Uncertainty-Aware Dimensionality Reduction for Channel Charting with Geodesic Loss

Florian Euchner, Phillip Stephan,
Stephan ten Brink

Institute of Telecommunications, University of Stuttgart

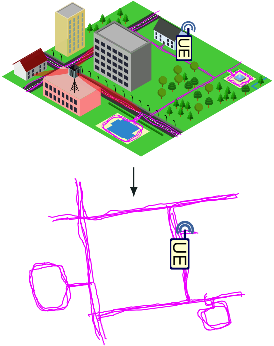
Asilomar Conference 2024
October 30th, 2024

University of Stuttgart
Institute of Telecommunications
Prof. Dr. Ing. Stephan ten Brink

Agenda

Agenda

Objective of Channel Charting [Studer et al., 2018]



Channel Charting is a **dimensionality reduction** technique that learns a mapping from a high-dimensional **channel state information (CSI)** manifold to a low-dimensional space, called **Channel Chart (CC)**, purely from data available at the base station.

- The mapping is called the **forward charting function**, usually $D = 2$ or $D = 3$:

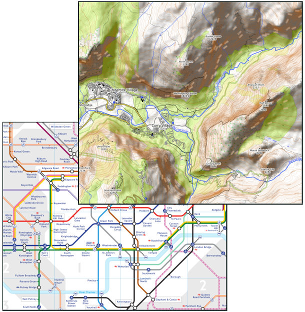
$$\mathbb{C}^{B \times M \times N_{\text{sub}}} \rightarrow \boxed{\text{FCF } C_\Theta} \rightarrow \mathbb{R}^D$$

- Channel Charting is self-supervised

What is a Channel Chart?

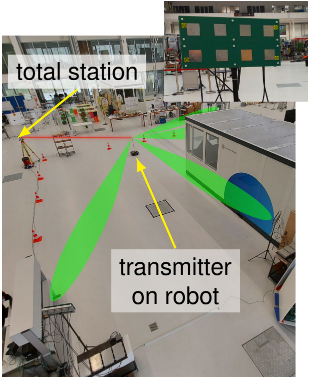
A channel chart is a *map* and C_Θ localizes the UE on the map:

- Maps depict (spatial) relationships
- Different maps preserve different properties, serve different purposes
- Applications:
  - UE clustering, pilot allocation
  - Channel prediction
  - CSI compression
  - **Localization (our focus!)**
  - ...



TfL Tube Map, OpenTopoMap

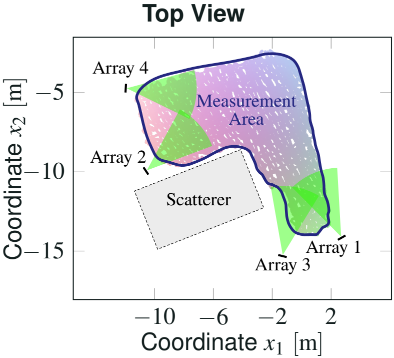
CSI Dataset: Distributed mMIMO in Factory



- Common benchmark:
  [Stahlke et al., 2023, Taner et al., 2023, Euchner et al., 2023, Taner et al., 2024, Vindas and Guillaud, 2024, Stephan et al., 2023, Chaaya et al., 2024, Euchner et al., 2024]
- DICHASUS (Distributed Channel Sounder by University of Stuttgart) dataset `dichasus-cf0x`, publicly available[a]
- Single position-tracked TX on robot
- $B = 4$ arrays with $M = 2 \times 4$ antennas
- $50\,\mathrm{MHz}$ bandwidth, $f_\mathrm{c} = 1.27\,\mathrm{GHz}$, OFDM

[a] `https://dichasus.inue.uni-stuttgart.de/datasets/data/dichasus-cf0x/`

CSI Dataset: Distributed mMIMO in Factory



University of Stuttgart
Institute of Telecommunications

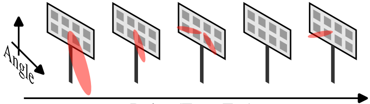
State of the Art

of dissimilarity-metric based Channel Charting

Dissimilarity Metrics

Angle-Delay-Profile Δ_{ADP}

- Same power from same angle at same delay $\rightarrow$ similar location



- Uses *time-domain* CSI $\mathbf{H} \in \mathbb{C}^{B \times M \times T}$:

$$\Delta_{\mathrm{ADP},i,j} = \sum_{b=1}^{B} \sum_{\tau=1}^{T} \left(1 - \frac{\left| \sum_{m=1}^{M} \left(\mathbf{H}_{b,m,\tau}^{(i)} \right)^* \mathbf{H}_{b,m,\tau}^{(j)} \right|^2}{\left(\sum_{m=1}^{M} \left| \mathbf{H}_{b,m,\tau}^{(i)} \right|^2 \right) \left(\sum_{m=1}^{M} \left| \mathbf{H}_{b,m,\tau}^{(j)} \right|^2 \right)} \right)$$

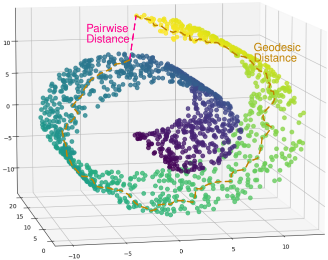
Timestamps Δ_{time}

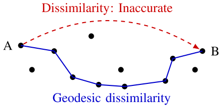
- Use time difference as dissimilarity

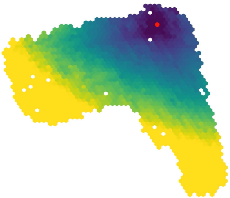
$$\Delta_{\mathrm{time},i,j} = \left| t^{(i)} - t^{(j)} \right|$$

University of Stuttgart
Institute of Telecommunications

Geodesic Dissimilarities Δ_{geo}

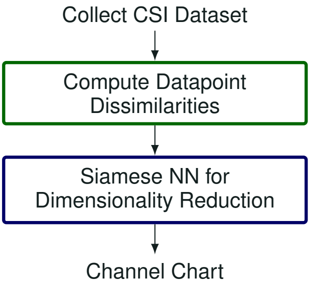


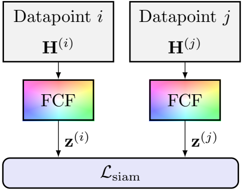




Dissimilarity metrics only locally accurate? Apply **shortest path algorithm** to a neighborhood graph!

Siamese Neural Network Training





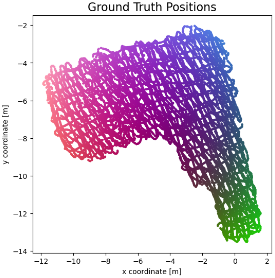
$$\mathcal{L}_{\mathrm{siam}} = \frac{1}{L^2} \sum_{i,j} \frac{\left(\left\| \mathbf{z}^{(i)} - \mathbf{z}^{(j)} \right\| - \Delta_{\mathrm{geo},i,j} \right)^2}{\Delta_{\mathrm{geo},i,j} + \beta}$$

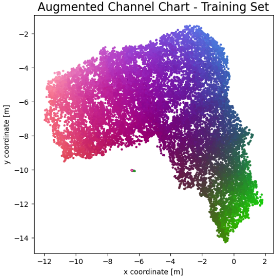
with CC locations $\mathbf{z}^{(l)} \in \mathbb{R}^2$, $l = 1, \ldots, |\mathcal{S}|$.

University of Stuttgart
Institute of Telecommunications

Exemplary Channel Charts

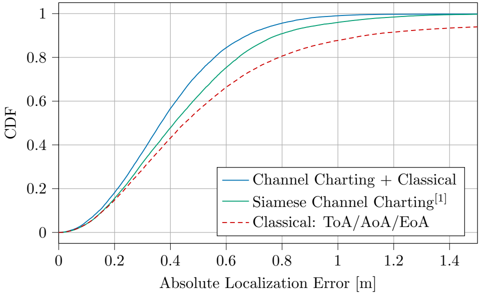
The Good (Asilomar ’23): CC + Classical Techniques





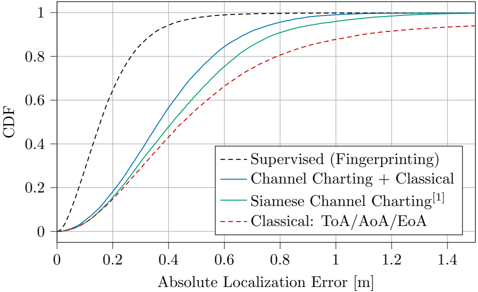
CC + AoA + ToA produces channel chart with $\mathrm{MAE} \approx 0.40\,\mathrm{m}$ [Euchner et al., 2023]

The Good (Asilomar ’23): CC + Classical Techniques



[1]Evaluated after optimal affine transform

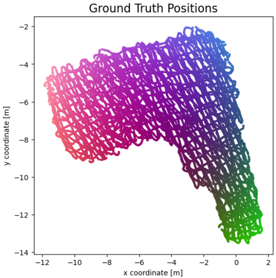
The Good (Asilomar ’23): CC + Classical Techniques

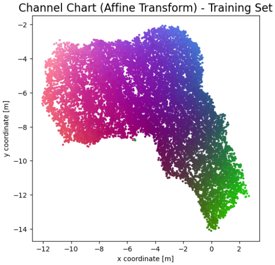


$\rightarrow$ Still room for improvement!

[1]Evaluated after optimal affine transform

The Bad (TCOM ’23): CC on nonconvex shape

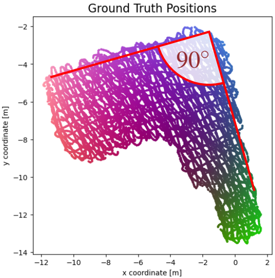


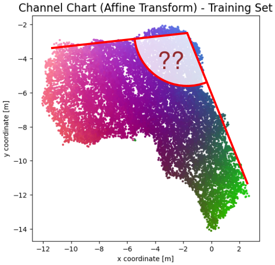


CC produces channel chart with $\mathrm{MAE} \approx 0.44\,\mathrm{m}$ [Stephan et al., 2023]

University of Stuttgart
Institute of Telecommunications

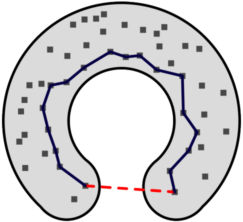
The Bad (TCOM ’23): CC on nonconvex shape





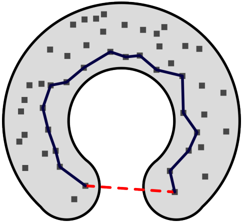
CC produces channel chart with $\mathrm{MAE} \approx 0.44\,\mathrm{m}$ [Stephan et al., 2023]

The Bad (TCOM ’23): CC on nonconvex shape



$$\text{Euclidean Distance} \overset{!}{\neq} \text{Geodesic Distance}$$

The Bad (TCOM '23): CC on nonconvex shape

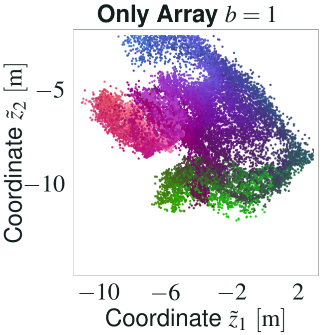


Euclidean Distance $\stackrel{!}{\neq}$ Geodesic Distance

$\rightarrow$ Well-known fundamental dimensionality reduction issue!

[Budninskiy et al., 2019, Rosman et al., 2010, Zha and Zhang, 2003, Schwartz and Talmon, 2019]

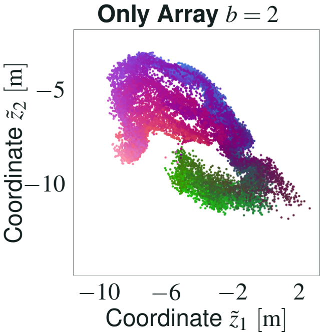
The Ugly: CC with a Single Antenna Array



Ant. Array	MAE [a]
$b = 1$	$\approx 2.04\,\mathrm{m}$

[a] Evaluated after optimal affine transform

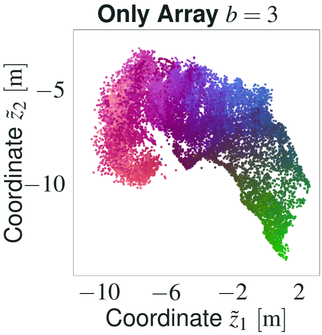
The Ugly: CC with a Single Antenna Array



Ant. Array	MAE [a]
$b = 1$	$\approx 2.04\,\mathrm{m}$
$b = 2$	$\approx 2.23\,\mathrm{m}$

[a] Evaluated after optimal affine transform

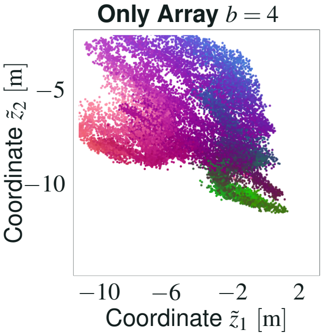
The Ugly: CC with a Single Antenna Array



Ant. Array	MAE [a]
$b = 1$	$\approx 2.04\,\mathrm{m}$
$b = 2$	$\approx 2.23\,\mathrm{m}$
$b = 3$	$\approx 1.37\,\mathrm{m}$

[a] Evaluated after optimal affine transform

The Ugly: CC with a Single Antenna Array

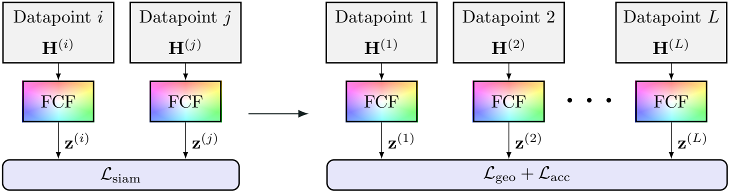


Ant. Array	MAE [a]
$b = 1$	$\approx 2.04\,\mathrm{m}$
$b = 2$	$\approx 2.23\,\mathrm{m}$
$b = 3$	$\approx 1.37\,\mathrm{m}$
$b = 4$	$\approx 1.62\,\mathrm{m}$

[a] Evaluated after optimal affine transform

Agenda

Siamese to Batch-Wise Loss

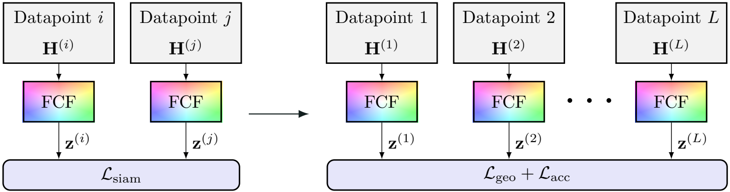


$2 \times |\mathcal{B}|$ FCF invocations

$|\mathcal{S}| = L$ FCF invocations

- Concept not entirely new [Hermans et al., 2017, Sohn, 2016]
- Not more complex if batches $\mathcal{B}$ were already large: $2 \times |\mathcal{B}| \approx |\mathcal{S}| = L$
- *Global view* (whole channel chart) of the problem in loss function

Siamese to Batch-Wise Loss

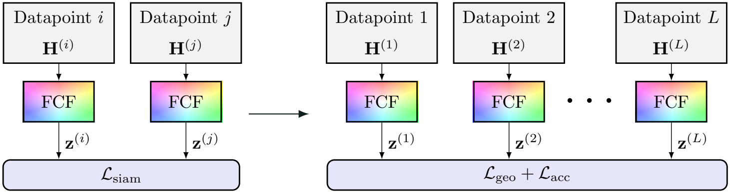


$2 \times |\mathcal{B}|$ FCF invocations

$|\mathcal{S}| = L$ FCF invocations

- Concept not entirely new [Hermans et al., 2017, Sohn, 2016]
- Not more complex if batches $\mathcal{B}$ were already large: $2 \times |\mathcal{B}| \approx |\mathcal{S}| = L$
- *Global view* (whole channel chart) of the problem in loss function

Siamese to Batch-Wise Loss

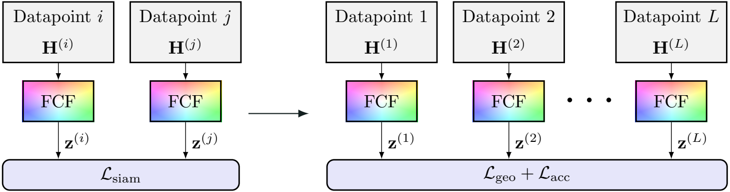


$2 \times |\mathcal{B}|$ FCF invocations

$|\mathcal{S}| = L$ FCF invocations

- Concept not entirely new [Hermans et al., 2017, Sohn, 2016]
- Not more complex if batches $\mathcal{B}$ were already large: $2 \times |\mathcal{B}| \approx |\mathcal{S}| = L$
- *Global view* (whole channel chart) of the problem in loss function

Siamese to Batch-Wise Loss



$2 \times |\mathcal{B}|$ FCF invocations

$|\mathcal{S}| = L$ FCF invocations

- Concept not entirely new [Hermans et al., 2017, Sohn, 2016]
- Not more complex if batches $\mathcal{B}$ were already large: $2 \times |\mathcal{B}| \approx |\mathcal{S}| = L$
- *Global view* (whole channel chart) of the problem in loss function

Improvement 1: Acceleration Constraint

- Exploit physical *inertia* of transmitter [Rappaport et al., 2021]
- Instantaneous velocity and acceleration in Channel Chart:

$$\mathbf{v}^{(l)} = \frac{\mathbf{z}^{(l)} - \mathbf{z}^{(l-1)}}{t^{(l)} - t^{(l-1)}} \qquad \mathbf{a}^{(l)} = \frac{\mathbf{v}^{(l)} - \mathbf{v}^{(l-1)}}{t^{(l)} - t^{(l-1)}}$$

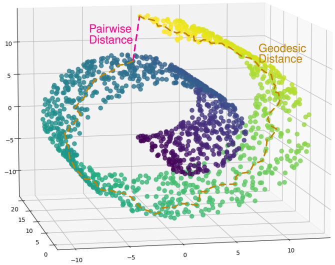
- Enforce log-likelihood acceleration constraint based on assumed distribution $a = \|\mathbf{a}\| \sim \mathcal{N}(\mu_{\mathrm{acc}}, \sigma_{\mathrm{acc}}^2)$:

$$\mathcal{L}_{\mathrm{acc}} = \frac{1}{L} \sum_{l=1}^{L} \frac{\left(\left\| \mathbf{a}^{(l)} \right\| - \mu_{\mathrm{acc}} \right)^2}{2\sigma_{\mathrm{acc}}^2}$$

Agenda

1. Introduction, Motivation, State of the Art
2. New Training Architecture
3. Geodesic Loss
4. Uncertainty (Concept Overview)
5. Evaluation and Results
6. Conclusion and Outlook

Improvement 2: Path-based Loss Function



Remember path $q_{i,j}[k]$ through manifold (*predecessor matrix*)!

- Paths $q_{i,j}[k]$: Index l of k-th datapoint in path from datapoint i to j
- Length of path through channel chart:

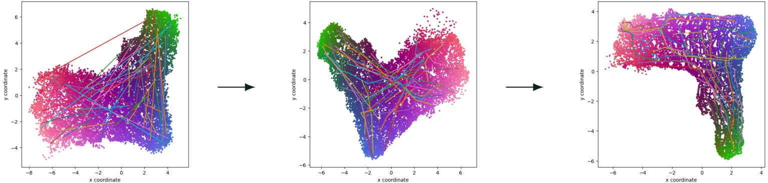
$$d_{\mathrm{path},i,j} = \sum_{k=1}^{K_{i,j}-1} \left\| \mathbf{z}^{(q_{i,j}[k])} - \mathbf{z}^{(q_{i,j}[k+1])} \right\|$$

- Compare geodesic distance on manifold to path length:

$$\mathcal{L}_{\mathrm{geo}} = \frac{1}{L^2} \sum_{i,j} \frac{(d_{\mathrm{path},i,j} - \Delta_{\mathrm{geo},i,j})^2}{\Delta_{\mathrm{geo},i,j} + \beta}$$

Implementation Issues

- $d_{\text{path},i,j}$ depends on $\mathbf{z}^{(q_{i,j}[1])}$ to $\mathbf{z}^{(q_{i,j}[K_{i,j}])}$ $\rightarrow$ many degrees of freedom
- Training may become unstable (optimizer runs into local minima)
- Solution: Train with sub-sampled paths $q_{i,j}^{(s)}[k] = q_{i,j}[k \cdot s]$, sub-sampling factor s depends on $d_{\text{path},i,j}$ and decreases during training



- Additional constraint to prevent loops / local minima, omitted here

Agenda

Sammon's Loss vs Gaussian Log-Likelihood

Sammon's / Siamese Loss:

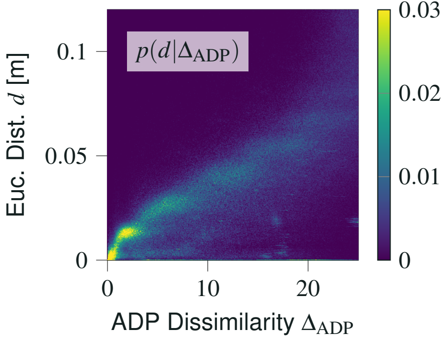
$$\mathcal{L}_{\text{Siam}} = \frac{1}{L^2} \sum_{i,j} \frac{\left(\left\|\mathbf{z}_i - \mathbf{z}_j\right\| - \Delta_{i,j}\right)^2}{\Delta_{i,j} + \beta}$$

Gaussian Log-Likelihood:

$$\mathcal{L}_{\text{ML}} = \text{const.} + \frac{1}{L^2} \sum_{i,j} \frac{\left(\left\|\mathbf{z}_i - \mathbf{z}_j\right\| - \Delta_{i,j}\right)^2}{2\sigma_{i,j}^2}$$

- Comparison reveals implicit assumption $\sigma_{i,j}^2 = \frac{\Delta_{i,j}+\beta}{2}$
- May have more information about uncertainty in $\Delta_{i,j}$
- Idea: **Probabilistic** model for dissimilarities Δ and distances d

Uncertainty: Formalization



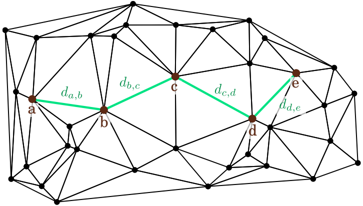
- Model dissimilarities Δ, distances d as random variables
- Assume known model

$$p(d_{i,j} \mid \Delta_{i,j})$$

for every type of dissimilarity

University of Stuttgart
Institute of Telecommunications

Many Questions



Q1. How to run shortest path algorithm if $d_{i,j}$ is non-deterministic?

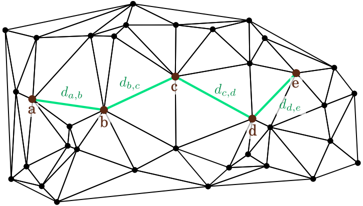
Q2. What is the distribution of

$$d_{\text{geo},i,j} = \sum_{k=1}^{K_{i,j}-1} d_{q_{i,j}[k],q_{i,j}[k+1]},$$

which is needed for $\mathcal{L}_{\text{geo}}$?

Q3. Correlations between $d_{a,b}$, $d_{b,c}$, ...?

Many Questions



Q1. How to run shortest path algorithm if $d_{i,j}$ is non-deterministic?

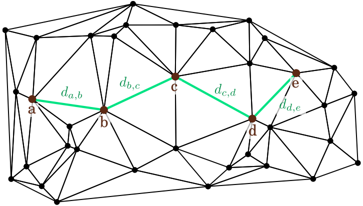
Q2. What is the distribution of

$$d_{\text{geo},i,j} = \sum_{k=1}^{K_{i,j}-1} d_{q_{i,j}[k],q_{i,j}[k+1]},$$

which is needed for $\mathcal{L}_{\text{geo}}$?

Q3. Correlations between $d_{a,b}$, $d_{b,c}$, ...?

Many Questions



Q1. How to run shortest path algorithm if $d_{i,j}$ is non-deterministic?

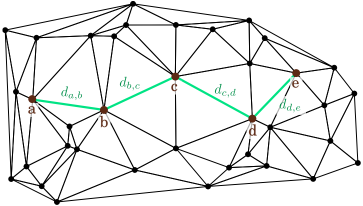
Q2. What is the distribution of

$$d_{\mathrm{geo},i,j} = \sum_{k=1}^{K_{i,j}-1} d_{q_{i,j}[k],q_{i,j}[k+1]},$$

which is needed for $\mathcal{L}_{\mathrm{geo}}$?

Q3. Correlations between $d_{a,b}$, $d_{b,c}$, ...?

Many Questions



Q1. How to run shortest path algorithm if $d_{i,j}$ is non-deterministic?

Q2. What is the distribution of

$$d_{\mathrm{geo},i,j} = \sum_{k=1}^{K_{i,j}-1} d_{q_{i,j}[k],q_{i,j}[k+1]},$$

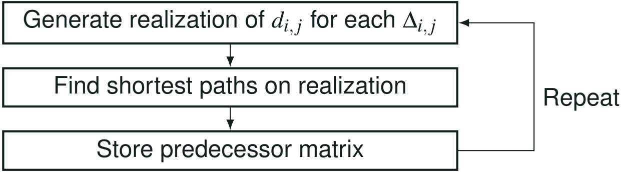
which is needed for $\mathcal{L}_{\mathrm{geo}}$?

Q3. Correlations between $d_{a,b}$, $d_{b,c}$, ...?

Need pragmatic answers!

(Q1) Shortest Path or Short Path

- Isomap, Siamese CC rely on *shortest* path and convex shape
- Geodesic Loss: Some *short* (no *shortest*) path is sufficient
- Algorithm to find *some short paths*:



- Now: Figure out assumed distribution of path lengths $d_{\mathrm{geo},i,j}$

(Q1) Shortest Path or Short Path

- Isomap, Siamese CC rely on *shortest* path and convex shape
- Geodesic Loss: Some *short* (no *shortest*) path is sufficient
- Algorithm to find *some short paths*:

Some short paths are sufficient!

Gene[illegible] $_{i,j}$ [illegible] $_{i,j}$

↓

Find shortest paths on realization

↓

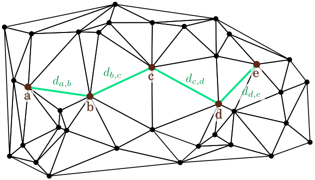
Store predecessor matrix

Repeat

- Now: Figure out assumed distribution of path lengths $d_{\mathrm{geo},i,j}$

(Q2) Distribution of $d_{\text{geo},i,j}$

- Arbitrary distributions $p(d|\Delta)$ are too hard to handle
- Assumption: Pairwise distances $d_{i,j}|\Delta_{i,j} \sim \mathcal{N}(\mu_{i,j}, \sigma^2_{i,j})$
- Hence: $d_{\text{geo},i,j} \sim \mathcal{N}(\mu_{\text{geo},i,j}, \sigma^2_{\text{geo},i,j})$ with

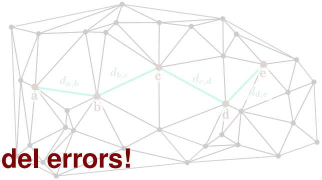


$$\mu_{\text{geo},i,j} = \sum_{k=1}^{K_{i,j}-1} \mu_{q_{i,j}[k],q_{i,j}[k+1]}$$

$$\sigma^2_{\text{geo},i,j} = \sum_{k=1}^{K_{i,j}-1} \sigma^2_{q_{i,j}[k],q_{i,j}[k+1]} + 2 \sum_{1 \leq a < b \leq k} \text{Cov}\left(d_{q_{i,j}[a],q_{i,j}[a+1]}, d_{q_{i,j}[b],q_{i,j}[b+1]}\right)$$

(Q2) Distribution of $d_{\text{geo},i,j}$

- Arbitrary distributions $p(d|\Delta)$ are too hard to handle
- Assumption: Pairwise distances $d_{i,j}|\Delta_{i,j} \sim \mathcal{N}(\mu_{i,j}, \sigma^2_{i,j})$
- Hence: $d_{\text{geo},i,j} \sim \mathcal{N}(\mu_{\text{geo},i,j}, \sigma^2_{\text{geo},i,j})$ with



Use normal distribution to model errors!

$$\mu_{\text{geo},i,j} = \sum_{k=1}^{K_{i,j}-1} \mu_{q_{i,j}[k],q_{i,j}[k+1]}$$

$$\sigma^2_{\text{geo},i,j} = \sum_{k=1}^{K_{i,j}-1} \sigma^2_{q_{i,j}[k],q_{i,j}[k+1]} + 2 \sum_{1\le a<b\le k} \text{Cov}(d_{q_{i,j}[a],q_{i,j}[a+1]}, d_{q_{i,j}[b],q_{i,j}[b+1]})$$

(Q3) Modeling Correlations

- **CSI-based** dissimilarities: Low correlation
  $\rightarrow$ Assume: $d|\Delta_{\mathrm{ADP}}$ **perfectly uncorrelated** to one another
- **Timestamp** dissimilarities: Displacement $d|\Delta_{\mathrm{time}}$ highly correlated to each other (unknown velocity)
  $\rightarrow$ Worst-case assumption: $d|\Delta_{\mathrm{t}}$ **perfectly correlated**
- Timestamp dissimilarities with velocity measurement: Displacement $d|\Delta_{\mathrm{time}}, v$ correlated only if velocity measurement error is correlated

(details omitted)

(Q3) Modeling Correlations

- **CSI-based** dissimilarities: Low correlation
  $\rightarrow$ Assume: $d|\Delta_{\mathrm{ADP}}$ **perfectly uncorrelated** to one another
- **Timestamp** dissimilarities: Displacement $d|\Delta_{\mathrm{time}}$ highly correlated to each other (unknown velocity)
  $\rightarrow$ Worst-case assumption: $d|\Delta_{\mathrm{t}}$ **perfectly correlated**
- Timestamp dissimilarities with velocity measurement: Displacement $d|\Delta_{\mathrm{time}}, v$ correlated only if velocity measurement error is correlated

Make reasonable worst-case assumptions!

(details omitted)

Improvement 3: Maximum Likelihood Geodesic Loss Function

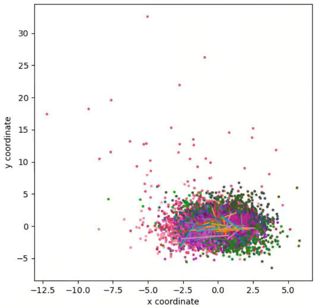
- Take variance $\sigma^2_{\mathrm{geo},i,j}$ of geodesic path length into account:

$$\mathcal{L}_{\mathrm{geo,ML}} = \frac{1}{L^2} \sum_{i,j} \frac{\left(\sum_{k=1}^{K_{i,j}^{(s)}-1} \left\| \mathbf{z}^{(q_{i,j}^{(s)}[k])} - \mathbf{z}^{(q_{i,j}^{(s)}[k+1])} \right\| - \mu_{\mathrm{geo},i,j} \right)^2}{2\sigma^2_{\mathrm{geo},i,j}}$$

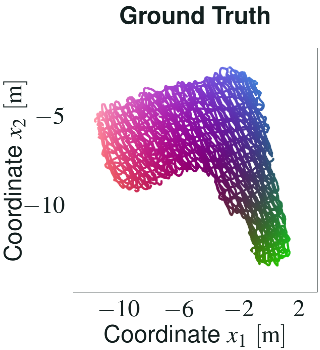
compares length of sub-sampled path through channel chart to likelihood function describing path length through manifold.

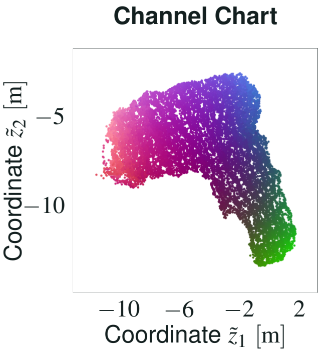
Agenda

Training Animation - All Arrays



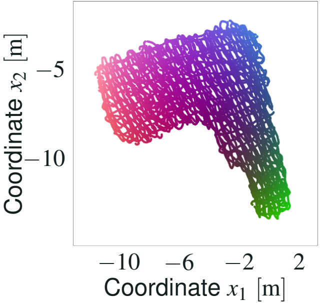
Evaluation - All Antenna Arrays



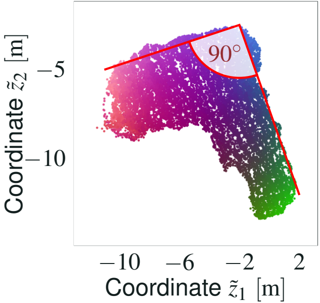


Evaluation - All Antenna Arrays

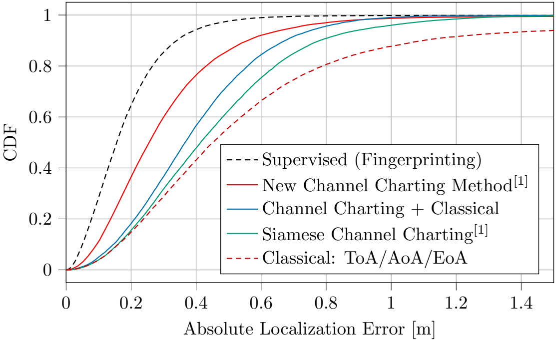
Ground Truth



Channel Chart



Evaluation - All Antenna Arrays



[1] Evaluated after optimal affine transform

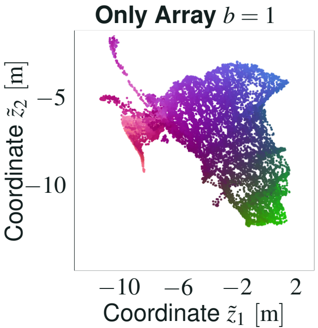
Evaluation - All Antenna Arrays

	MAE ↓[2]	DRMS ↓[2]	CEP ↓[2]	R95 ↓[2]	KS ↓[2]	CT/TW ↑[2]
Classical ToA + AoA	0.676 m	1.228 m	0.462 m	1.763 m	0.214	0.965/0.970
Siamese CC[1]	0.490 m	0.584 m	0.441 m	1.026 m	0.071	0.996/0.996
Classical + CC	0.401 m	0.483 m	0.369 m	0.789 m	0.070	0.995/0.995
New Method[1]	**0.295** m	**0.375** m	**0.234** m	**0.719** m	**0.069**	**0.998/0.998**

[1] MAE / DRMS / CEP / R95 evaluated after optimal affine transform

[2] MAE = mean absolute error, DRMS = distance root mean squared, CEP = circular error probable, R95 = 95^{th} error percentile, KS = Kruskal Stress, CT/TW = Continuity / Trustworthiness

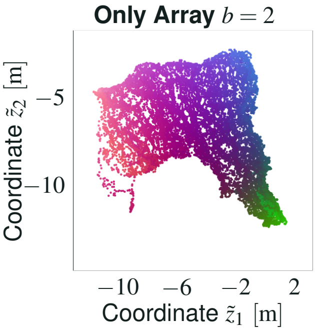
Evaluation - Single Antenna Array



Ant. Array	MAE [1]	MAE Siam. [1]
$b = 1$	$\approx 1.05\,\mathrm{m}$	$\approx 2.04\,\mathrm{m}$
$b = 2$		$\approx 2.23\,\mathrm{m}$
$b = 3$		$\approx 1.37\,\mathrm{m}$
$b = 4$		$\approx 1.62\,\mathrm{m}$

[1]Evaluated after optimal affine transform

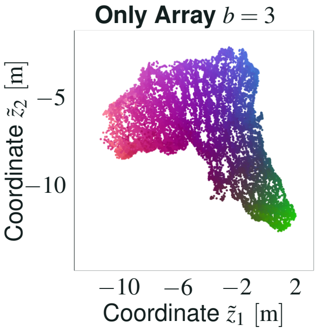
Evaluation - Single Antenna Array



Ant. Array	MAE [1]	MAE Siam. [1]
$b = 1$	$\approx 1.05\,\mathrm{m}$	$\approx 2.04\,\mathrm{m}$
$b = 2$	$\approx 0.44\,\mathrm{m}$	$\approx 2.23\,\mathrm{m}$
$b = 3$		$\approx 1.37\,\mathrm{m}$
$b = 4$		$\approx 1.62\,\mathrm{m}$

[1]Evaluated after optimal affine transform

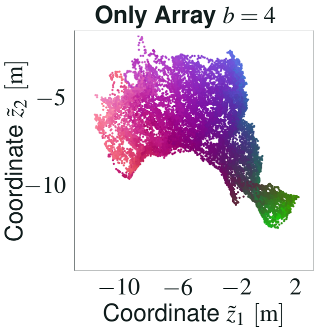
Evaluation - Single Antenna Array



Ant. Array	MAE [1]	MAE Siam. [1]
$b = 1$	$\approx 1.05\,\mathrm{m}$	$\approx 2.04\,\mathrm{m}$
$b = 2$	$\approx 0.44\,\mathrm{m}$	$\approx 2.23\,\mathrm{m}$
$b = 3$	$\approx 0.44\,\mathrm{m}$	$\approx 1.37\,\mathrm{m}$
$b = 4$		$\approx 1.62\,\mathrm{m}$

[1]Evaluated after optimal affine transform

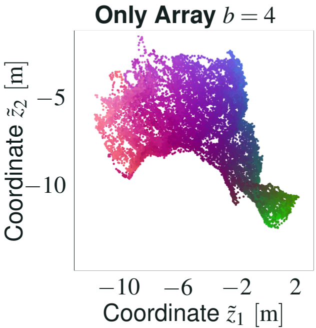
Evaluation - Single Antenna Array



Ant. Array	MAE [1]	MAE Siam. [1]
$b = 1$	$\approx 1.05\,\mathrm{m}$	$\approx 2.04\,\mathrm{m}$
$b = 2$	$\approx 0.44\,\mathrm{m}$	$\approx 2.23\,\mathrm{m}$
$b = 3$	$\approx 0.44\,\mathrm{m}$	$\approx 1.37\,\mathrm{m}$
$b = 4$	$\approx 0.73\,\mathrm{m}$	$\approx 1.62\,\mathrm{m}$

[1] Evaluated after optimal affine transform

Evaluation - Single Antenna Array

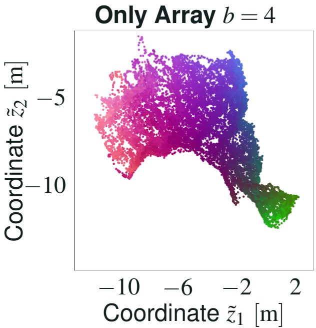


Ant. Array	MAE [1]	MAE Siam. [1]
$b = 1$	$\approx 1.05\,\mathrm{m}$	$\approx 2.04\,\mathrm{m}$
$b = 2$	$\approx 0.44\,\mathrm{m}$	$\approx 2.23\,\mathrm{m}$
$b = 3$	$\approx 0.44\,\mathrm{m}$	$\approx 1.37\,\mathrm{m}$
$b = 4$	$\approx 0.73\,\mathrm{m}$	$\approx 1.62\,\mathrm{m}$

$\rightarrow$ Good single array LoS localization!

[1] Evaluated after optimal affine transform

Evaluation - Single Antenna Array



Ant. Array	MAE [1]	MAE Siam. [1]
$b = 1$	$\approx 1.05\,\mathrm{m}$	$\approx 2.04\,\mathrm{m}$
$b = 2$	$\approx 0.44\,\mathrm{m}$	$\approx 2.23\,\mathrm{m}$
$b = 3$	$\approx 0.44\,\mathrm{m}$	$\approx 1.37\,\mathrm{m}$
$b = 4$	$\approx 0.73\,\mathrm{m}$	$\approx 1.62\,\mathrm{m}$

$\rightarrow$ Good single array LoS localization!
$\rightarrow$ Some NLoS localization capability!

[1] Evaluated after optimal affine transform

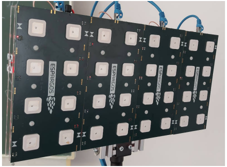
Agenda

1. Introduction, Motivation, State of the Art
2. New Training Architecture
3. Geodesic Loss
4. Uncertainty (Concept Overview)
5. Evaluation and Results
6. Conclusion and Outlook

Conclusion and Outlook

- New batch-wise training architecture
- New geodesic loss taking into account uncertainty
- Significant localization performance improvements!

- Uncertainty model: Integrate additional information sources (e.g., velocity measurements)
- Reproduce results on other CSI datasets (WiFi/ESPARGOS, DICHASUS, third-party)



Thank you for your attention! Questions?

Source Code (GitHub)

Dataset (DICHASUS)

References

[Budninskiy et al., 2019] Budninskiy, M., Yin, G., Feng, L., Tong, Y., and Desbrun, M. (2019).
Parallel transport unfolding: a connection-based manifold learning approach.
SIAM Journal on Applied Algebra and Geometry, 3(2):266–291.

[Chaaya et al., 2024] Chaaya, C. B., Girgis, A. M., and Bennis, M. (2024).
Learning Latent Wireless Dynamics from Channel State Information.

[Euchner et al., 2023] Euchner, F., Stephan, P., and ten Brink, S. (2023).
Augmenting Channel Charting with Classical Wireless Source Localization Techniques.

[Euchner et al., 2024] Euchner, F., Stephan, P., and Ten Brink, S. (2024).
Leveraging the doppler effect for channel charting.
In *2024 IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 731–735.

[Hermans et al., 2017] Hermans, A., Beyer, L., and Leibe, B. (2017).
In Defense of the Triplet Loss for Person Re-Identification.
arXiv preprint arXiv:1703.07737.

[Rappaport et al., 2021] Rappaport, B., Gönültaş, E., Hoydis, J., Arnold, M., Srinath, P. K., and Studer, C. (2021).
Improving Channel Charting using a Split Triplet Loss and an Inertial Regularizer.
In *2021 17th International Symposium on Wireless Communication Systems (ISWCS)*, pages 1–6.

[Rosman et al., 2010] Rosman, G., Bronstein, M. M., Bronstein, A. M., and Kimmel, R. (2010).
Nonlinear dimensionality reduction by topologically constrained isometric embedding.
International Journal of Computer Vision, 89(1):56–68.

[Schwartz and Talmon, 2019] Schwartz, A. and Talmon, R. (2019).
Intrinsic isometric manifold learning with application to localization.
SIAM Journal on Imaging Sciences, 12(3):1347–1391.

[Sohn, 2016] Sohn, K. (2016).
Improved Deep Metric Learning with Multi-class N-pair Loss Objective.
Advances in neural information processing systems, 29.

[Stahlke et al., 2023] Stahlke, M., Yammine, G., Feigl, T., Eskofier, B. M., and Mutschler, C. (2023).
Velocity-Based Channel Charting with Spatial Distribution Map Matching.

[Stephan et al., 2023] Stephan, P., Euchner, F., and Ten Brink, S. (2023).
Angle-Delay Profile-Based and Timestamp-Aided Dissimilarity Metrics for Channel Charting.
IEEE Transactions on Communications.

[Studer et al., 2018] Studer, C., Medjkouh, S., Gonultaş, E., Goldstein, T., and Tirkkonen, O. (2018).
Channel Charting: Locating Users Within the Radio Environment Using Channel State Information.
IEEE Access, 6:47682–47698.

[Taner et al., 2023] Taner, S., Guillaud, M., Tirkkonen, O., and Studer, C. (2023).
Channel Charting for Streaming CSI Data.

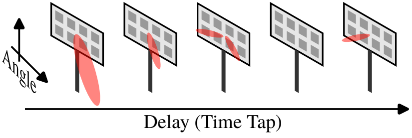
[Taner et al., 2024] Taner, S., Palhares, V., and Studer, C. (2024).
Channel Charting in Real-World Coordinates with Distributed MIMO.

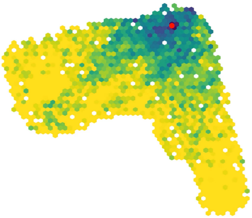
[Vindas and Guillaud, 2024] Vindas, Y. and Guillaud, M. (2024).
Multi-Site Wireless Channel Charting Through Latent Space Alignment.
In *2024 IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 826–830. IEEE.

[Zha and Zhang, 2003] Zha, H. and Zhang, Z. (2003).
Isometric embedding and continuum isomap.
In *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pages 864–871.

Dissimilarity Metric: Angle-Delay-Profile Δ_{ADP}

- Same power from same angle at same delay $\rightarrow$ similar location





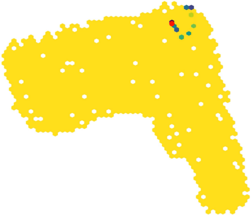
- Uses *time-domain* CSI $\mathbf{H} \in \mathbb{C}^{B \times M \times T}$:

$$\Delta_{\text{ADP},i,j} = \sum_{b=1}^{B} \sum_{\tau=1}^{T} \left(1 - \frac{\left| \sum_{m=1}^{M} \left(\mathbf{H}_{b,m,\tau}^{(i)} \right)^* \mathbf{H}_{b,m,\tau}^{(j)} \right|^2}{\left(\sum_{m=1}^{M} \left| \mathbf{H}_{b,m,\tau}^{(i)} \right|^2 \right) \left(\sum_{m=1}^{M} \left| \mathbf{H}_{b,m,\tau}^{(j)} \right|^2 \right)} \right)$$

Dissimilarity Metric: Timestamp Difference Δ_{time}

- Use time difference as dissimilarity

$$\Delta_{\mathrm{time},i,j} = \left| t^{(i)} - t^{(j)} \right|$$



Fusing Dissimilarity Metrics

- Metrics may be combined into a unified dissimilarity matrix, e.g.:

$$\Delta_{\text{fuse},i,j} = \min\left\{\Delta_{\text{ADP},i,j}, \Delta_{\text{time},i,j}\right\}$$

- Requires scaling (e.g., express all dissimilarities in meters)