{\partial b}\mathrm{i}\right)$$

Functions and Derivatives (II)

Derivatives:

- in a number of situations ***derivatives*** are required
  e.g., for optimization
- the derivative of real-valued functions of real-valued variables is well known

Quaternionic Functions:

- the concept of Wirtinger derivatives can be generalized to quaternions
- thereby, due to non-commutativity, ***left*** and ***right*** derivatives have to be distinguished
- ***generalized (left) Wirtinger derivatives*** (w.r.t. to q and q^*) [XJT'15], [MJT'11], [XM'15]

$$\frac{\partial f}{\partial q} \overset{\text{def}}{=} \frac{1}{4}\left(\frac{\partial f}{\partial a} - \frac{\partial f}{\partial b}\mathrm{i} - \frac{\partial f}{\partial c}\mathrm{j} - \frac{\partial f}{\partial d}\mathrm{k}\right)$$

$$\frac{\partial f}{\partial q^*} \overset{\text{def}}{=} \frac{1}{4}\left(\frac{\partial f}{\partial a} + \frac{\partial f}{\partial b}\mathrm{i} + \frac{\partial f}{\partial c}\mathrm{j} + \frac{\partial f}{\partial d}\mathrm{k}\right)$$

(eight different derivatives exist; we only require these two)

Functions and Derivatives (III)

Real-Valued

for $C \in \mathbb{R}$

- $$\frac{\partial}{\partial x} Cx = C$$
- $$\frac{\partial}{\partial x} x^2 = 2x$$

Complex-Valued

for $C = C_a + C_b\mathrm{i} \in \mathbb{C}$

- $$\frac{\partial}{\partial z} Cz = C$$ $$\frac{\partial}{\partial z^*} Cz = 0$$
- $$\frac{\partial}{\partial z} (z^*C + C^*z) = C^*$$ $$\frac{\partial}{\partial z^*} (z^*C + C^*z) = C$$
- $$\frac{\partial}{\partial z} zz^* = z^*$$ $$\frac{\partial}{\partial z^*} zz^* = z$$

Quaternionic

for $C = C_a + C_b\mathrm{i} + C_c\mathrm{j} + C_d\mathrm{k} \in \mathbb{H}$

- $$\frac{\partial}{\partial q} Cq = C$$ $$\frac{\partial}{\partial q^*} Cq = -\tfrac{1}{2}C$$
- $$\frac{\partial}{\partial q} (q^*C + C^*q) = \tfrac{1}{2}C^*$$ $$\frac{\partial}{\partial q^*} (q^*C + C^*q) = \tfrac{1}{2}C$$
- $$\frac{\partial}{\partial q} qq^* = \tfrac{1}{2}q^*$$ $$\frac{\partial}{\partial q^*} qq^* = \tfrac{1}{2}q$$

Functions and Derivatives (IV)

Gradients:

- for functions $f(q_1, \dots, q_m) = f(\boldsymbol{q})$ we define the *gradients* ($\mathbb{R}$), *Wirtinger gradients* ($\mathbb{C}$), and *generalized Wirtinger gradients* ($\mathbb{H}$) as

row vector
$$\frac{\partial f(\boldsymbol{q})}{\partial \boldsymbol{q}} \overset{\text{def}}{=} \begin{bmatrix} \frac{\partial f}{\partial q_1} & \cdots & \frac{\partial f}{\partial q_m} \end{bmatrix}$$

column vector
$$\frac{\partial f(\boldsymbol{q})}{\partial \boldsymbol{q}^{\mathsf{H}}} \overset{\text{def}}{=} \begin{bmatrix} \frac{\partial f}{\partial q_1^*} \\ \vdots \\ \frac{\partial f}{\partial q_m^*} \end{bmatrix}$$

- similar for vector-valued functions $\boldsymbol{f}(\boldsymbol{q}) = [f_1(\boldsymbol{q}) \cdots f_n(\boldsymbol{q})]^{\mathsf{T}}$

- for real-valued functions, $\boldsymbol{f}(\boldsymbol{q}) \in \mathbb{R}^n$, we have

$$\frac{\partial \boldsymbol{f}(\boldsymbol{q})}{\partial \boldsymbol{q}} = \left(\frac{\partial \boldsymbol{f}(\boldsymbol{q})}{\partial \boldsymbol{q}^{\mathsf{H}}}\right)^{\mathsf{H}}$$

Functions and Derivatives (V)

Real-Valued

for $\boldsymbol{c} \in \mathbb{R}^{m\times 1}$

- $$\frac{\partial}{\partial \boldsymbol{x}} \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x} = \boldsymbol{c}^{\mathsf{T}}$$
  $$\frac{\partial}{\partial \boldsymbol{x}^{\mathsf{T}}} \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x} = \boldsymbol{c}$$

- $$\frac{\partial}{\partial \boldsymbol{x}} (\boldsymbol{x}^{\mathsf{T}}\boldsymbol{c} + \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x}) = \frac{\partial}{\partial \boldsymbol{x}} 2\boldsymbol{c}^{\mathsf{T}}\boldsymbol{x} = 2\boldsymbol{c}^{\mathsf{T}}$$

for $\boldsymbol{M} = \boldsymbol{M}^{\mathsf{T}} \in \mathbb{R}^{m\times m}$

- $$\frac{\partial}{\partial \boldsymbol{x}} \boldsymbol{x}^{\mathsf{T}}\boldsymbol{M}\boldsymbol{x} = 2\boldsymbol{x}^{\mathsf{T}}\boldsymbol{M}$$

Complex-Valued

for $\boldsymbol{c} \in \mathbb{C}^{m\times 1}$

- $$\frac{\partial}{\partial \boldsymbol{z}} \boldsymbol{c}^{\mathsf{H}}\boldsymbol{z} = \boldsymbol{c}^{\mathsf{H}}$$
  $$\frac{\partial}{\partial \boldsymbol{z}^{\mathsf{H}}} \boldsymbol{c}^{\mathsf{H}}\boldsymbol{z} = \boldsymbol{0}$$

- $$\frac{\partial}{\partial \boldsymbol{z}} (\boldsymbol{z}^{\mathsf{H}}\boldsymbol{c} + \boldsymbol{c}^{\mathsf{H}}\boldsymbol{z}) = \frac{\partial}{\partial \boldsymbol{z}} 2\mathrm{Re}\{\boldsymbol{c}^{\mathsf{H}}\boldsymbol{z}\} = \boldsymbol{c}^{\mathsf{H}}$$

for $\boldsymbol{M} = \boldsymbol{M}^{\mathsf{H}} \in \mathbb{C}^{m\times m}$

- $$\frac{\partial}{\partial \boldsymbol{z}} \boldsymbol{z}^{\mathsf{H}}\boldsymbol{M}\boldsymbol{z} = \boldsymbol{z}^{\mathsf{H}}\boldsymbol{M}$$

Quaternionic

for $\boldsymbol{c} \in \mathbb{H}^{m\times 1}$

- $$\frac{\partial}{\partial \boldsymbol{q}} \boldsymbol{c}^{\mathsf{H}}\boldsymbol{q} = \boldsymbol{c}^{\mathsf{H}}$$
  $$\frac{\partial}{\partial \boldsymbol{q}^{\mathsf{H}}} \boldsymbol{c}^{\mathsf{H}}\boldsymbol{q} = -\tfrac{1}{2}\boldsymbol{c}^{*}$$

- $$\frac{\partial}{\partial \boldsymbol{q}} (\boldsymbol{q}^{\mathsf{H}}\boldsymbol{c} + \boldsymbol{c}^{\mathsf{H}}\boldsymbol{q}) = \frac{\partial}{\partial \boldsymbol{q}} 2\mathrm{Re}\{\boldsymbol{c}^{\mathsf{H}}\boldsymbol{q}\} = \tfrac{1}{2}\boldsymbol{c}^{\mathsf{H}}$$

for $\boldsymbol{M} = \boldsymbol{M}^{\mathsf{H}} \in \mathbb{H}^{m\times m}$

- $$\frac{\partial}{\partial \boldsymbol{q}} \boldsymbol{q}^{\mathsf{H}}\boldsymbol{M}\boldsymbol{q} = \tfrac{1}{2}\boldsymbol{q}^{\mathsf{H}}\boldsymbol{M}$$

Functions and Derivatives (V)

Real-Valued

for $\boldsymbol{c} \in \mathbb{R}^{m\times 1}$

- $$\frac{\partial}{\partial \boldsymbol{x}} \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x} = \boldsymbol{c}^{\mathsf{T}}$$
  $$\frac{\partial}{\partial \boldsymbol{x}^{\mathsf{T}}} \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x} = \boldsymbol{c}$$
- $D = 1$

Complex-Valued

for $\boldsymbol{c} \in \mathbb{C}^{m\times 1}$

- $$\frac{\partial}{\partial \boldsymbol{z}} \boldsymbol{c}^{\mathsf{H}}\boldsymbol{z} = \boldsymbol{c}^{\mathsf{H}}$$
  $$\frac{\partial}{\partial \boldsymbol{z}^{\mathsf{H}}} \boldsymbol{c}^{\mathsf{H}}\boldsymbol{z} = \boldsymbol{0}$$
- $D = 2$

Quaternionic

for $\boldsymbol{c} \in \mathbb{H}^{m\times 1}$

- $$\frac{\partial}{\partial \boldsymbol{q}} \boldsymbol{c}^{\mathsf{H}}\boldsymbol{q} = \boldsymbol{c}^{\mathsf{H}}$$
  $$\frac{\partial}{\partial \boldsymbol{q}^{\mathsf{H}}} \boldsymbol{c}^{\mathsf{H}}\boldsymbol{q} = -\tfrac{1}{2}\boldsymbol{c}^{*}$$
- $D = 4$

$$\frac{\partial}{\partial \boldsymbol{z}} \operatorname{Re}\{\boldsymbol{c}^{\mathsf{H}}\boldsymbol{z}\} = \frac{1}{D}\boldsymbol{c}^{\mathsf{H}}$$

for $\boldsymbol{M} = \boldsymbol{M}^{\mathsf{T}} \in \mathbb{R}^{m\times m}$

for $\boldsymbol{M} = \boldsymbol{M}^{\mathsf{H}} \in \mathbb{C}^{m\times m}$

for $\boldsymbol{M} = \boldsymbol{M}^{\mathsf{H}} \in \mathbb{H}^{m\times m}$

$$\frac{\partial}{\partial \boldsymbol{z}} \boldsymbol{z}^{\mathsf{H}}\boldsymbol{M}\boldsymbol{z} = \frac{2}{D}\boldsymbol{z}^{\mathsf{H}}\boldsymbol{M}$$

Arithmetics Rules:

- product and chain rule are much more intricate over $\mathbb{C}$ and $\mathbb{H}$ than over $\mathbb{R}$
- however, if the function is real-valued, simpler expressions are obtained

Product Rule:

- let $\boldsymbol{g}(\boldsymbol{q}) \in \mathbb{H}^n$ and $h(\boldsymbol{q}) \in \mathbb{R}$
- it can be shown that

$$\frac{\partial}{\partial \boldsymbol{q}}\Big(h \cdot \boldsymbol{g}\Big) = \boldsymbol{g}(\boldsymbol{q}) \cdot \left(\frac{\partial}{\partial \boldsymbol{q}} h(\boldsymbol{q})\right) + \left(\frac{\partial}{\partial \boldsymbol{q}} \boldsymbol{g}(\boldsymbol{q})\right) \cdot h(\boldsymbol{q})$$

Chain Rule:

- let $\boldsymbol{g}(\boldsymbol{q}) \in \mathbb{R}^n$ and $h(\boldsymbol{w}) \in \mathbb{R}$
- it can be shown that

$$\frac{\partial}{\partial \boldsymbol{q}} h(\boldsymbol{g}(\boldsymbol{q})) = \left.\frac{\partial h(\boldsymbol{w})}{\partial \boldsymbol{w}}\right|_{\boldsymbol{w}=\boldsymbol{g}(\boldsymbol{q})} \cdot \frac{\partial \boldsymbol{g}(\boldsymbol{q})}{\partial \boldsymbol{q}}$$

Random Vectors

Real-Valued Gaussian Random Vector:

- real-valued Gaussian random vector $\boldsymbol{n}$ of dimension M
- the Gaussian random vector is completely characterized by its *mean* and *covariance matrix*

$$\boldsymbol{\mu_n} = \mathrm{E}\{\boldsymbol{n}\}\,, \qquad \boldsymbol{\Phi_{nn}} = \mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^\mathsf{T}\}$$

- the probability density function (pdf) of $\boldsymbol{n}$ is given by

$$\mathrm{f}_{\boldsymbol{n}}(\boldsymbol{n}) = \frac{1}{\sqrt{(2\pi)^M\,|\boldsymbol{\Phi_{nn}}|}}\,\mathrm{e}^{-\frac{1}{2}\,(\boldsymbol{n}-\boldsymbol{\mu_n})^\mathsf{T}\boldsymbol{\Phi_{nn}}^{-1}\,(\boldsymbol{n}-\boldsymbol{\mu_n})}$$

Random Vectors (II)

Complex-Valued Gaussian Random Vector:

- complex-valued Gaussian random vector $\boldsymbol{n} = \boldsymbol{a} + \boldsymbol{b}\mathrm{i}$ of dimension M
- the Gaussian random vector is completely characterized by the ***mean*** and ***covariance matrix*** of the stacked vector $\overline{\boldsymbol{n}} \overset{\text{def}}{=} \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix}$

$$\boldsymbol{\mu}_{\overline{\boldsymbol{n}}} = \mathrm{E}\left\{ \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix} \right\}, \qquad \boldsymbol{\Phi}_{\overline{\boldsymbol{nn}}} = \mathrm{E}\left\{ \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix} \begin{bmatrix} \boldsymbol{a}^{\mathsf{T}} & \boldsymbol{b}^{\mathsf{T}} \end{bmatrix} \right\} = \begin{bmatrix} \boldsymbol{\Phi}_{\boldsymbol{aa}} & \boldsymbol{\Phi}_{\boldsymbol{ab}} \\ \boldsymbol{\Phi}_{\boldsymbol{ba}} & \boldsymbol{\Phi}_{\boldsymbol{bb}} \end{bmatrix}$$

or using the augmented vector $\underline{\boldsymbol{n}} \overset{\text{def}}{=} \begin{bmatrix} \boldsymbol{n} \\ \boldsymbol{n}^* \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}+\boldsymbol{b}\mathrm{i} \\ \boldsymbol{a}-\boldsymbol{b}\mathrm{i} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{I}\mathrm{i} \\ \boldsymbol{I} & -\boldsymbol{I}\mathrm{i} \end{bmatrix} \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix}$ by

$$\boldsymbol{\mu}_{\underline{\boldsymbol{n}}} = \mathrm{E}\left\{ \begin{bmatrix} \boldsymbol{n} \\ \boldsymbol{n}^* \end{bmatrix} \right\}, \qquad \boldsymbol{\Phi}_{\underline{\boldsymbol{nn}}} = \mathrm{E}\left\{ \begin{bmatrix} \boldsymbol{n} \\ \boldsymbol{n}^* \end{bmatrix} \begin{bmatrix} \boldsymbol{n}^{\mathsf{H}} & \boldsymbol{n}^{\mathsf{T}} \end{bmatrix} \right\} = \begin{bmatrix} \boldsymbol{\Phi}_{\boldsymbol{nn}} & \boldsymbol{\Psi}_{\boldsymbol{nn}} \\ \boldsymbol{\Psi}^*_{\boldsymbol{nn}} & \boldsymbol{\Phi}^*_{\boldsymbol{nn}} \end{bmatrix}$$

with ***covariance matrix*** and ***pseudo covariance matrix***

$$\boldsymbol{\Phi}_{\boldsymbol{nn}} \overset{\text{def}}{=} \mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^{\mathsf{H}}\}, \qquad \boldsymbol{\Psi}_{\boldsymbol{nn}} \overset{\text{def}}{=} \mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\}$$

Random Vectors (II)

Complex-Valued Gaussian Random Vector:

- complex-valued Gaussian random vector $\boldsymbol{n} = \boldsymbol{a} + \boldsymbol{b}\mathrm{i}$ of dimension M
- the random vector is called ***proper*** if the pseudo covariance matrix vanishes [NM'93]

$$\boldsymbol{\Psi}_{\boldsymbol{nn}} = \mathbf{0}$$

- a zero-mean Gaussian random vector is then ***rotationally invariant***
  ($\boldsymbol{n}$ and $\boldsymbol{n}\mathrm{e}^{\mathrm{i}\phi}$, $\phi \in \mathbb{R}$, have the same statistics)
- the proper Gaussian random vector is completely characterized by its ***mean*** and ***covariance matrix***

$$\boldsymbol{\mu}_{\boldsymbol{n}} = \mathrm{E}\{\boldsymbol{n}\} , \qquad \boldsymbol{\Phi}_{\boldsymbol{nn}} = \mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^{\mathsf{H}}\}$$

- the probability density function (pdf) of $\boldsymbol{n}$ is then formally given by

$$\mathrm{f}_{\boldsymbol{n}}(\boldsymbol{n}) = \frac{1}{\pi^M \, |\boldsymbol{\Phi}_{\boldsymbol{nn}}|} \, \mathrm{e}^{-(\boldsymbol{n}-\boldsymbol{\mu}_{\boldsymbol{n}})^{\mathsf{H}} \boldsymbol{\Phi}_{\boldsymbol{nn}}^{-1} (\boldsymbol{n}-\boldsymbol{\mu}_{\boldsymbol{n}})}$$

Random Vectors (III)

Quaternionic Gaussian Random Vector:

- quaternionic Gaussian random vector $\boldsymbol{n} = \boldsymbol{a} + \boldsymbol{b}\mathrm{i} + \boldsymbol{c}\mathrm{j} + \boldsymbol{d}\mathrm{k}$ of dimension M
- the Gaussian random vector is completely characterized by the ***mean*** and ***covariance matrix*** of the augmented vector ($\boldsymbol{n}^{(\mathrm{i})} = -\mathrm{i}\boldsymbol{n}\mathrm{i}$, similar for j and k)

$$\underline{\boldsymbol{n}} \stackrel{\text{def}}{=} \begin{bmatrix} \boldsymbol{n} \\ \boldsymbol{n}^{(\mathrm{i})} \\ \boldsymbol{n}^{(\mathrm{j})} \\ \boldsymbol{n}^{(\mathrm{k})} \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}+\boldsymbol{b}\mathrm{i}+\boldsymbol{c}\mathrm{j}+\boldsymbol{d}\mathrm{k} \\ \boldsymbol{a}+\boldsymbol{b}\mathrm{i}-\boldsymbol{c}\mathrm{j}-\boldsymbol{d}\mathrm{k} \\ \boldsymbol{a}-\boldsymbol{b}\mathrm{i}+\boldsymbol{c}\mathrm{j}-\boldsymbol{d}\mathrm{k} \\ \boldsymbol{a}-\boldsymbol{b}\mathrm{i}-\boldsymbol{c}\mathrm{j}+\boldsymbol{d}\mathrm{k} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{I}\mathrm{i} & \boldsymbol{I}\mathrm{j} & \boldsymbol{I}\mathrm{k} \\ \boldsymbol{I} & \boldsymbol{I}\mathrm{i} & -\boldsymbol{I}\mathrm{j} & -\boldsymbol{I}\mathrm{k} \\ \boldsymbol{I} & -\boldsymbol{I}\mathrm{i} & \boldsymbol{I}\mathrm{j} & -\boldsymbol{I}\mathrm{k} \\ \boldsymbol{I} & -\boldsymbol{I}\mathrm{i} & -\boldsymbol{I}\mathrm{j} & \boldsymbol{I}\mathrm{k} \end{bmatrix} \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \\ \boldsymbol{c} \\ \boldsymbol{d} \end{bmatrix}$$

$$\boldsymbol{\mu}_{\underline{\boldsymbol{n}}} = \mathrm{E}\left\{\begin{bmatrix} \boldsymbol{n} \\ \boldsymbol{n}^{(\mathrm{i})} \\ \boldsymbol{n}^{(\mathrm{j})} \\ \boldsymbol{n}^{(\mathrm{k})} \end{bmatrix}\right\}, \qquad \boldsymbol{\Phi}_{\underline{\boldsymbol{nn}}} = \mathrm{E}\left\{\begin{bmatrix} \boldsymbol{n} \\ \boldsymbol{n}^{(\mathrm{i})} \\ \boldsymbol{n}^{(\mathrm{j})} \\ \boldsymbol{n}^{(\mathrm{k})} \end{bmatrix} \begin{bmatrix} \boldsymbol{n}^{\mathrm{H}} & \boldsymbol{n}^{(\mathrm{i})\mathrm{H}} & \boldsymbol{n}^{(\mathrm{j})\mathrm{H}} & \boldsymbol{n}^{(\mathrm{k})\mathrm{H}} \end{bmatrix}\right\} = \begin{bmatrix} \boldsymbol{\Phi}_{\boldsymbol{nn}} & \boldsymbol{\Psi}_{\boldsymbol{nn}^{(\mathrm{i})}} & \boldsymbol{\Psi}_{\boldsymbol{nn}^{(\mathrm{j})}} & \boldsymbol{\Psi}_{\boldsymbol{nn}^{(\mathrm{k})}} \\ \boldsymbol{\Psi}^{(\mathrm{i})}_{\boldsymbol{nn}^{(\mathrm{i})}} & \boldsymbol{\Phi}^{(\mathrm{i})}_{\boldsymbol{nn}} & \boldsymbol{\Psi}^{(\mathrm{i})}_{\boldsymbol{nn}^{(\mathrm{k})}} & \boldsymbol{\Psi}^{(\mathrm{i})}_{\boldsymbol{nn}^{(\mathrm{j})}} \\ \boldsymbol{\Psi}^{(\mathrm{j})}_{\boldsymbol{nn}^{(\mathrm{j})}} & \boldsymbol{\Psi}^{(\mathrm{j})}_{\boldsymbol{nn}^{(\mathrm{k})}} & \boldsymbol{\Phi}^{(\mathrm{j})}_{\boldsymbol{nn}} & \boldsymbol{\Psi}^{(\mathrm{j})}_{\boldsymbol{nn}^{(\mathrm{i})}} \\ \boldsymbol{\Psi}^{(\mathrm{k})}_{\boldsymbol{nn}^{(\mathrm{k})}} & \boldsymbol{\Psi}^{(\mathrm{k})}_{\boldsymbol{nn}^{(\mathrm{j})}} & \boldsymbol{\Psi}^{(\mathrm{k})}_{\boldsymbol{nn}^{(\mathrm{i})}} & \boldsymbol{\Phi}^{(\mathrm{k})}_{\boldsymbol{nn}} \end{bmatrix}$$

with ***covariance matrix*** and ***three pseudo covariance matrices***

$$\boldsymbol{\Phi}_{\boldsymbol{nn}} \stackrel{\text{def}}{=} \mathrm{E}\{\boldsymbol{nn}^{\mathrm{H}}\}, \quad \boldsymbol{\Psi}_{\boldsymbol{nn}^{(\mathrm{i})}} \stackrel{\text{def}}{=} \mathrm{E}\{\boldsymbol{nn}^{(\mathrm{i})\mathrm{H}}\}, \quad \boldsymbol{\Psi}_{\boldsymbol{nn}^{(\mathrm{j})}} \stackrel{\text{def}}{=} \mathrm{E}\{\boldsymbol{nn}^{(\mathrm{j})\mathrm{H}}\}, \quad \boldsymbol{\Psi}_{\boldsymbol{nn}^{(\mathrm{k})}} \stackrel{\text{def}}{=} \mathrm{E}\{\boldsymbol{nn}^{(\mathrm{k})\mathrm{H}}\}$$

Random Vectors (III)

Quaternionic Gaussian Random Vector:

- quaternionic Gaussian random vector $\boldsymbol{n} = \boldsymbol{a} + \boldsymbol{b}\mathrm{i} + \boldsymbol{c}\mathrm{j} + \boldsymbol{d}\mathrm{k}$ of dimension M
- the random vector is called ***$\mathbb{Q}$-proper*** if all pseudo covariance matrices vanish [VRS'10]

$$\mathrm{E}\{\boldsymbol{n}(-\mathrm{i}\boldsymbol{n}^{\mathsf{H}}\mathrm{i})\} = \mathbf{0}\,, \qquad \mathrm{E}\{\boldsymbol{n}(-\mathrm{j}\boldsymbol{n}^{\mathsf{H}}\mathrm{j})\} = \mathbf{0}\,, \qquad \mathrm{E}\{\boldsymbol{n}(-\mathrm{k}\boldsymbol{n}^{\mathsf{H}}\mathrm{k})\} = \mathbf{0}$$

- a zero-mean Gaussian random vector is then ***right rotationally invariant***
  ($\boldsymbol{n}$ and $\boldsymbol{n}\mathrm{e}^{\eta\phi}$, $\phi \in \mathbb{R}$, $\eta = \eta_{\mathrm{i}}\mathrm{i} + \eta_{\mathrm{j}}\mathrm{j} + \eta_{\mathrm{k}}\mathrm{k}$, $|\eta| = 1$, $\eta^2 = -1$, have the same statistics)
- the Gaussian random vector is completely characterized by its ***mean*** and ***covariance matrix***

$$\boldsymbol{\mu}_{\boldsymbol{n}} = \mathrm{E}\{\boldsymbol{n}\}\,, \qquad \boldsymbol{\Phi}_{\boldsymbol{nn}} = \mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^{\mathsf{H}}\}$$

- the probability density function (pdf) of $\boldsymbol{n}$ is formally given by [VRS'10]

$$\mathrm{f}_{\boldsymbol{n}}(\boldsymbol{n}) = \frac{1}{(\pi/2)^{2M}\,|\boldsymbol{\Phi}_{\boldsymbol{nn}}|^2}\,\mathrm{e}^{-2\,(\boldsymbol{n}-\boldsymbol{\mu}_{\boldsymbol{n}})^{\mathsf{H}}\boldsymbol{\Phi}_{\boldsymbol{nn}}^{-1}\,(\boldsymbol{n}-\boldsymbol{\mu}_{\boldsymbol{n}})}$$

(the determinant of the quaternionic covariance matrix has to be understood as that of the equivalent complex-valued matrix of doubled dimensions) [Zha'97], [SLF'22]

Random Vectors (IV)

Real-Valued

real-valued Gaussian random vector $\boldsymbol{n}$

- *mean* and *covariance matrix*

$$\begin{aligned}\boldsymbol{\mu}_{\boldsymbol{n}} &= \mathrm{E}\{\boldsymbol{n}\}\\ \boldsymbol{\Phi}_{\boldsymbol{nn}} &= \mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\}\end{aligned}$$

- pdf $D = 1$

Complex-Valued

complex-valued Gaussian random vector $\boldsymbol{n} = \boldsymbol{a} + \boldsymbol{b}\mathrm{i}$

- *mean* and *covariance matrix*

$$\begin{aligned}\boldsymbol{\mu}_{\boldsymbol{n}} &= \mathrm{E}\{\boldsymbol{n}\}\\ \boldsymbol{\Phi}_{\boldsymbol{nn}} &= \mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^{\mathsf{H}}\}\end{aligned}$$

- proper if

$$\mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\} = \mathbf{0}$$

- pdf $D = 2$

Quaternionic

quaternionic Gaussian random vector $\boldsymbol{n} = \boldsymbol{a} + \boldsymbol{b}\mathrm{i} + \boldsymbol{c}\mathrm{j} + \boldsymbol{d}\mathrm{k}$

- *mean* and *covariance matrix*

$$\begin{aligned}\boldsymbol{\mu}_{\boldsymbol{n}} &= \mathrm{E}\{\boldsymbol{n}\}\\ \boldsymbol{\Phi}_{\boldsymbol{nn}} &= \mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^{\mathsf{H}}\}\end{aligned}$$

- proper if

$$\begin{aligned}\mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^{(\mathrm{i})\mathsf{H}}\} &= \mathbf{0}\\ \mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^{(\mathrm{j})\mathsf{H}}\} &= \mathbf{0}\\ \mathrm{E}\{\boldsymbol{n}\boldsymbol{n}^{(\mathrm{k})\mathsf{H}}\} &= \mathbf{0}\end{aligned}$$

- pdf $D = 4$

$$\mathrm{f}_{\boldsymbol{n}}(\boldsymbol{n}) = \frac{1}{(2\pi/D)^{MD/2}\,|\boldsymbol{\Phi}_{\boldsymbol{nn}}|^{D/2}}\,\mathrm{e}^{-\frac{D}{2}(\boldsymbol{n}-\boldsymbol{\mu}_{\boldsymbol{n}})^{\mathsf{H}}\boldsymbol{\Phi}_{\boldsymbol{nn}}^{-1}(\boldsymbol{n}-\boldsymbol{\mu}_{\boldsymbol{n}})}$$

Exponential Family

Observation Model:

- vector-valued observation model ($\mathbb{X} \in \{\mathbb{R},\ \mathbb{C},\ \mathbb{H}\}$)

measurement matrix $\in \mathbb{X}^{M \times N}$

zero-mean proper Gaussian random vector $\in \mathbb{X}^{M}$, independent from $\boldsymbol{x}$

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$$

measurement vector $\in \mathbb{X}^{M}$

signal vector, i.i.d. elements, marginal pdf $\mathrm{f}_x(x)$

- via Bayes' rule, the a-posteriori pdf of $\boldsymbol{x}$ given $\boldsymbol{y}$ reads

$$\mathrm{f}_{\boldsymbol{x}|\boldsymbol{y}}(\boldsymbol{x}) = \frac{1}{\mathrm{f}_{\boldsymbol{y}}(\boldsymbol{y})} \mathrm{f}_{\boldsymbol{x}}(\boldsymbol{x})\, \mathrm{f}_{\boldsymbol{n}}(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x})$$

- noise pdf ($\mathbb{R}$: $D = 1$; $\mathbb{C}$: $D = 2$; $\mathbb{H}$: $D = 4$)

$$\mathrm{f}_{\boldsymbol{n}}(\boldsymbol{n}) = \frac{1}{(2\pi/D)^{MD/2}\, |\boldsymbol{\Phi}_{\boldsymbol{nn}}|^{D/2}} \mathrm{e}^{-\frac{D}{2}\boldsymbol{n}^{\mathsf{H}}\boldsymbol{\Phi}_{\boldsymbol{nn}}^{-1}\boldsymbol{n}}$$

Exponential Family (II)

A-posteriori pdf:

- write the a-posteriori pdf in the form of an *exponential family*

natural parameters; $\boldsymbol{\lambda} = \mathrm{fct}(\boldsymbol{y})$ — sufficient statistics

$$\mathsf{f}_{\boldsymbol{x}|\boldsymbol{\lambda}}(\boldsymbol{x}) = h(\boldsymbol{x}) \cdot \mathrm{e}^{\mathrm{Re}\{\boldsymbol{\lambda}^{\mathsf{H}} \boldsymbol{g}(\boldsymbol{x})\} - A(\boldsymbol{\lambda})}$$

carrying density — log-partition function

- *natural parameter* $\boldsymbol{\lambda} \stackrel{\text{def}}{=} \boldsymbol{H}^{\mathsf{H}} \boldsymbol{\Phi}_{\boldsymbol{nn}}^{-1} \boldsymbol{y}$
- *sufficient statistics* $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{x}$

- a-posteriori pdf ($\mathbb{R}$: $D = 1$; $\mathbb{C}$: $D = 2$; $\mathbb{H}$: $D = 4$)

$$\mathsf{f}_{\boldsymbol{x}|\boldsymbol{\lambda}}(\boldsymbol{x}) = \underbrace{\mathsf{f}_{\boldsymbol{x}}(\boldsymbol{x})\, \mathrm{e}^{-\frac{D}{2}\boldsymbol{x}^{\mathsf{H}}\boldsymbol{H}^{\mathsf{H}}\boldsymbol{\Phi}_{\boldsymbol{nn}}^{-1}\boldsymbol{H}\boldsymbol{x}}}_{h(\boldsymbol{x})} \cdot \mathrm{e}^{D\,\mathrm{Re}\{\boldsymbol{\lambda}^{\mathsf{H}}\boldsymbol{x}\}} \cdot \underbrace{\frac{1}{(2\pi/D)^{MD/2}|\boldsymbol{\Phi}_{\boldsymbol{nn}}|^{D/2}} \frac{1}{\mathsf{f}_{\boldsymbol{y}}(\boldsymbol{\Phi}_{\boldsymbol{nn}}\boldsymbol{H}^{-\mathsf{H}}\boldsymbol{\lambda})}\, \mathrm{e}^{-\frac{D}{2}\boldsymbol{\lambda}^{\mathsf{H}}\boldsymbol{H}^{-1}\boldsymbol{\Phi}_{\boldsymbol{nn}}\boldsymbol{H}^{-\mathsf{H}}\boldsymbol{\lambda}}}_{\mathrm{e}^{-A(\boldsymbol{\lambda})}}$$

MMSE Estimators

Sketch of Proof: ($\mathbb{R}$: $D = 1$; $\mathbb{C}$: $D = 2$; $\mathbb{H}$: $D = 4$)

- $\mathsf{f}_{\boldsymbol{x}|\boldsymbol{\lambda}}(\boldsymbol{x})$ is a valid pdf, i.e.,

$$1 = \int \mathsf{f}_{\boldsymbol{x}|\boldsymbol{\lambda}}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int h(\boldsymbol{x}) \, \mathrm{e}^{D \, \mathrm{Re}\{\boldsymbol{\lambda}^{\mathsf{H}} \boldsymbol{x}\} - A(\boldsymbol{\lambda})} \, \mathrm{d}\boldsymbol{x}$$

(integration w.r.t. $\boldsymbol{x} = [x_1, \ldots, x_N]^{\mathsf{T}}$ means integration over all elements x_i and all components of each element)

- taking the ***(generalized) Wirtinger gradient*** w.r.t. $\boldsymbol{\lambda}^{\mathsf{H}}$ on both sides leads to

$$\mathbf{0} = \frac{\partial}{\partial \boldsymbol{\lambda}^{\mathsf{H}}} \int h(\boldsymbol{x}) \, \mathrm{e}^{D \, \mathrm{Re}\{\boldsymbol{\lambda}^{\mathsf{H}} \boldsymbol{x}\} - A(\boldsymbol{\lambda})} \, \mathrm{d}\boldsymbol{x} = \quad \ldots \quad = \mathrm{E}\{\boldsymbol{x} \mid \boldsymbol{\lambda}\} - \frac{\partial}{\partial \boldsymbol{\lambda}^{\mathsf{H}}} A(\boldsymbol{\lambda})$$

- hence, the ***conditional mean estimator*** can be written as

$$\mathrm{E}\{\boldsymbol{x} \mid \boldsymbol{\lambda}\} = \frac{\partial}{\partial \boldsymbol{\lambda}^{\mathsf{H}}} A(\boldsymbol{\lambda})$$

MMSE Estimators

Sketch of Proof: ($\mathbb{R}$: $D = 1$; $\mathbb{C}$: $D = 2$; $\mathbb{H}$: $D = 4$)

- $\mathsf{f}_{\boldsymbol{x}|\boldsymbol{\lambda}}(\boldsymbol{x})$ is a valid pdf, i.e.,

$$1 = \int \mathsf{f}_{\boldsymbol{x}|\boldsymbol{\lambda}}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int h(\boldsymbol{x}) \, \mathrm{e}^{D \, \mathrm{Re}\{\boldsymbol{\lambda}^{\mathsf{H}} \boldsymbol{x}\} - A(\boldsymbol{\lambda})} \, \mathrm{d}\boldsymbol{x}$$

(integration w.r.t. $\boldsymbol{x} = [x_1, \ldots, x_N]^{\mathsf{T}}$ means integration over all elements x_i and all components of each element)

- taking again the ***(generalized) Wirtinger gradient*** now w.r.t. $\boldsymbol{\lambda}$ on both sides leads to

$$\mathbf{0} = \frac{\partial}{\partial \boldsymbol{\lambda}} \frac{\partial}{\partial \boldsymbol{\lambda}^{\mathsf{H}}} \int h(\boldsymbol{x}) \, \mathrm{e}^{D \, \mathrm{Re}\{\boldsymbol{\lambda}^{\mathsf{H}} \boldsymbol{x}\} - A(\boldsymbol{\lambda})} \, \mathrm{d}\boldsymbol{x} = \quad \ldots \quad = \mathrm{E}\{\boldsymbol{e}\boldsymbol{e}^{\mathsf{H}}\} - \frac{\partial}{\partial \boldsymbol{\lambda}} \frac{\partial}{\partial \boldsymbol{\lambda}^{\mathsf{H}}} A(\boldsymbol{\lambda})$$

- hence, the ***error covariance matrix*** can be written as

$$\boldsymbol{\Phi}_{\boldsymbol{e}\boldsymbol{e}} = \mathrm{E}\{\boldsymbol{e}\boldsymbol{e}^{\mathsf{H}}\} = \frac{\partial}{\partial \boldsymbol{\lambda}} \frac{\partial}{\partial \boldsymbol{\lambda}^{\mathsf{H}}} A(\boldsymbol{\lambda}) = \frac{\partial}{\partial \boldsymbol{\lambda}} \mathrm{E}\{\boldsymbol{x} \mid \boldsymbol{\lambda}\}$$

- Formally, we have the same results for the real, the complex, and the quaternionic case! [Efr'23], [SF'25], [FS'26]

Application to Quaternionic Compressed Sensing

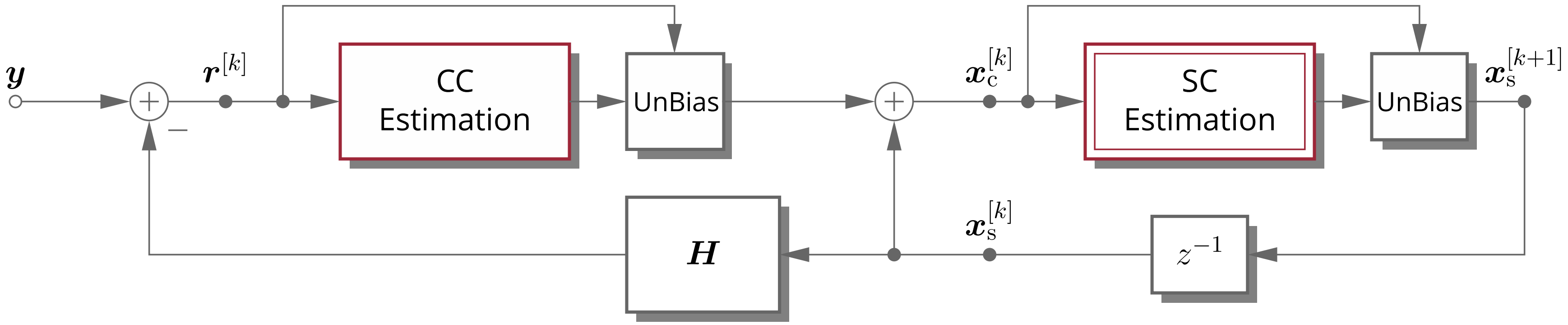
Problem and Algorithms:

- quaternionic compressed sensing problem

$$y = Hx + n$$

where:
- H: sensing matrix $\in \mathbb{H}^{M\times N}$
- n: zero-mean quaternionic proper Gaussian random vector $\in \mathbb{H}^{M}$, variance σ_n^2, independent from x
- y: measurement vector $\in \mathbb{H}^{M}$
- x: sparse signal vector $\in \mathbb{H}^{N}$, i.i.d. elements, marginal pdf $f_x(x)$ with Dirac component at $x = 0$

- Vector Approximate Message Passing (VAMP) / Orthogonal AMP (OAMP) [RSF'19], [MP'17]
  - "channel-constrained (CC)" (consider H, ignore $f_x(x)$) and "signal-constrained (SC)" (consider $f_x(x)$, ignore H) MMSE estimation problems
  - *bias* has to be removed / calculation of ***extrinsics***



Application to Quaternionic Compressed Sensing

Problem and Algorithms:

- quaternionic compressed sensing problem

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$$

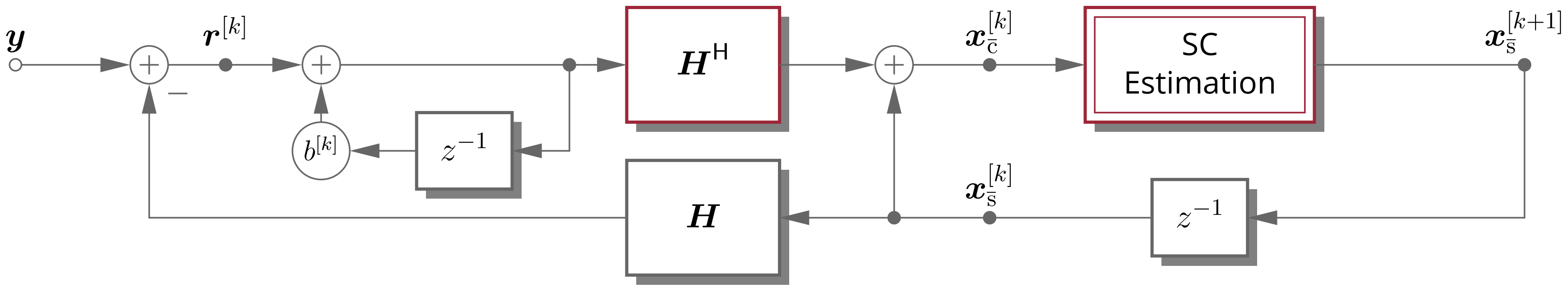
sensing matrix $\in \mathbb{H}^{M \times N}$

zero-mean quaternionic proper Gaussian random vector $\in \mathbb{H}^{M}$, variance σ_n^2, independent from $\boldsymbol{x}$

measurement vector $\in \mathbb{H}^{M}$

sparse signal vector $\in \mathbb{H}^{N}$, i.i.d. elements marginal pdf $f_x(x)$ with Dirac component at $x = 0$

- Approximate Message Passing (AMP) [DMM'10]
  - Matched filter (MF) and "signal-constrained (SC)" MMSE estimation (BAMP) or soft thresholding ("denoising")
  - Onsager correction

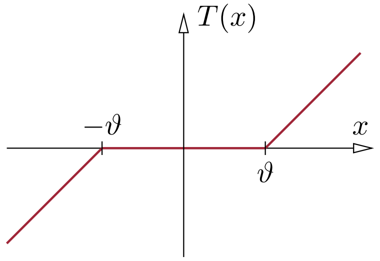


Application to Quaternionic Compressed Sensing (II)

Real-Valued

- soft thresholding

$$T(x) = \begin{cases} x - \vartheta \frac{x}{|x|}, & |x| > \vartheta \\ 0, & \text{else} \end{cases}$$



- variance calculation using

$$\left| \frac{\mathrm{d}}{\mathrm{d}x} T(x) \right| = 1 \,, \quad |x| > \vartheta$$

- adjustment of threshold

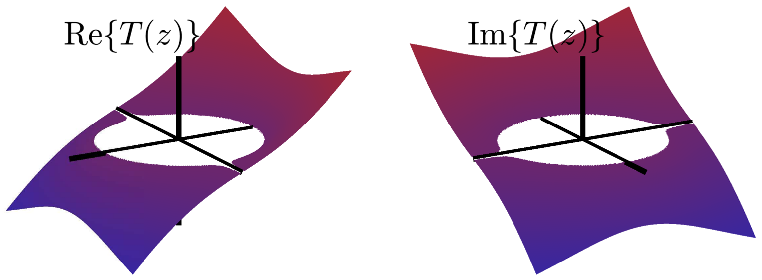
$$\vartheta = \frac{\operatorname{median}(|\boldsymbol{x}_{\bar{c}}|)}{\sqrt{2}\,\operatorname{erf}^{-1}(1/2)}$$

(erf^{-1}: inverse of the error function)

Complex-Valued

- soft thresholding

$$T(z) = \begin{cases} z - \vartheta \frac{z}{|z|}, & |z| > \vartheta \\ 0, & \text{else} \end{cases}$$



- variance calculation using

$$\left| \frac{\partial}{\partial z} T(z) \right| = 1 - \frac{\vartheta}{2|z|} \,, \quad |z| > \vartheta$$

- adjustment of threshold

$$\vartheta = \frac{\operatorname{median}(|\boldsymbol{x}_{\bar{c}}|)}{\sqrt{\log(2)}}$$

Quaternionic

- soft thresholding

$$T(q) = \begin{cases} q - \vartheta \frac{q}{|q|}, & |q| > \vartheta \\ 0, & \text{else} \end{cases}$$

- variance calculation using

$$\left| \frac{\partial}{\partial q} T(q) \right| = 1 - \frac{3\vartheta}{4|q|} \,, \quad |q| > \vartheta$$

- adjustment of threshold

$$\vartheta = \frac{\operatorname{median}(|\boldsymbol{x}_{\bar{c}}|)}{\sqrt{(-W(-\frac{1}{2\mathrm{e}}) - 1)/2}}$$

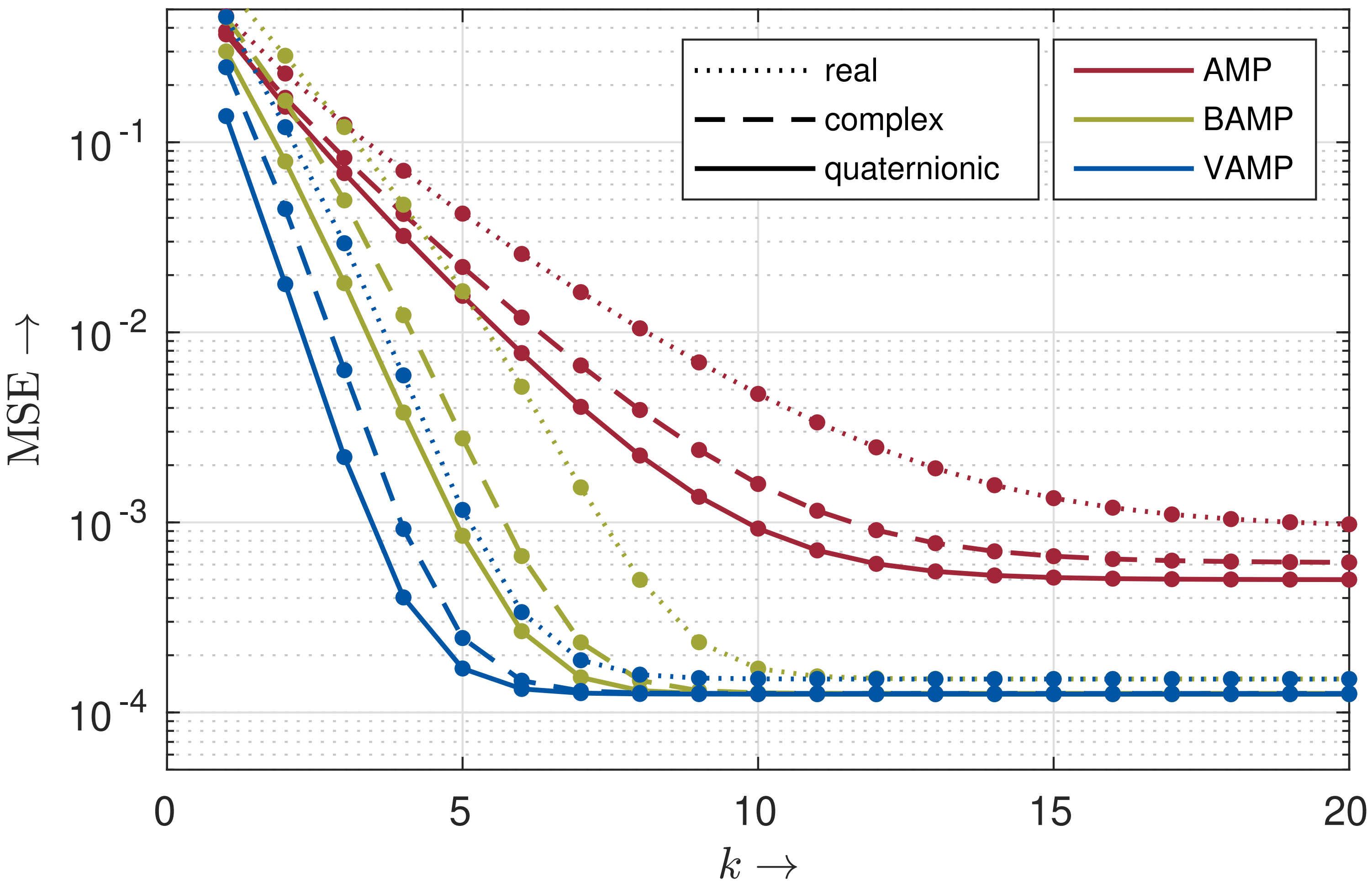
($W(z)$: branch -1 of Lambert W function)

Numerical Results:

- parameters:
  - signal dimension $N = 1024$
  - observation dimension $M = 512$
  - relative sparsity $\tau = 102/1024 \approx 0.1$
  - Bernoulli–Gaussian quaternionic signal
  - signal-to-noise ratio $30\ \mathrm{dB}$
  - elements of $\boldsymbol{H}$ i.i.d. quaternionic Gaussian
  - performance metric: MSE per element

- algorithms:

	CC	SC	Correct
AMP	MF	ST	Onsager
BAMP	MF	MMSE	Onsager
VAMP	MMSE	MMSE	unbiasing



Summary

Quaternionic MMSE Estimation:

- quaternions are a handy tool whenever two complex signals are treated jointly
- quaternion algebra has some peculiarities
- MMSE estimation of quaternionic signals occurs in various fields

We Have Presented:

- a concise method for calculating complex-valued and quaternionic MMSE estimators
- gradients have to be defined suitably
- formally the same results are obtained in the real, complex, and quaternionic case

Application:

- the derived MMSE estimators can directly be applied in iterative (AMP-type) algorithms
- e.g., for compressed sensing

References

[AMO'13] L. Anitori, A. Maleki, M. Otten, R.G. Baraniuk, P. Hoogeboom. Design and Analysis of Compressed Sensing Radar Detectors. *IEEE Transactions on Signal Processing*, vol. 61, no. 4, pp. 813–827, 2013.

[CS'03] J.H. Conway, D.A. Smith. *On Quaternions and Octonions*. CRC Press, Boca Raton, FL, USA, 2003.

[DMM'10] D.L. Donoho, A. Maleki, A. Montanari. Message Passing Algorithms for Compressed Sensing: I. Motivation and Construction; II. Analysis and Validation. In *IEEE Information Theory Workshop (ITW)*, Cato, Egypt, 2010.

[Efr'23] B. Efron. *Exponential Families in Theory and Practice*, Cambridge University Press, Cambridge, 2023.

[Fis'02] R.F.H. Fischer. *Precoding and Signal Shaping for Digital Transmission*, John Wiley & Sons, New York, 2002.

[FS'26] R.F.H. Fischer, E. Sterk. Quaternionic MMSE Estimation and Iterative Algorithms for Quaternionic Compressed Sensing. 2026 International Zurich Seminar on Information and Communication, Zurich, Feb. 2026.

[GMW'23] A. Grathwohl, B. Meinecke, M. Widmann, J. Kanz, C. Waldschmidt. UAV-Based Bistatic SAR-Imaging Using a Stationary Repeater. *IEEE Journal of Microwaves*, vol. 3, no. 2, pp. 625–634, April 2023.

[KA'16] M. Karlsson, E. Agrell. Multidimensional Optimized Optical Modulation Formats. Chapter 2 in *Enabling Technologies for High Spectral-Efficiency Coherent Optical Communication Networks*, editors X. Zhou, C. Xie, Wiley, Hoboken, NJ, USA, 2016.

[KFL'01] F.R. Kschischang, B.J. Frey, H.-A. Loeliger. Factor Graphs and the Sum-Product Algorithm. *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 498–519, Feb. 2001.

[LDH'07] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, F.R. Kschischang. The Factor Graph Approach to Model-Based Signal Processing, *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1295–1322, June 2007.

[MP'17] J. Ma, L. Ping. Orthogonal AMP. *IEEE Access*, vol. 5, pp. 2020–2033, 2017.

[MJT'11] D.P. Mandic, C. Jahanchahi, C.C. Took. A Quaternion Gradient Operator and Its Applications. *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 47–50, 2011.

[MBM'06] S. Miron, N. Le Bihan, J.I. Mars. Quaternion-MUSIC for Vector-Sensor Array Processing. *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1218–1229, 2006.

References

[NM'93] F.D. Neeser, J.L. Massey. Proper Complex Random Processes with Applications to Information Theory. *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1293–1302, July 1993.

[RSF'19] S. Rangan, P. Schniter, A.K. Fletcher. Vector Approximate Message Passing. *IEEE Transactions on Information Theory*, vol. 65, no. 10, pp. 6664–6684, 2019.

[Sab'06] A.M. Sabatini. Quaternion-Based Extended Kalman Filter for determining Orientation by Inertial and Magnetic Sensing. *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 7, pp. 1346–1356, 2006.

[SF'25] E. Sterk, R.F.H. Fischer. Iterative Recovery Algorithms for Complex-Valued Distributed Compressed Sensing. In *14th International ITG Conference on Systems, Communications, and Coding (SCC)*, Karlsruhe, Germany, 2025.

[SLF'22] S. Stern, C. Ling, R.F.H. Fischer. Algorithms and Bounds for Complex and Quaternionic Lattices With Application to MIMO Transmission. *IEEE Transactions on Information Theory*, vol. 68, no. 7, pp. 4491–4517, July 2022.

[TC'09] W. Tan, A.R. Calderbank. Multiuser Detection of Alamouti Signals. *IEEE Transactions on Communications*, vol. 57, no. 7, pp. 2080–2089, 2009.

[SPMag'24] N.A. Valous, E. Hitzer, S. Vitabile, S. Bernstein, C. Lavor D. Abbott, M.E. Luna-Elizarrarás, W. Elena (editors). Hypercomplex Signal and Image Processing: Part 1 and Part 2. *IEEE Signal Processing Magazine*, vol. 41, nos. 2 and 3, 2024.

[VRS'10] J. Vía, D. Ramírez, I. Santamaría. Properness and Widely Linear Processing of Quaternion Random Vectors. *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3502–3515, July 2010.

[WX'19] G. Wang, R. Xue. Comments on "The Quaternion LMS Algorithm for Adaptive Filtering of Hypercomplex Processes". *IEEE Transactions on Signal Processing*, vol. 67, no. 7, pp. 1957–1958, 2019.

[XJT'15] D. Xu, C. Jahanchahi, C.C. Took, D.P. Mandic. Enabling Quaternion Derivatives: The Generalized HR Calculus. *Royal Society Open Science*, vol. 2, no. 8, 2015.

[XM'15] D. Xu, D.P. Mandic. The Theory of Quaternion Matrix Derivatives. *IEEE Transactions on Signal Processing*, vol 63, no. 6, pp. 1543–1556, 2015.

[Zha'97] F. Zhang. Quaternions and Matrices of Quaternions. *Linear Algebra and its Applications*, vol. 251, pp. 21–57, 1997.