



Gaussian Process Summer School

Manchester
14/09/2026

Bayesian non-parametric dynamical clustering of time series

Adrián Pérez Herrero



Outline

Motivation

Hypothesis

Background

The HDP-GPC model

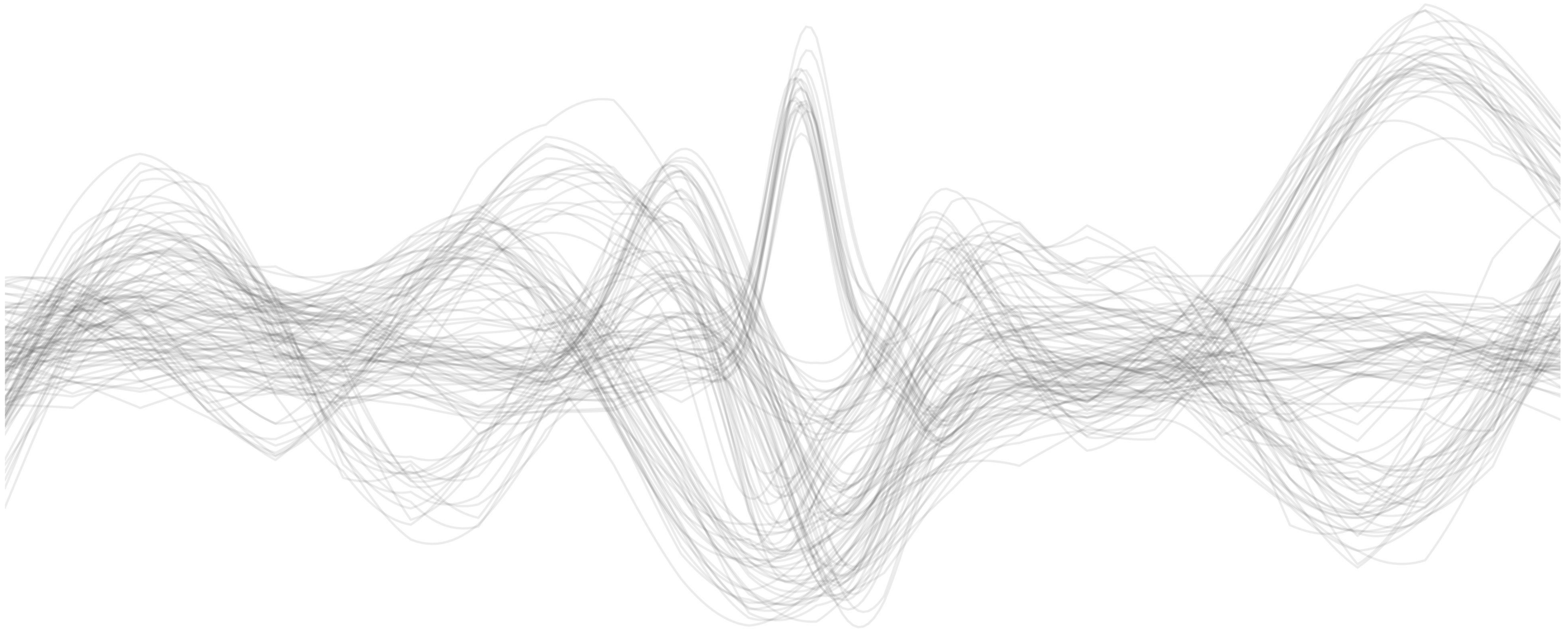
Applications

Conclusions

Motivation

Motivation

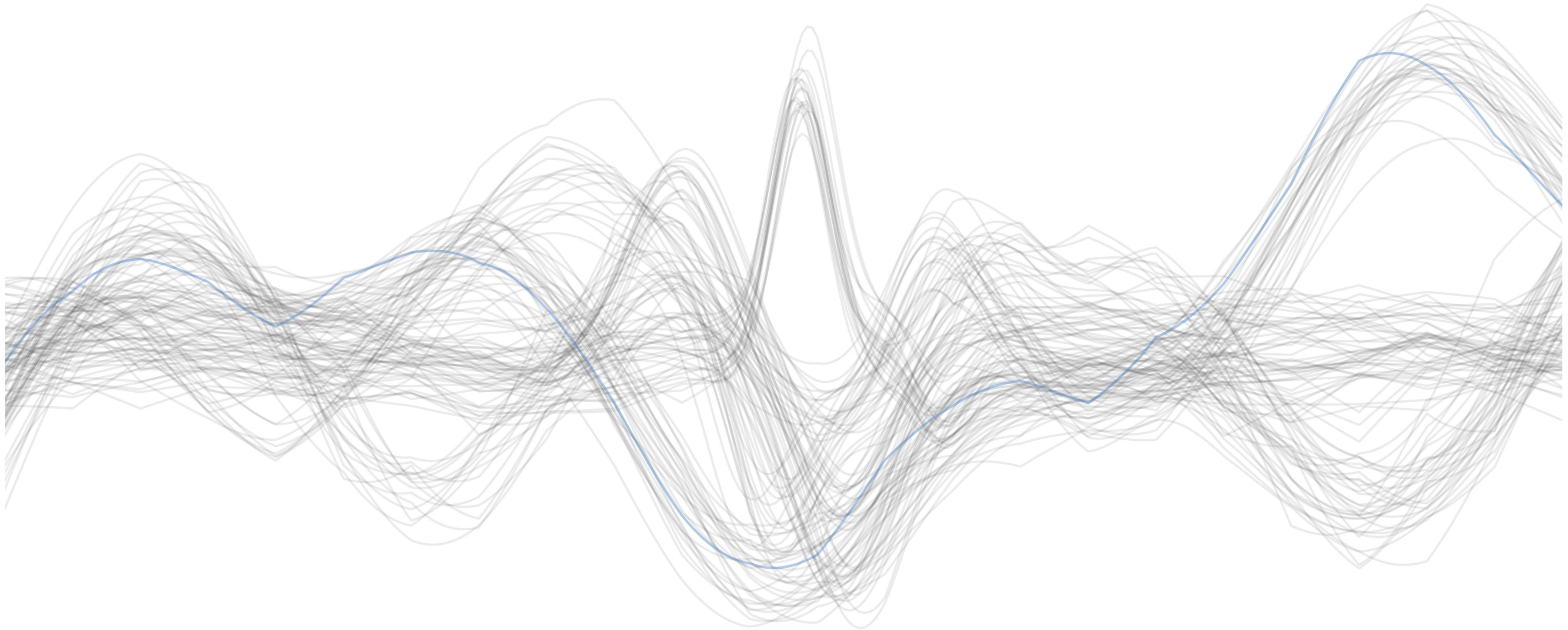
Why clustering?



Motivation

Why clustering?

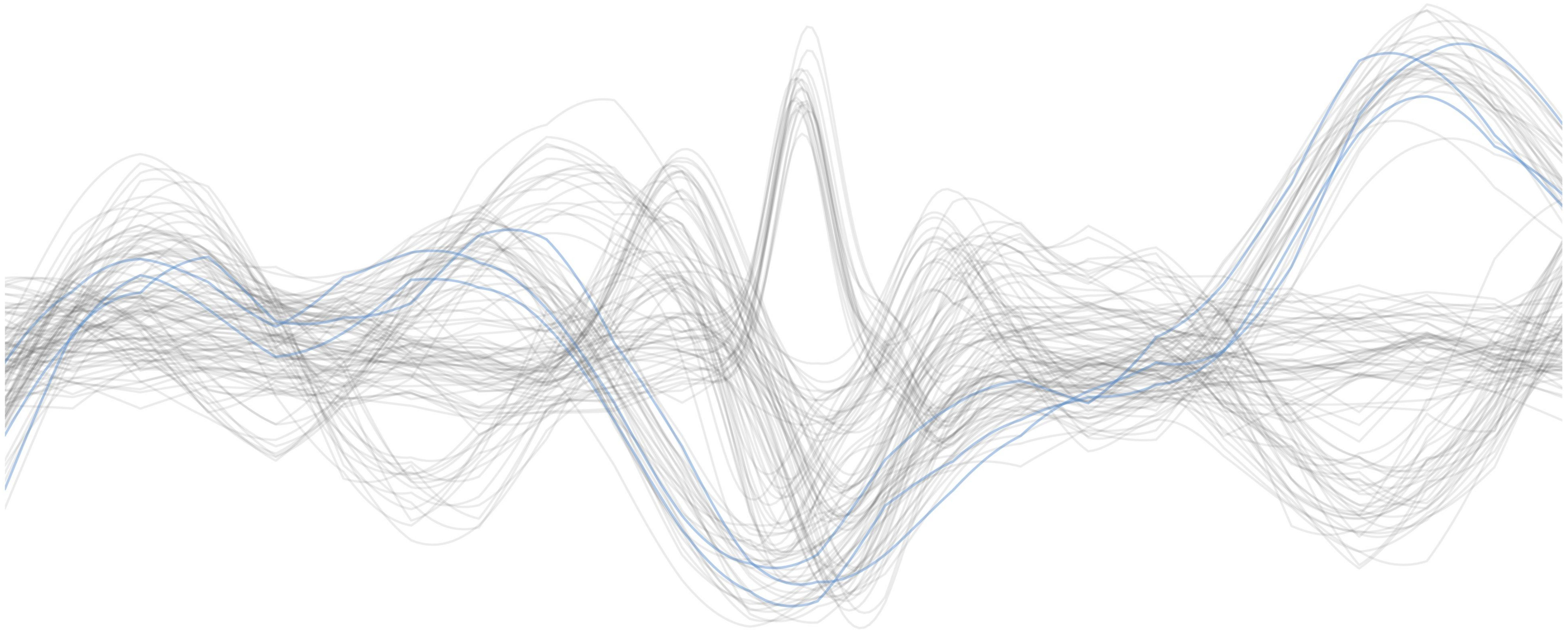
1



Motivation

Why clustering?

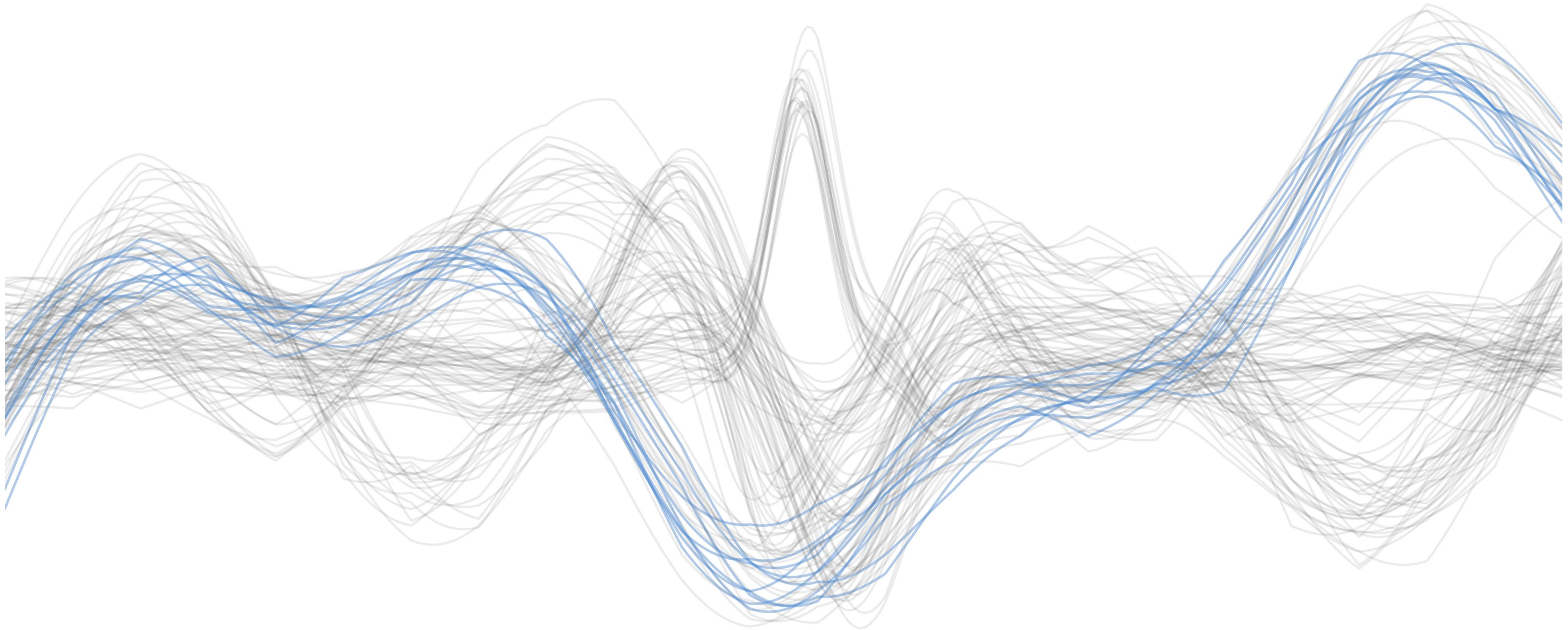
1



Motivation

Why clustering?

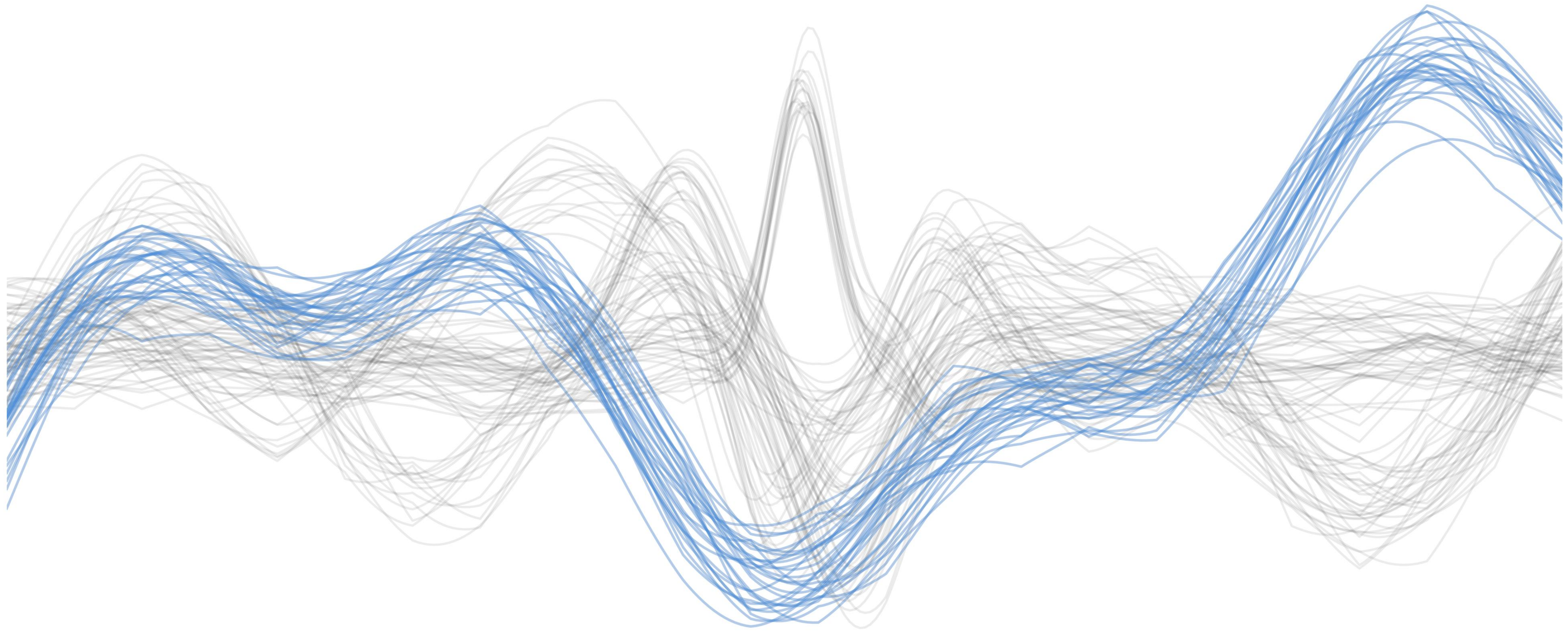
1



Motivation

Why clustering?

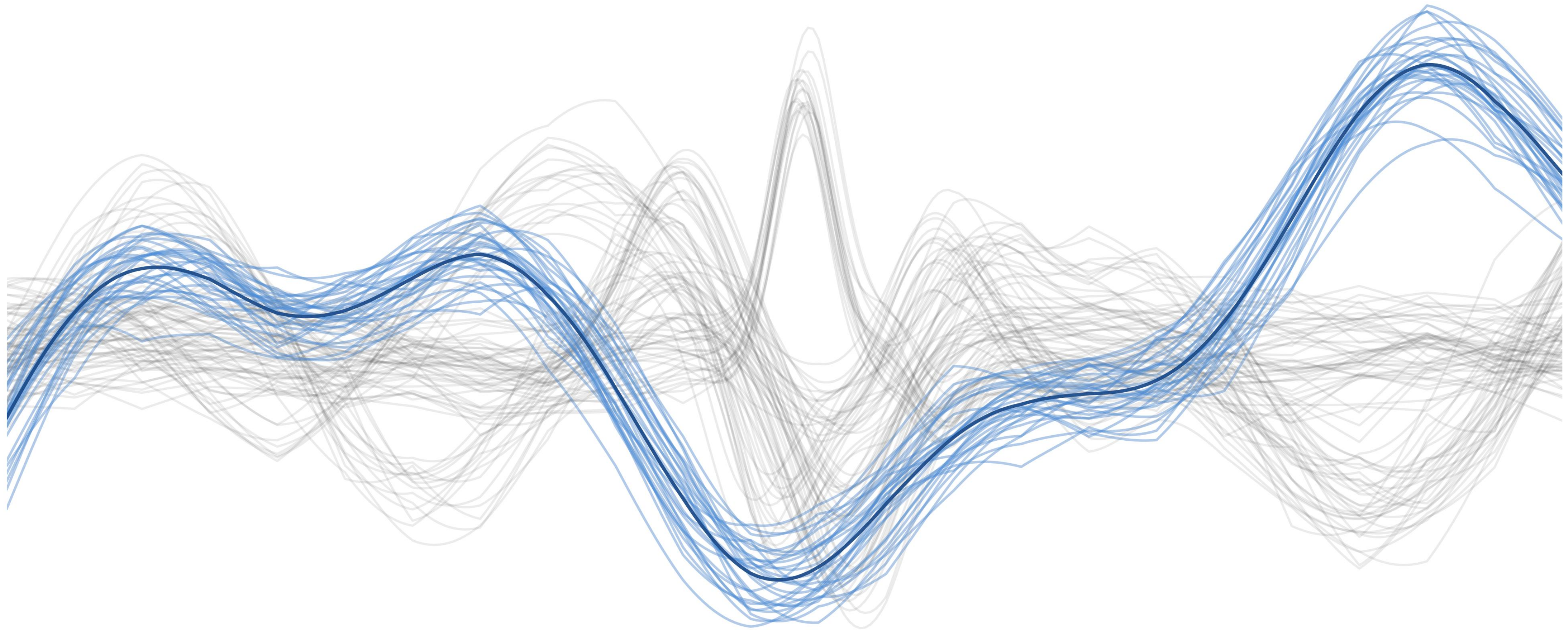
1



Motivation

Why clustering?

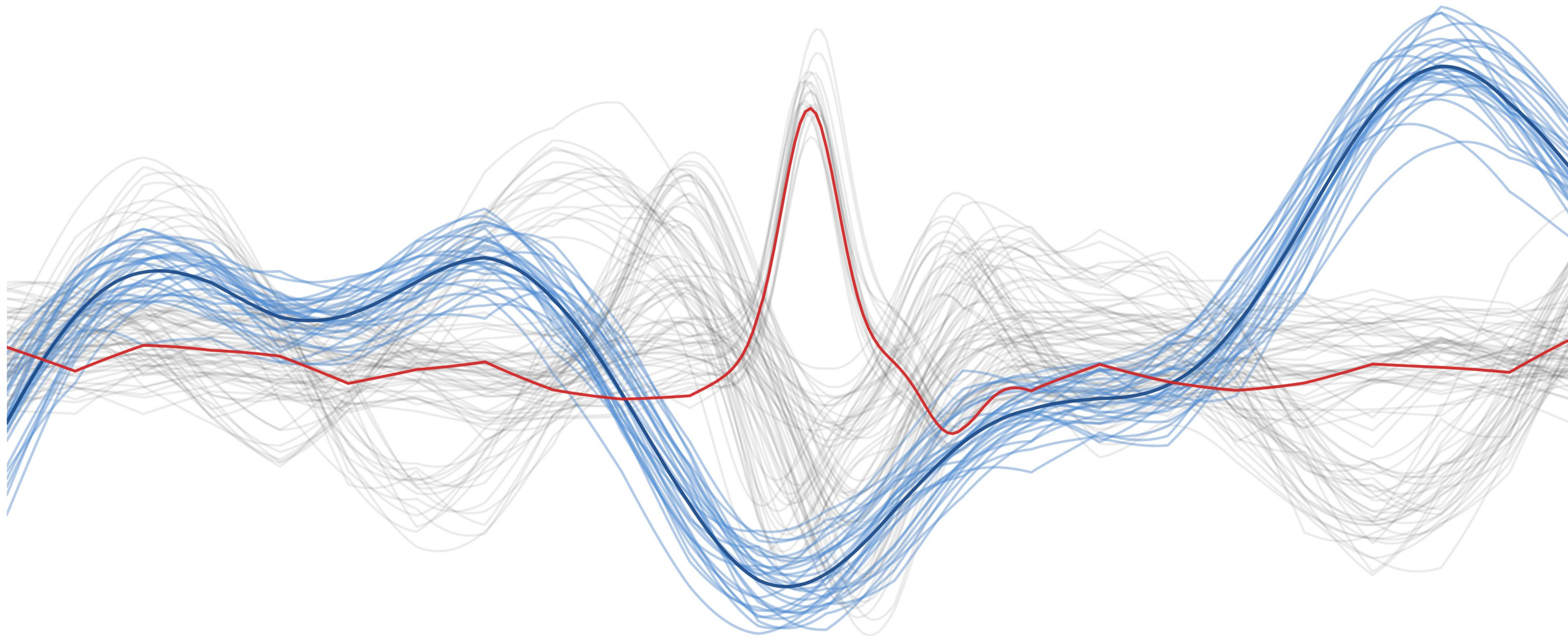
1



Motivation

Why clustering?

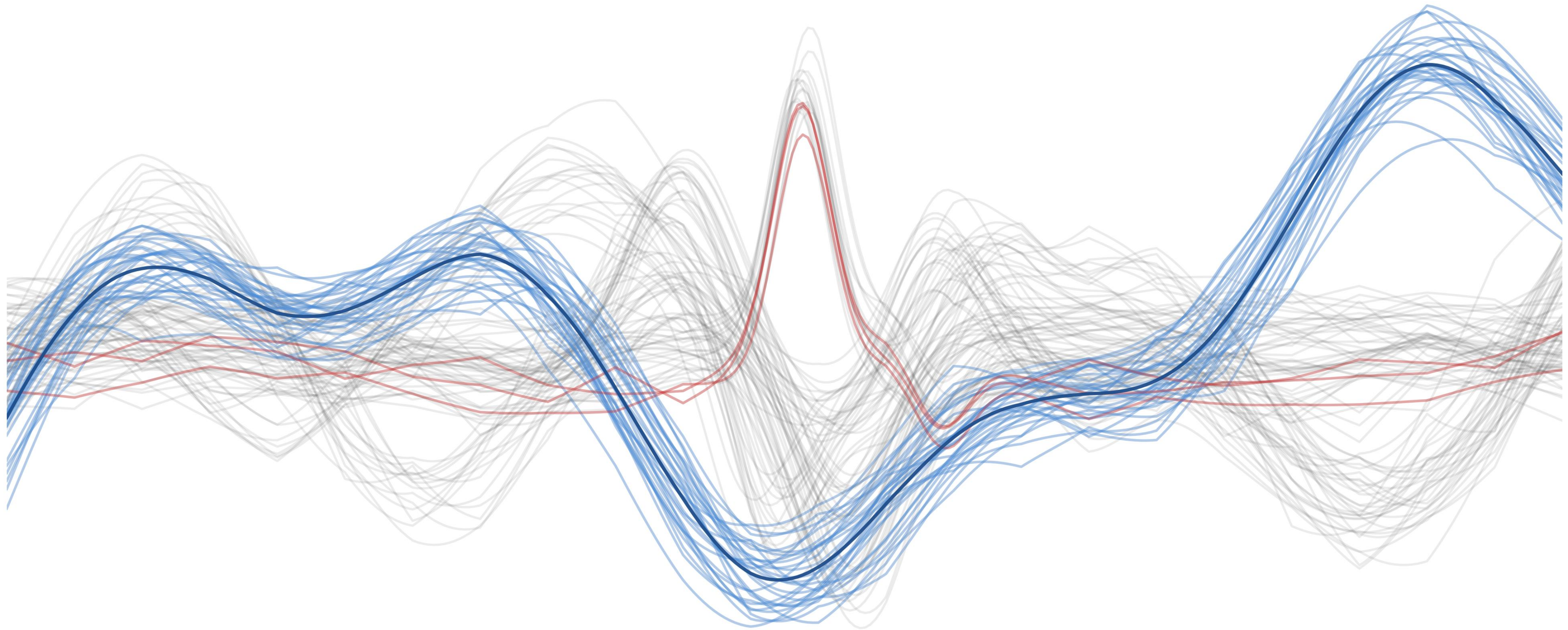
1



Motivation

Why clustering?

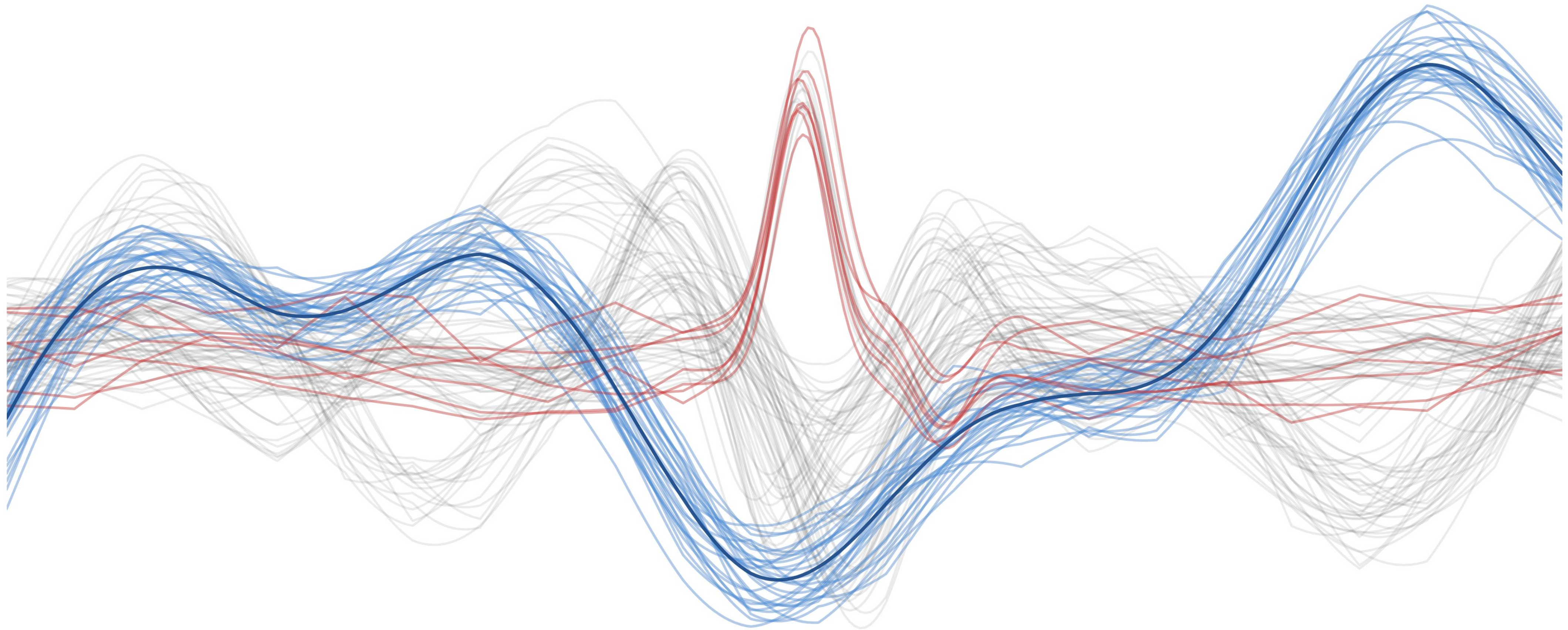
1



Motivation

Why clustering?

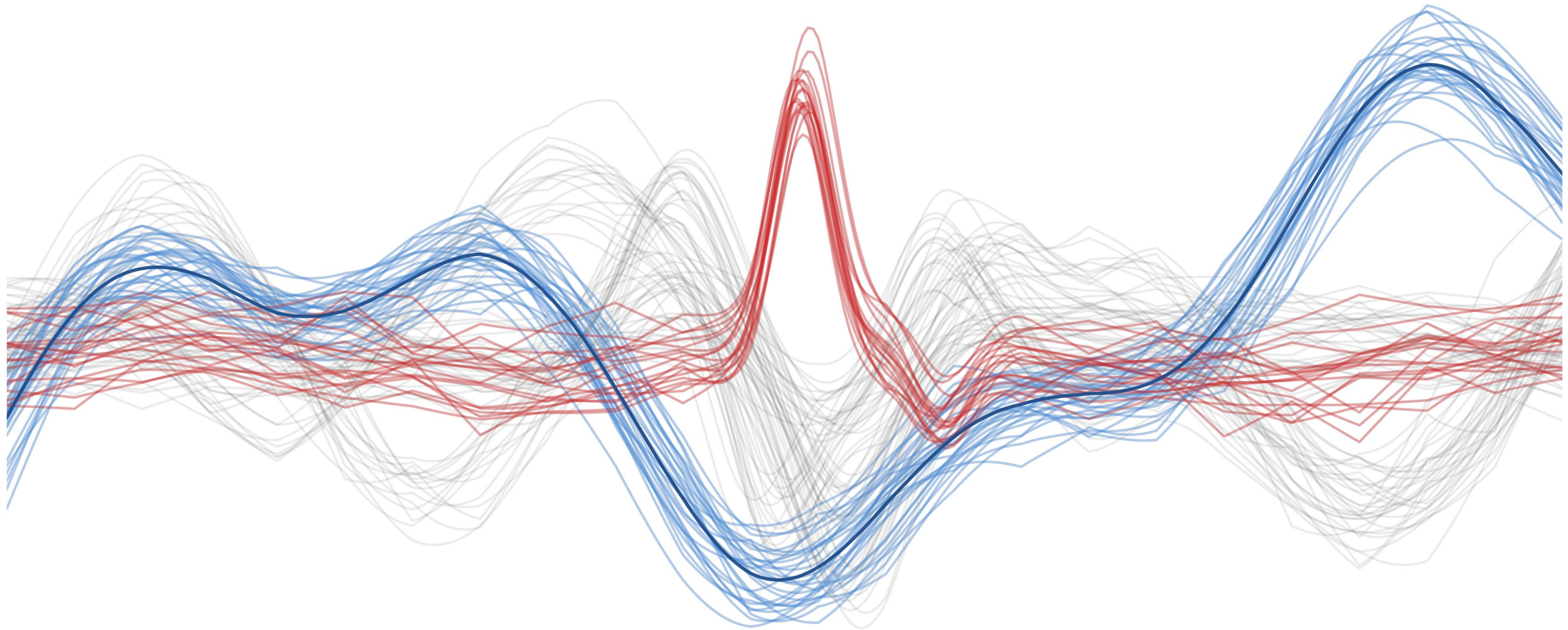
1



Motivation

Why clustering?

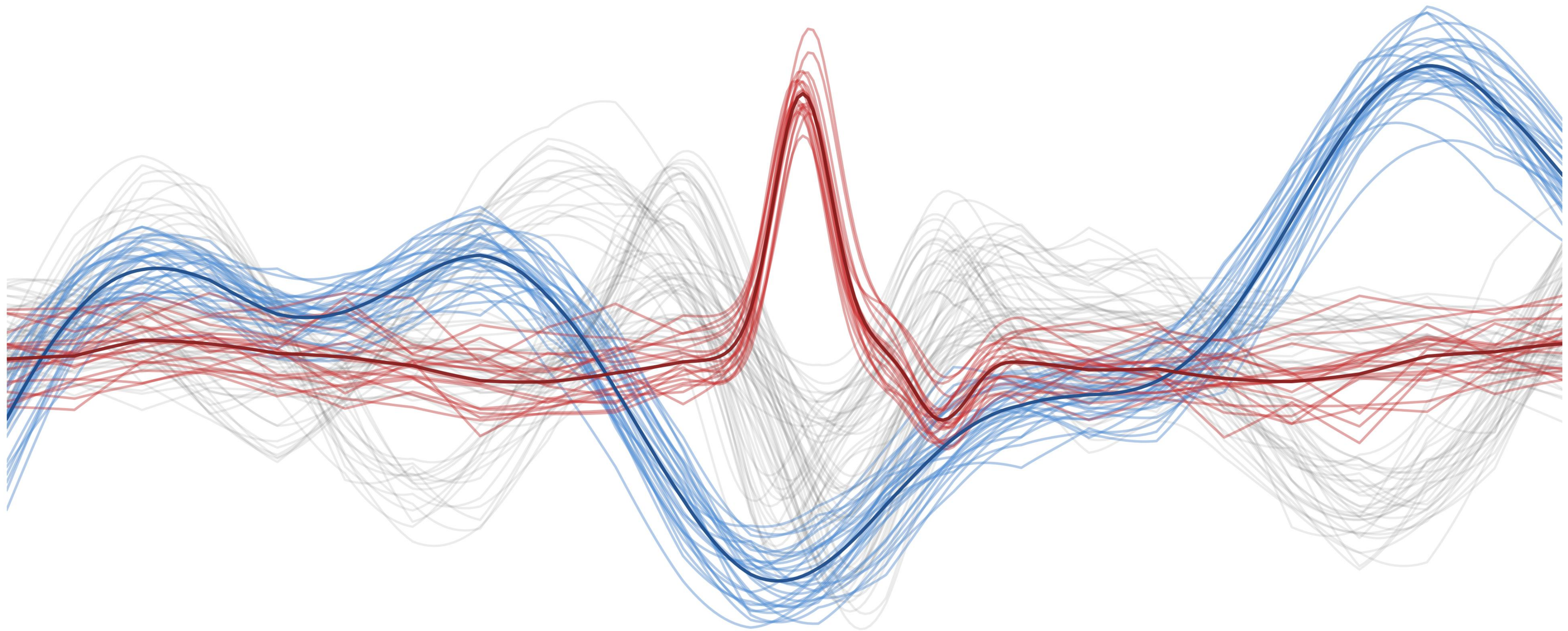
1



Motivation

Why clustering?

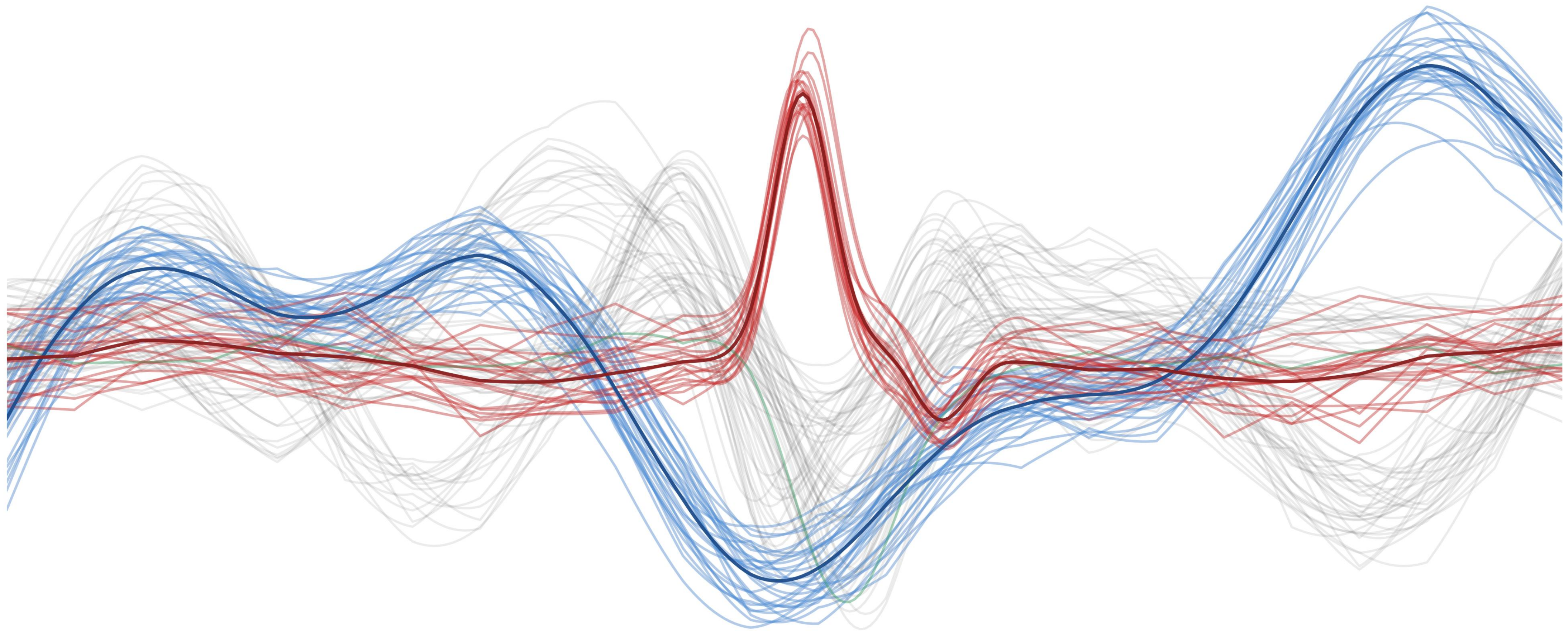
1



Motivation

Why clustering?

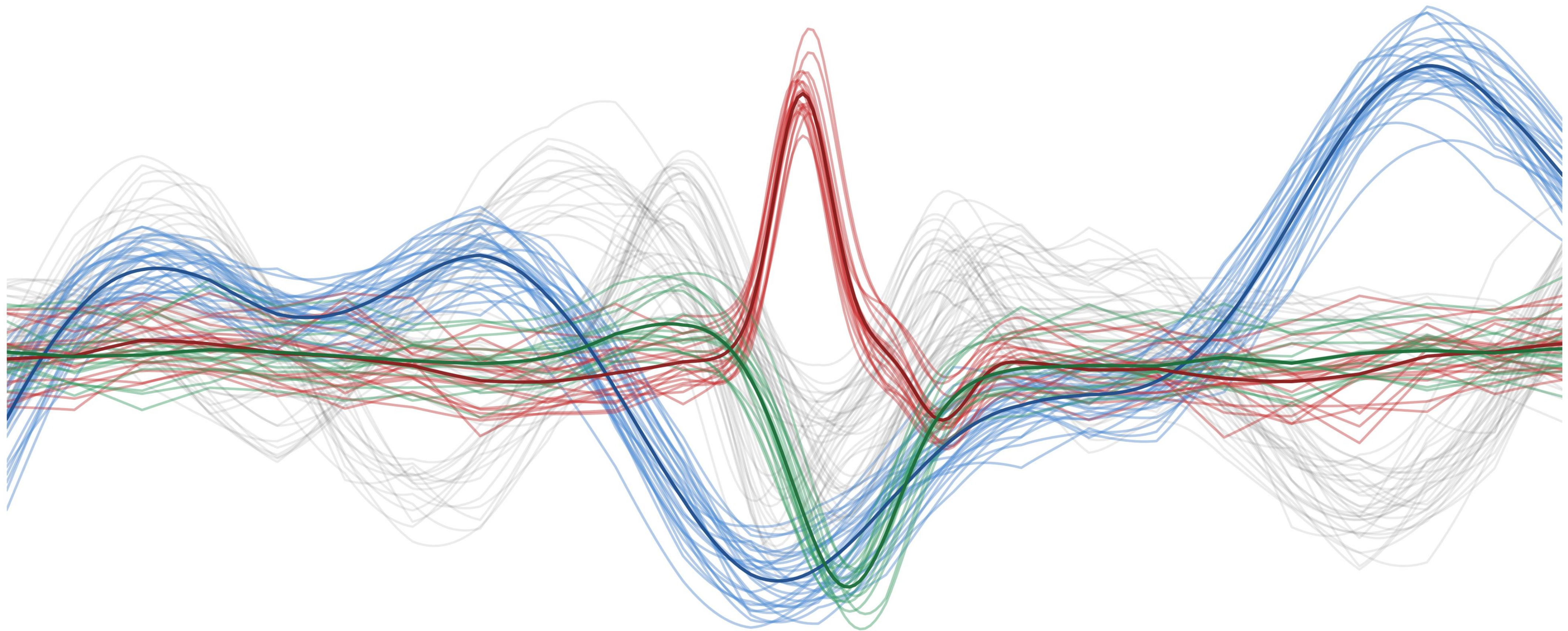
1



Motivation

Why clustering?

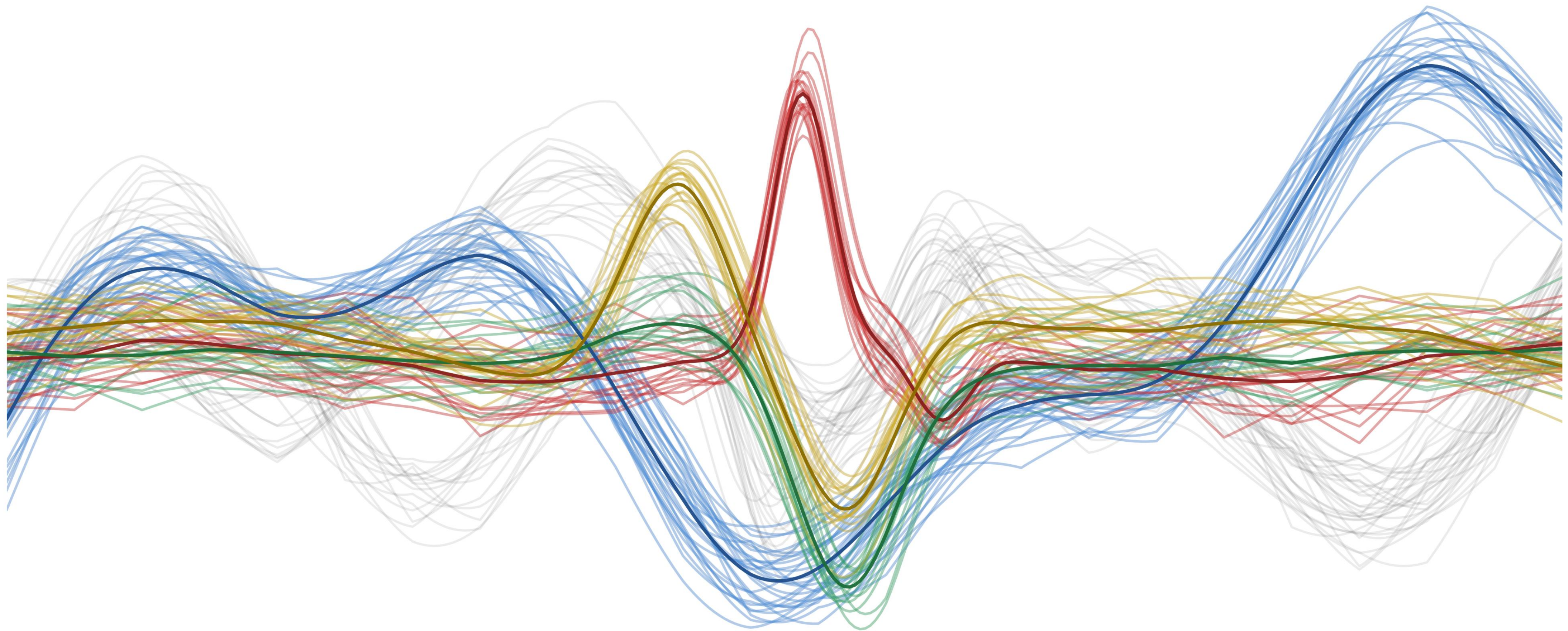
1



Motivation

Why clustering?

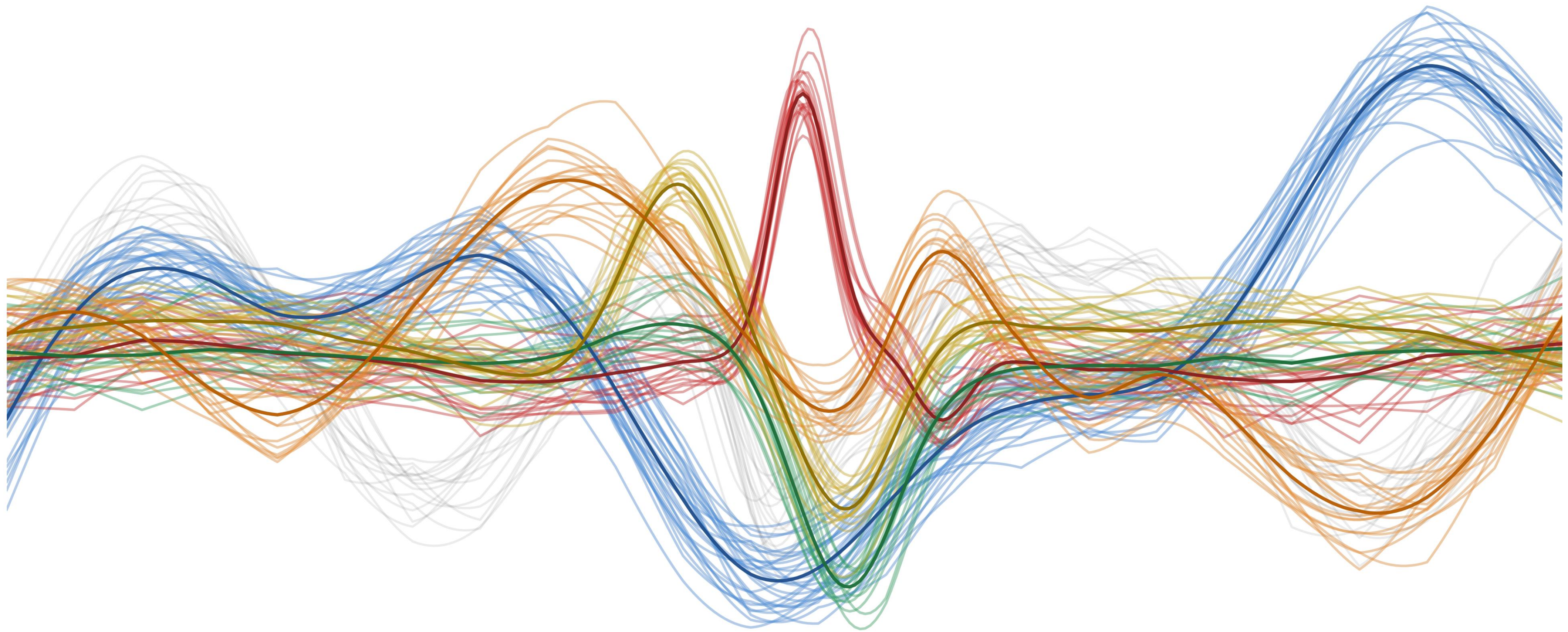
1



Motivation

Why clustering?

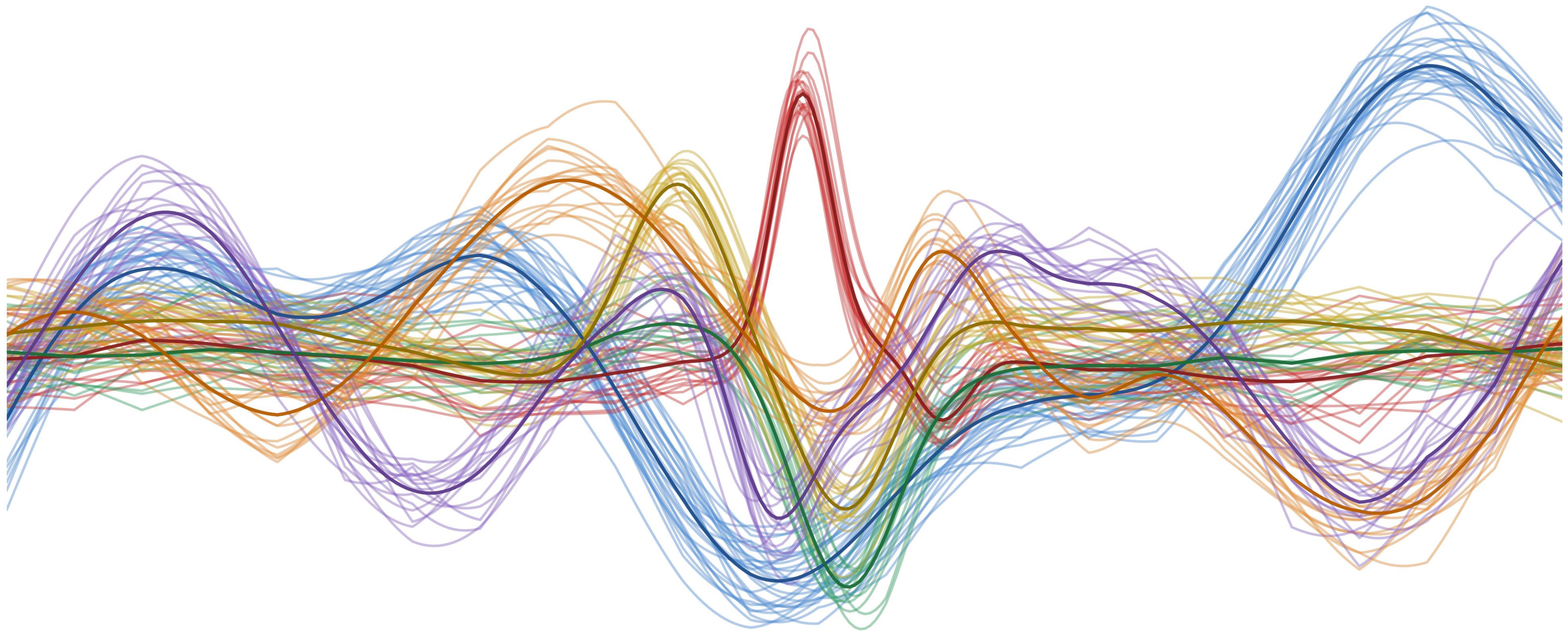
1



Motivation

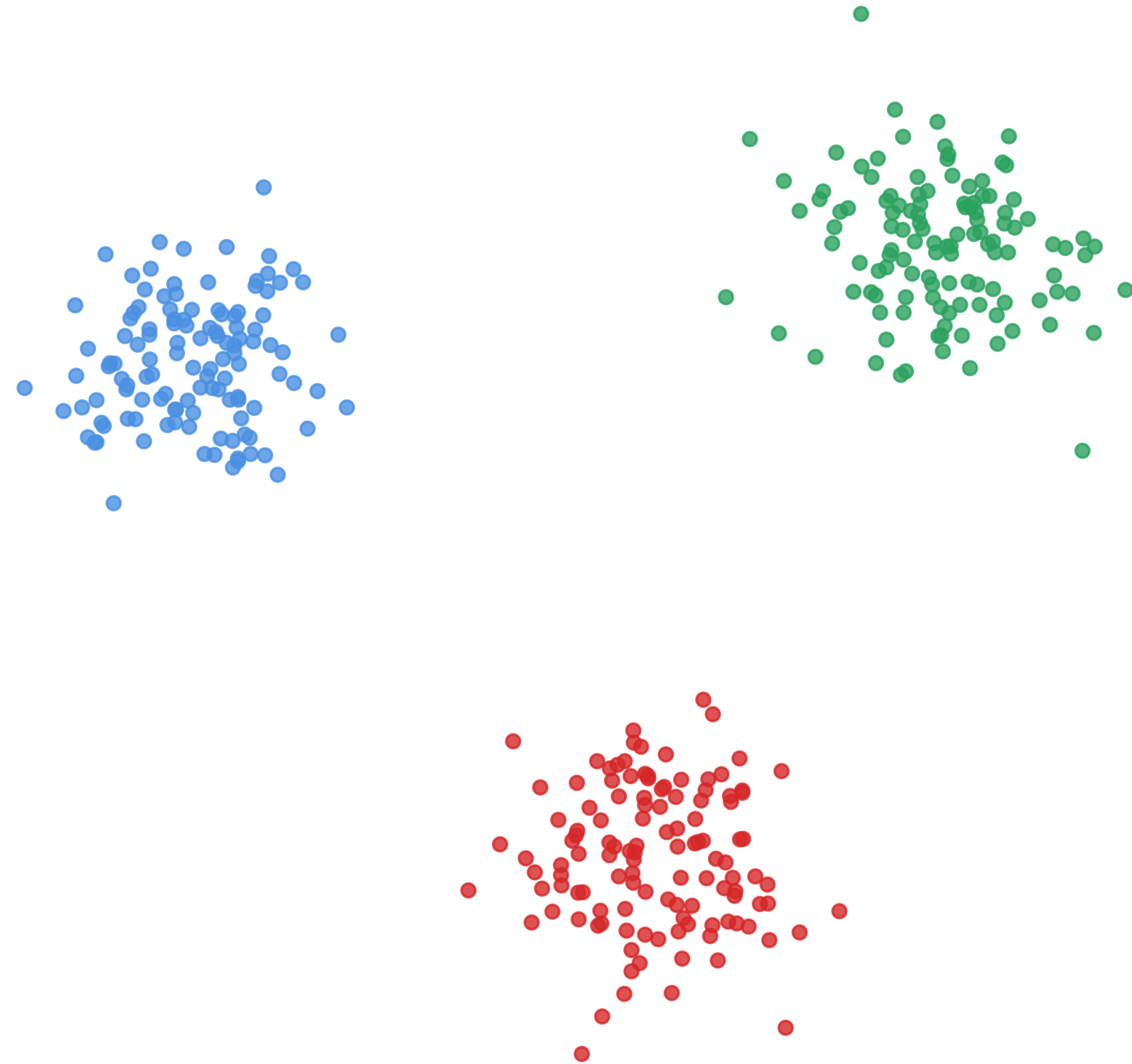
Why clustering?

1



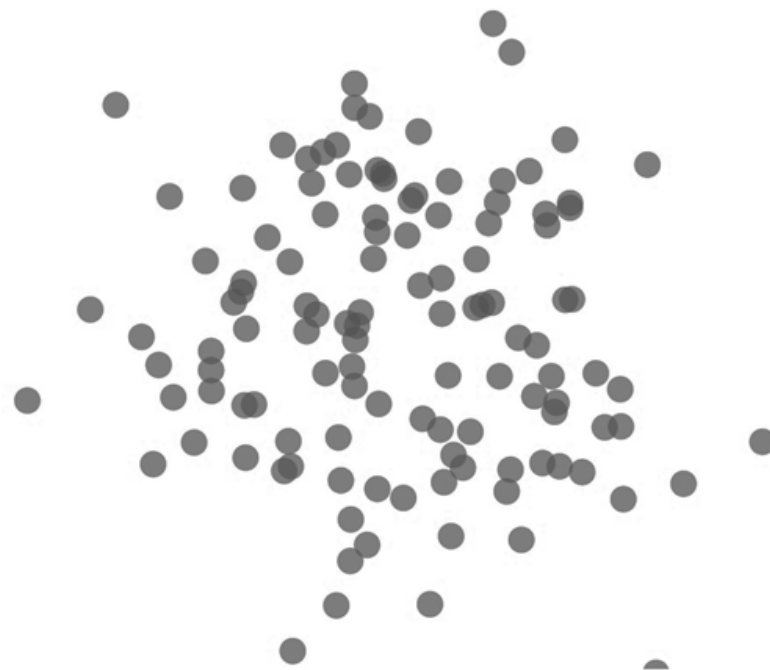
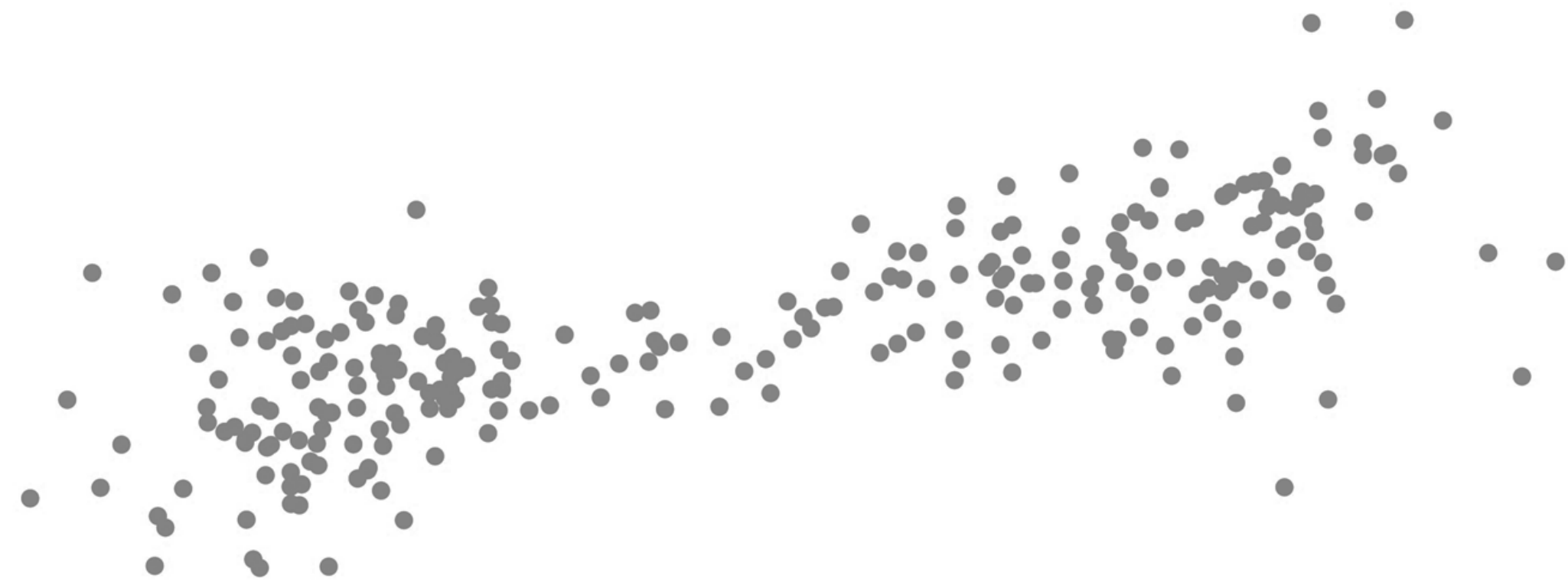
Motivation

Clustering: classic approaches (i.i.d)



Motivation

Clustering: classic approaches (i.i.d)



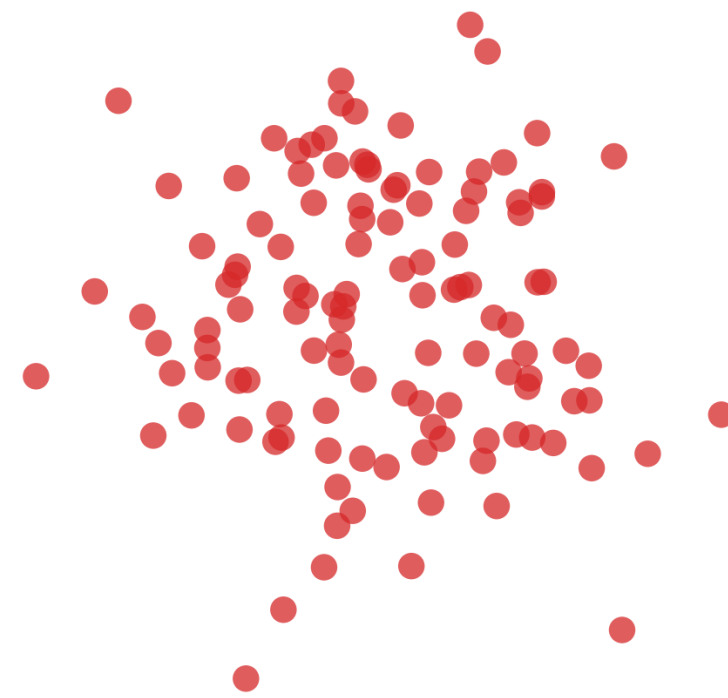
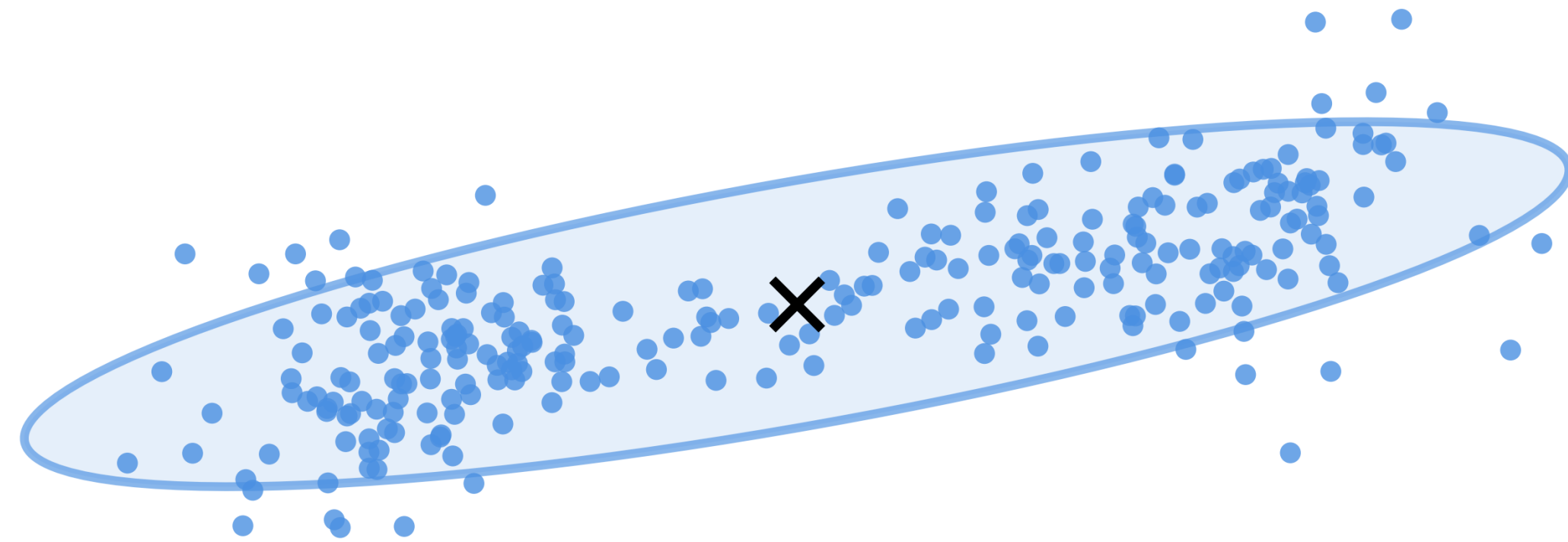
Motivation

Clustering: classic approaches (i.i.d)



Motivation

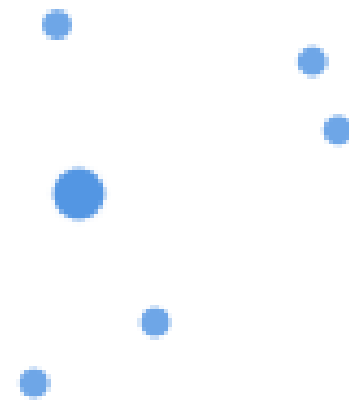
Clustering: classic approaches (i.i.d)



Motivation

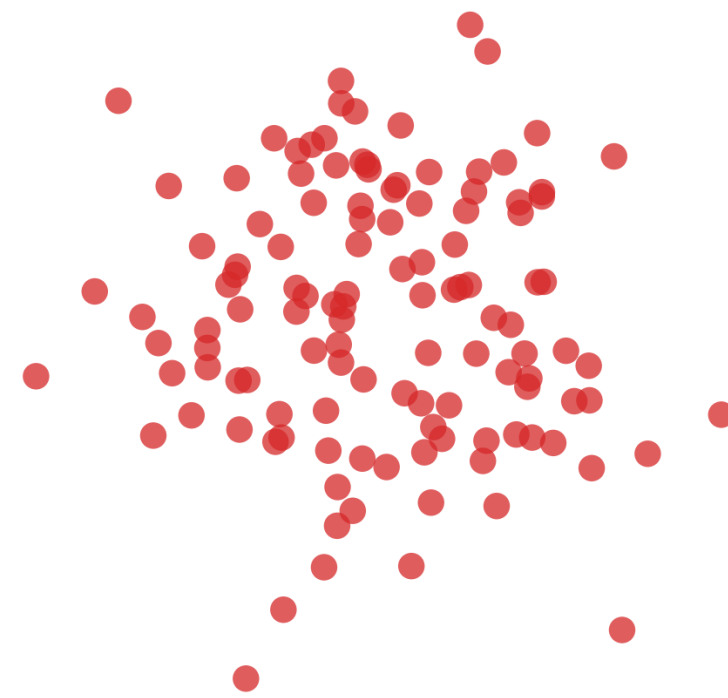
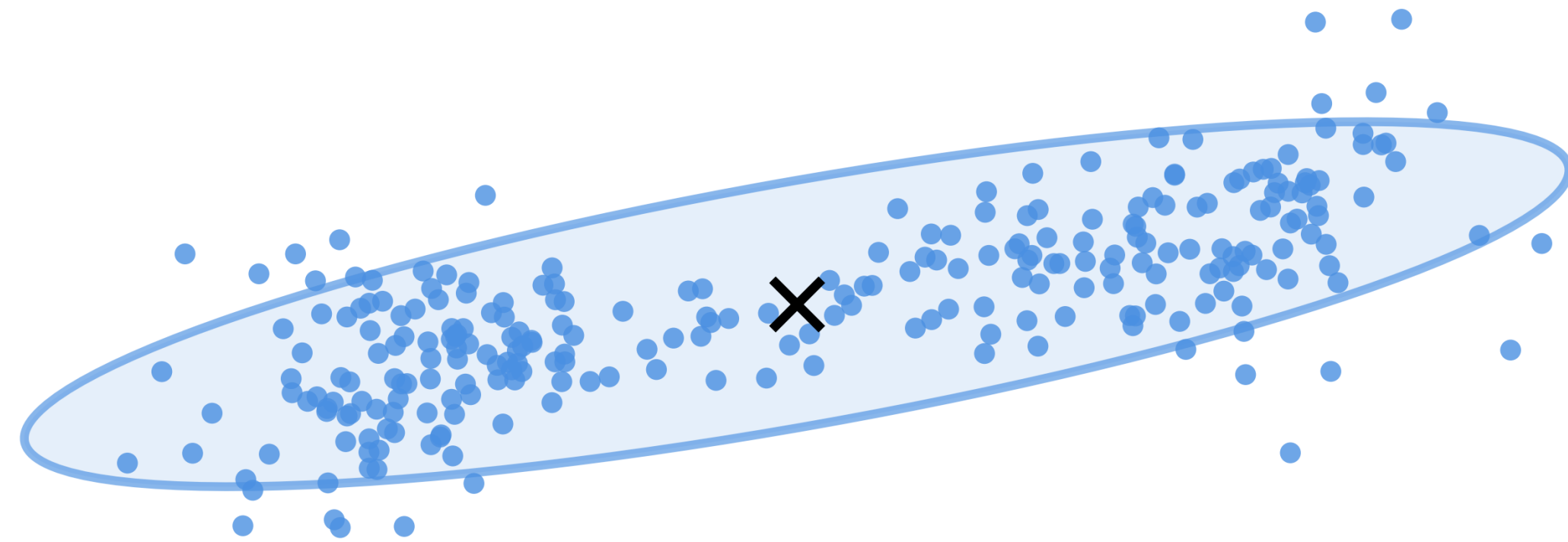
Dynamical clustering

2



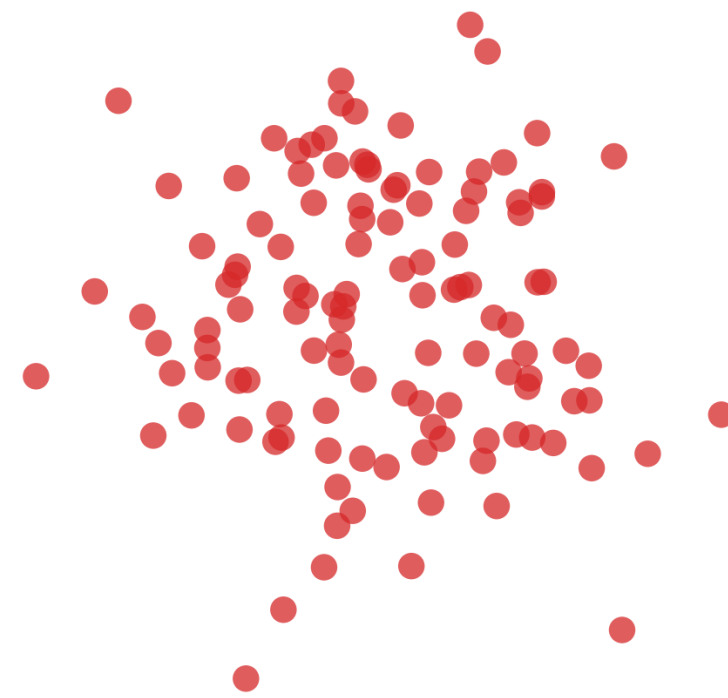
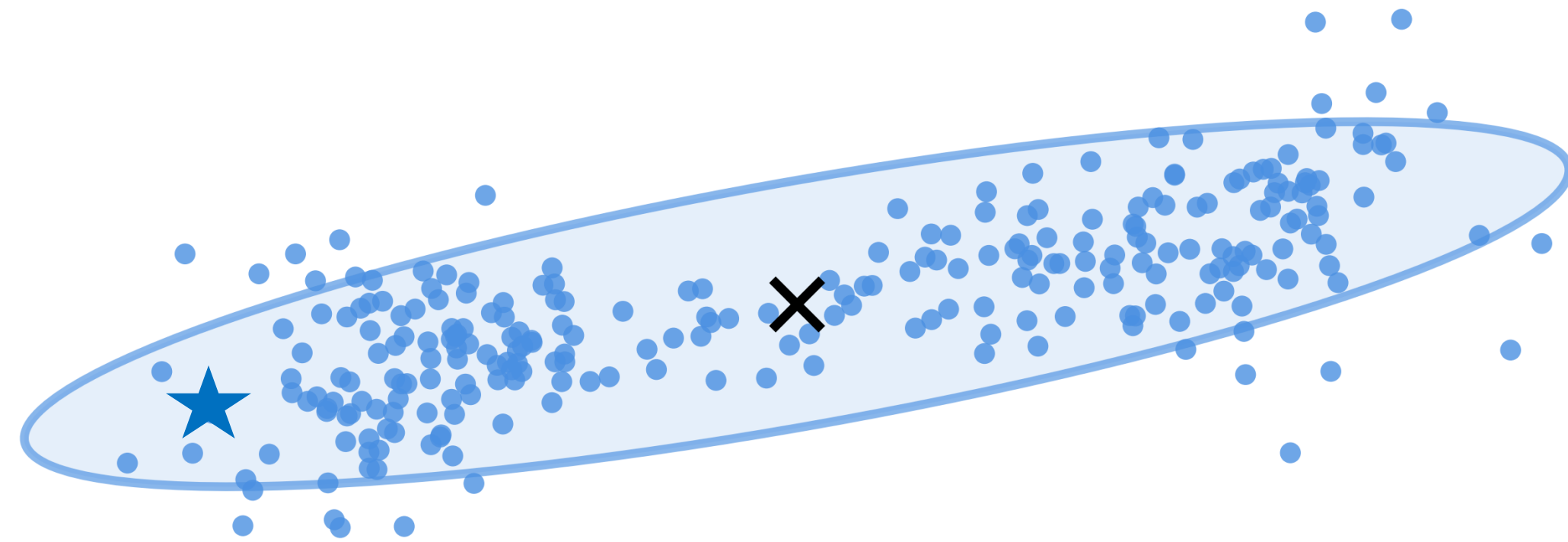
Motivation

Clustering: classic approaches (i.i.d)



Motivation

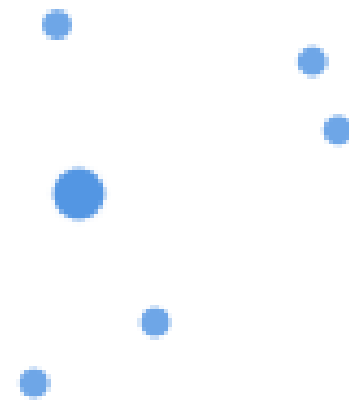
Clustering: classic approaches (i.i.d)



Motivation

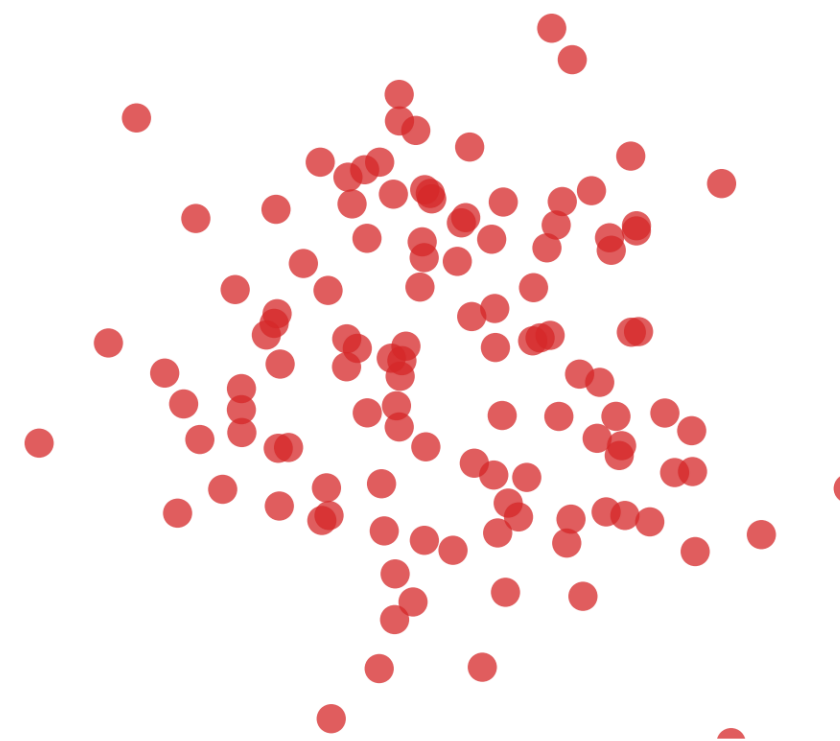
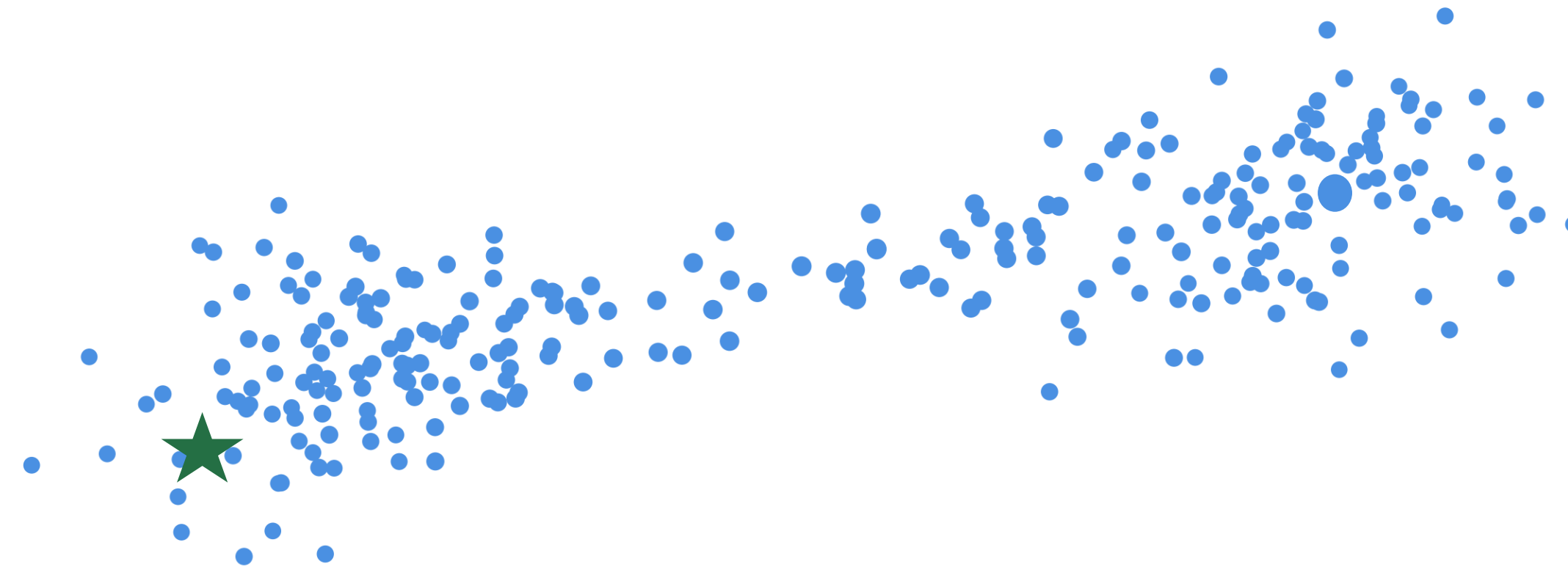
Dynamical clustering

2



Motivation

Dynamical clustering



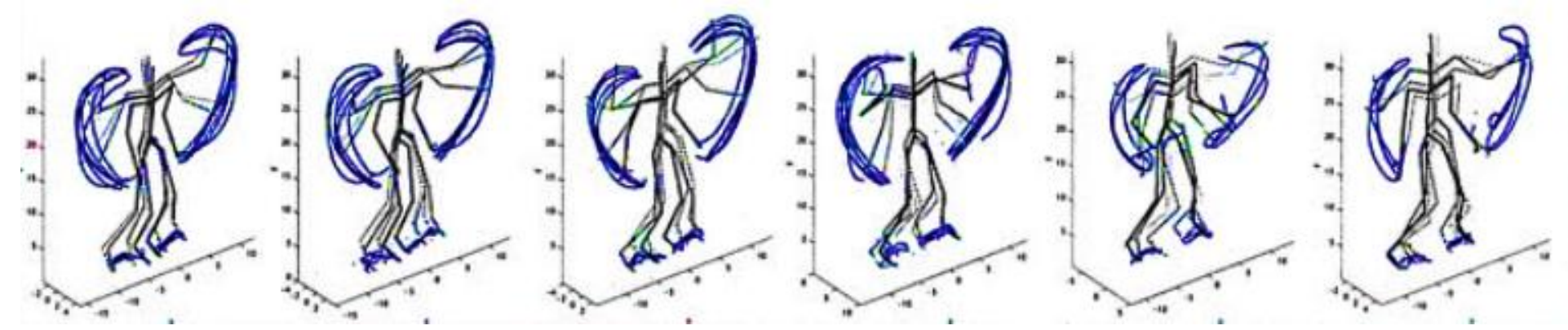
Motivation

Time series

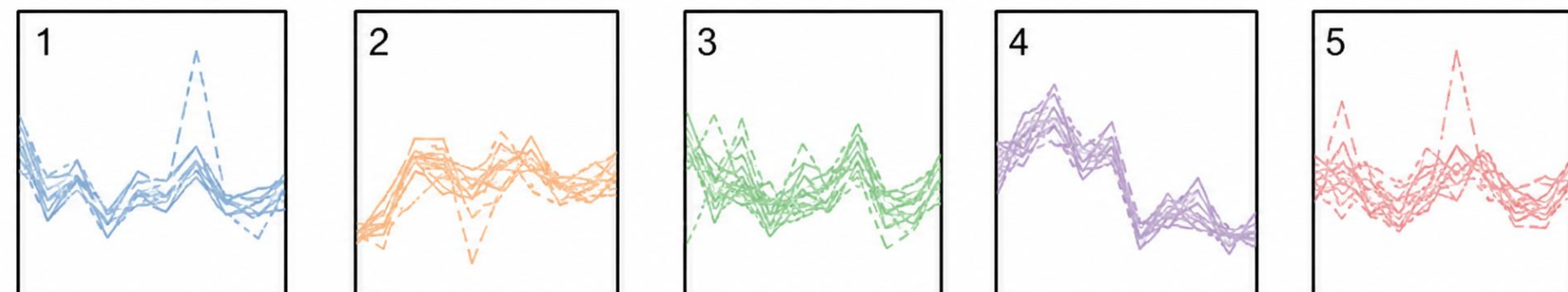
ECG morphologies



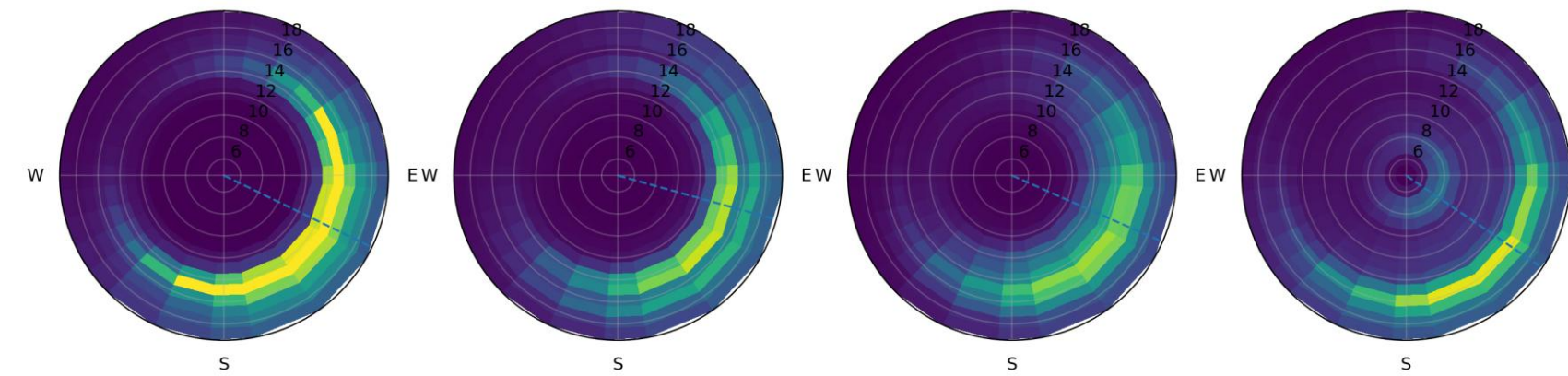
Human tasks



Gene expressions



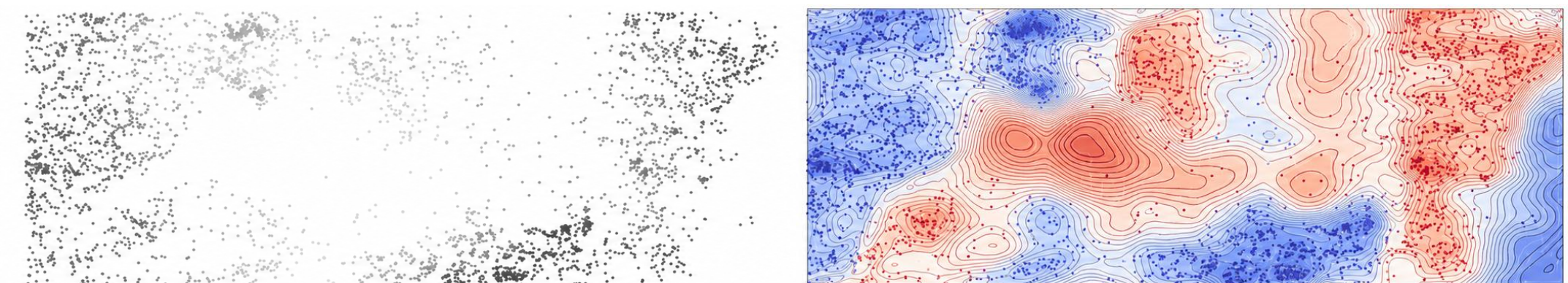
Wave spectra



Epilepsy episodes from EEG



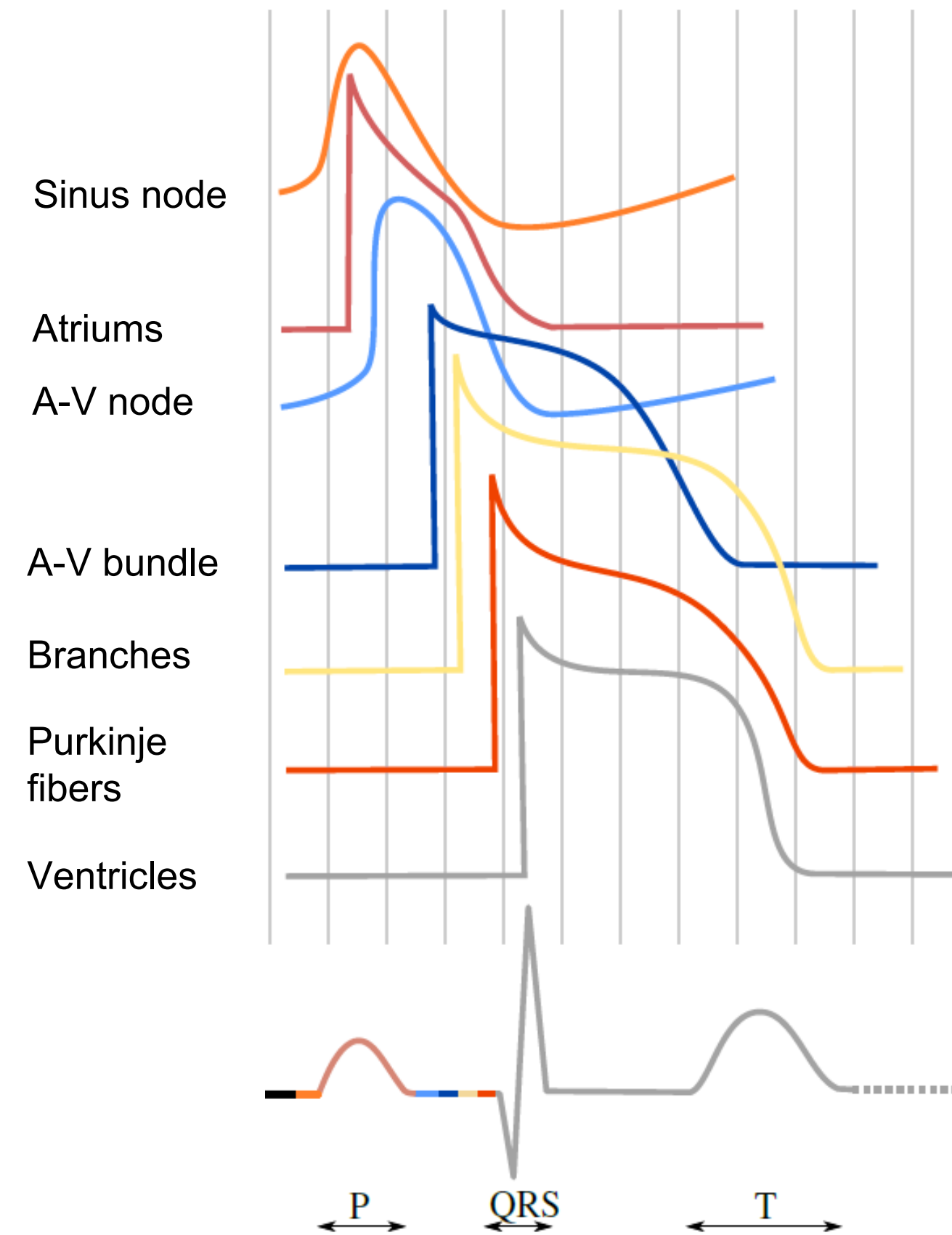
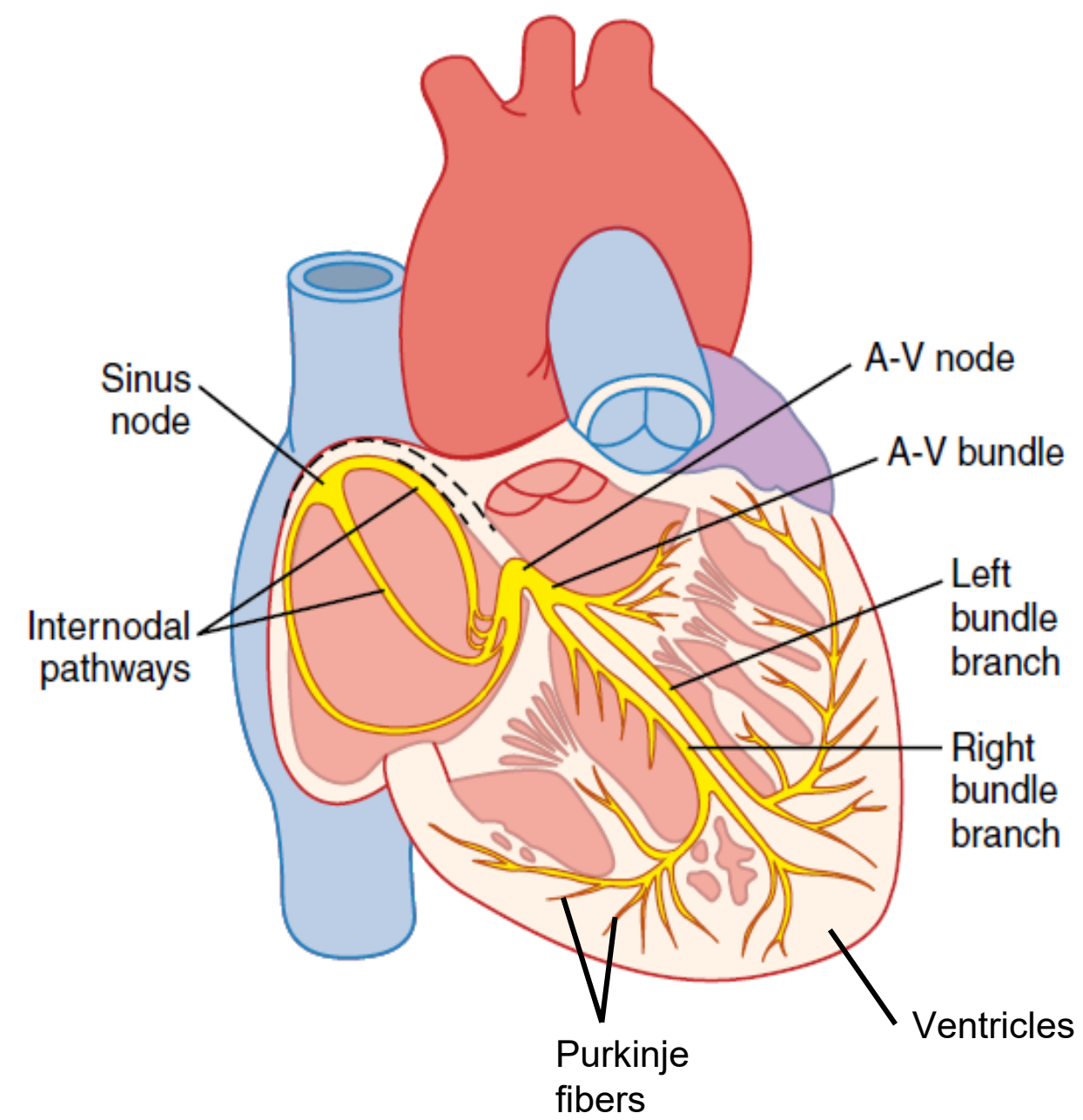
Climate patterns



Motivation

Electrocardiogram

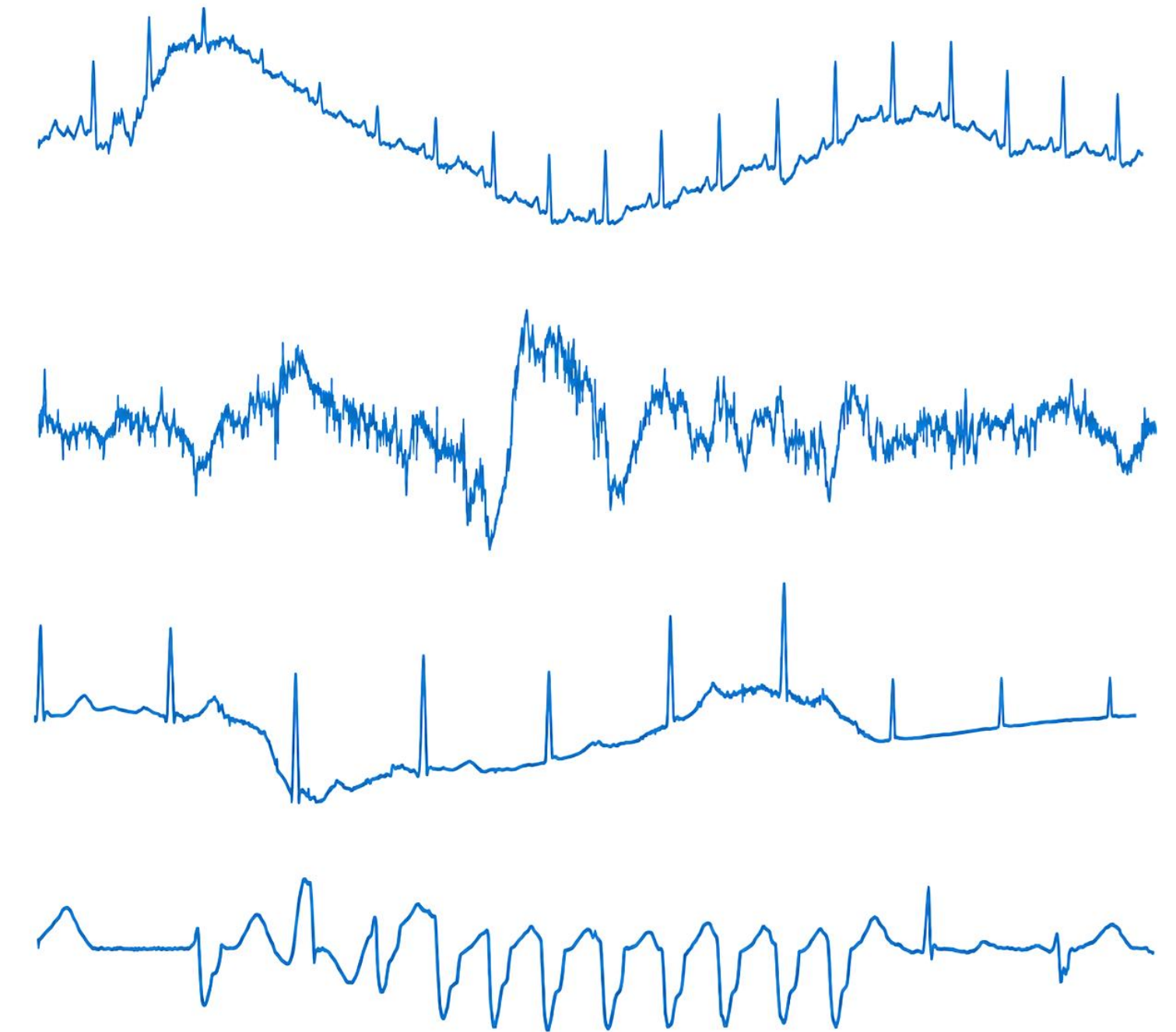
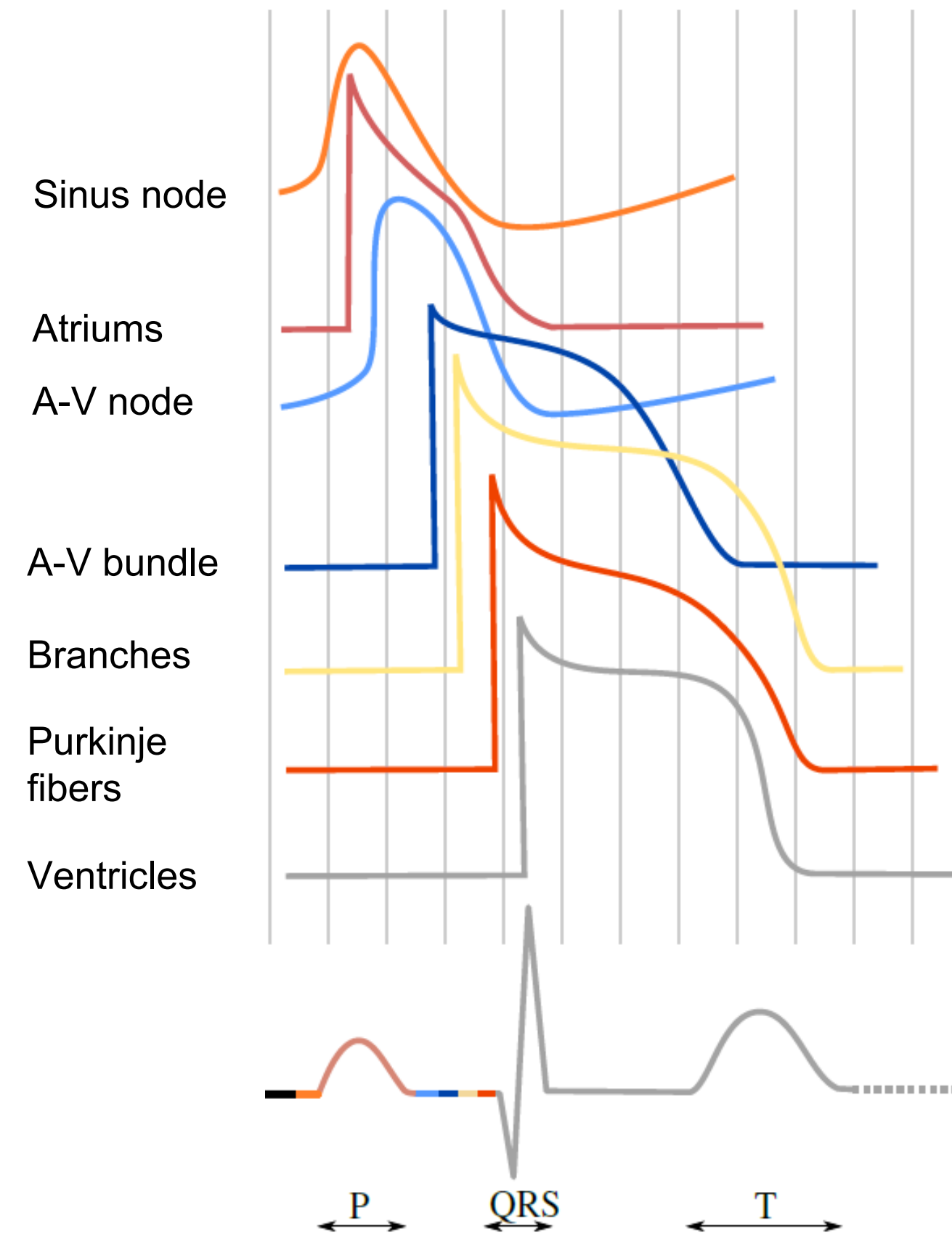
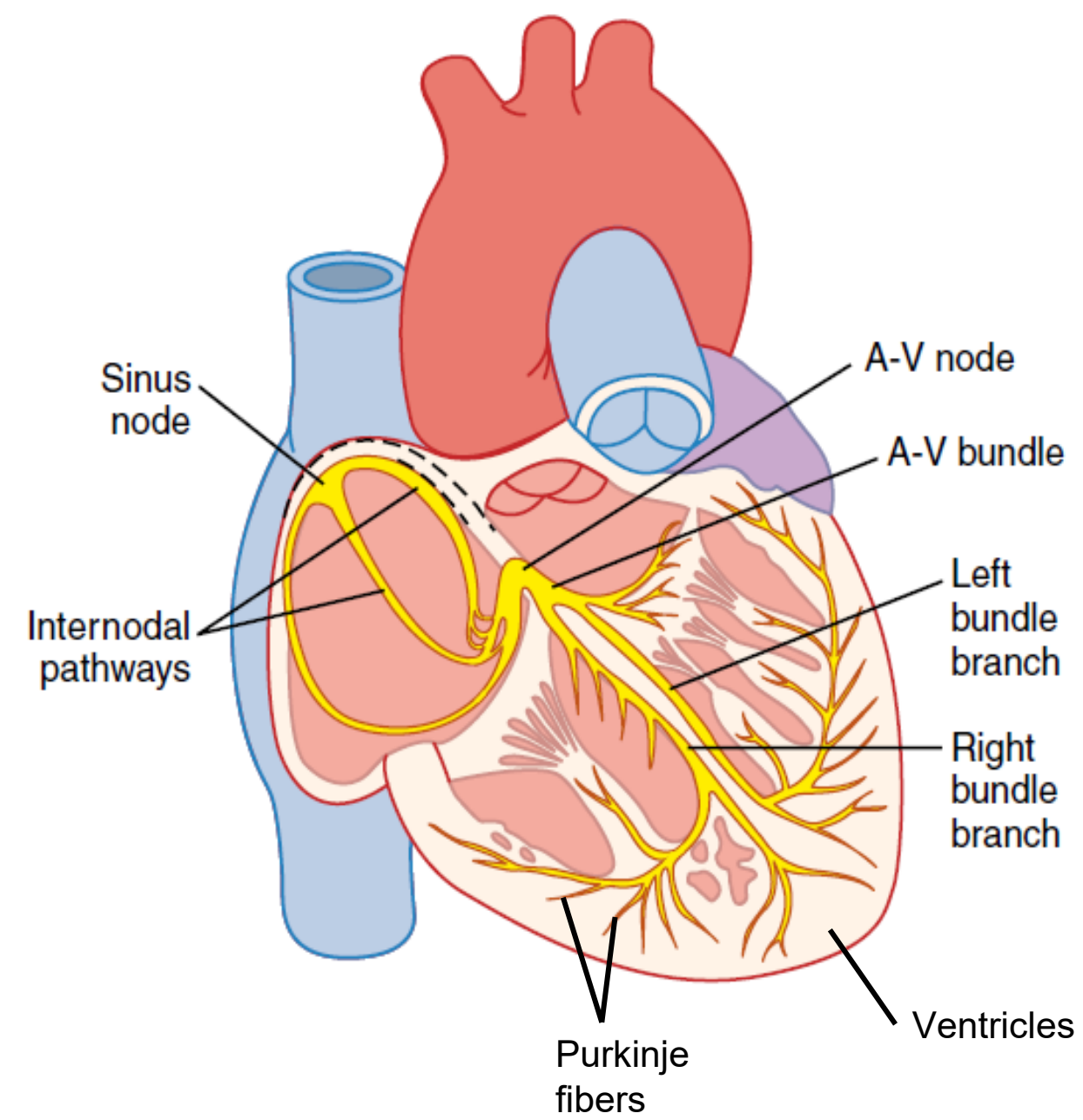
4



Motivation

Electrocardiogram

4



Records 103 and 105 of MIT-BIH Arrhythmia Database

Motivation

Evolving morphologies

5



N



N



N



N



V

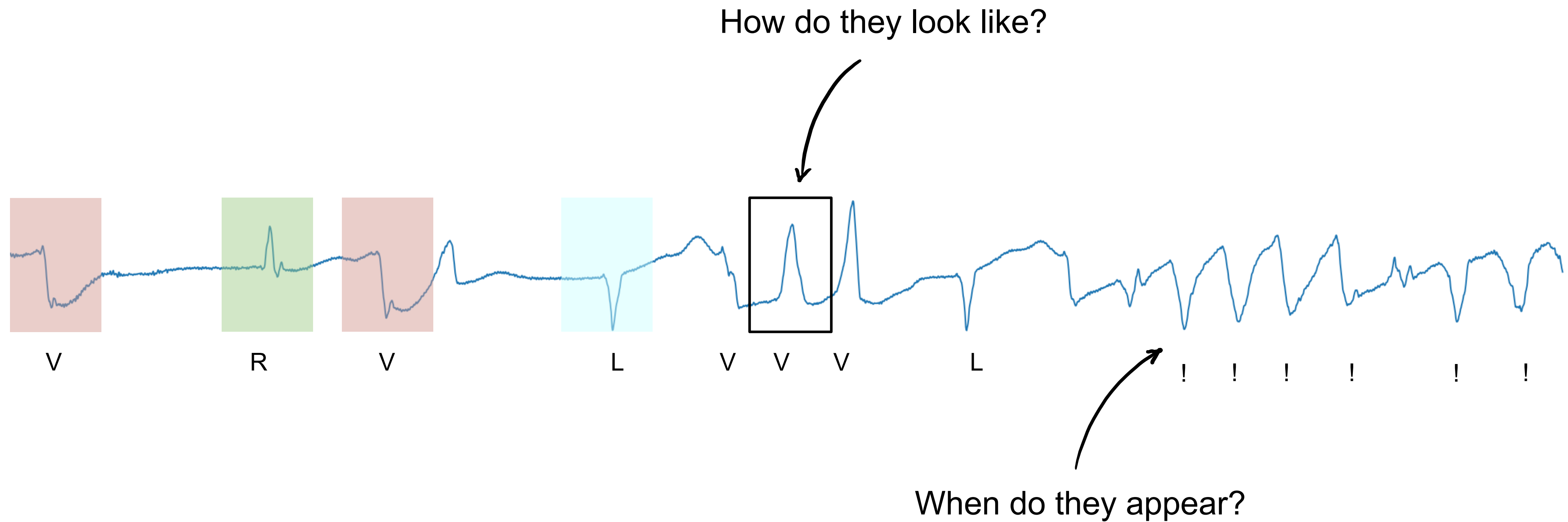


N

Motivation

Emergence of new morphologies

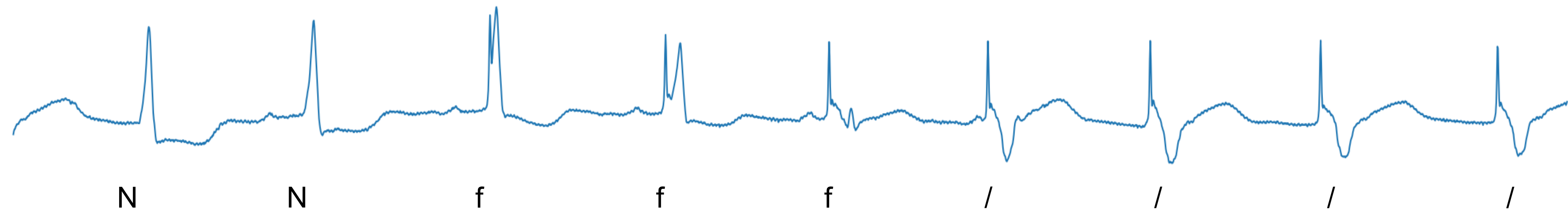
6



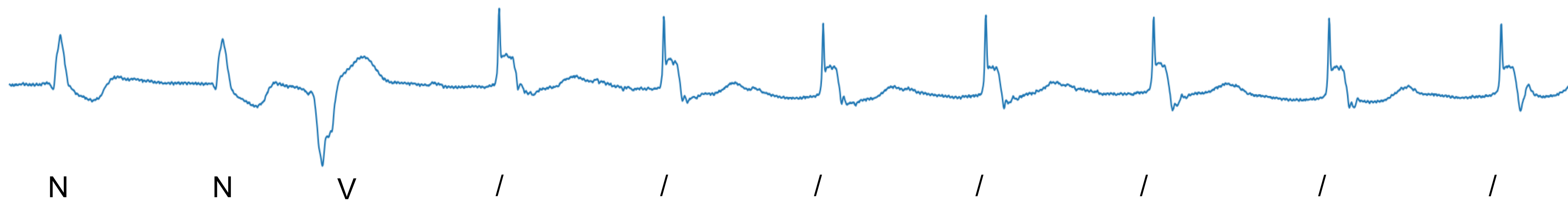
Motivation

High variability

Intra-patient



Inter-patient

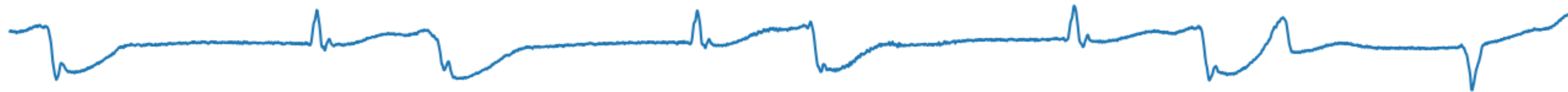


Hypothesis

Hypothesis

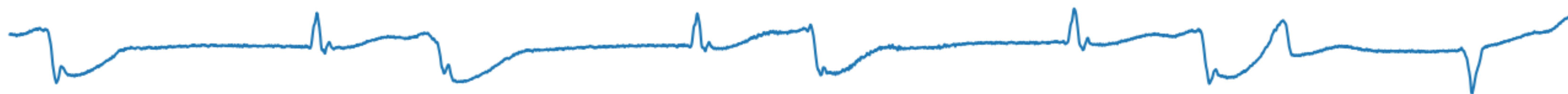
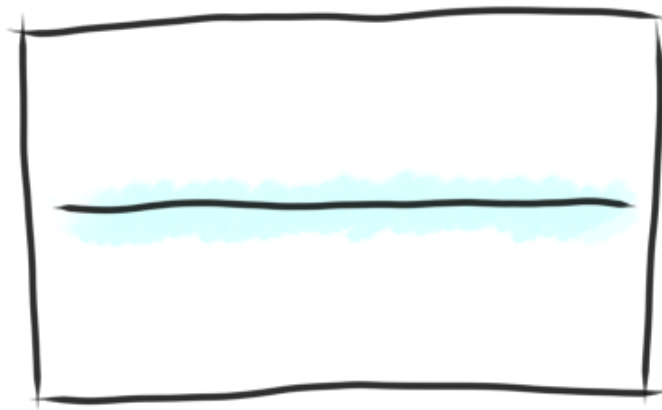
My thesis in a napkin

8



Hypothesis

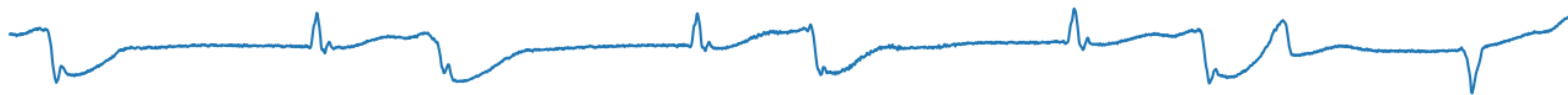
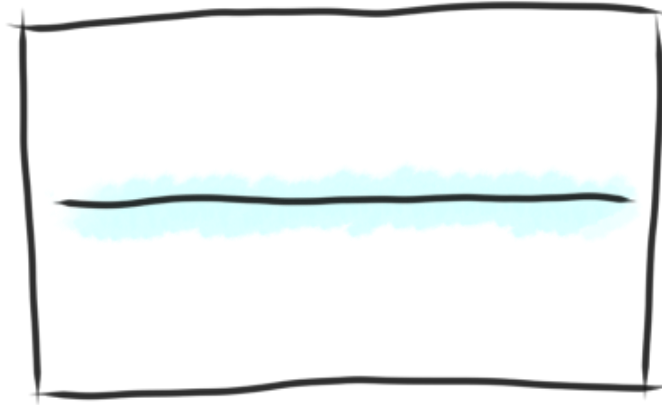
My thesis in a napkin



Hypothesis

My thesis in a napkin

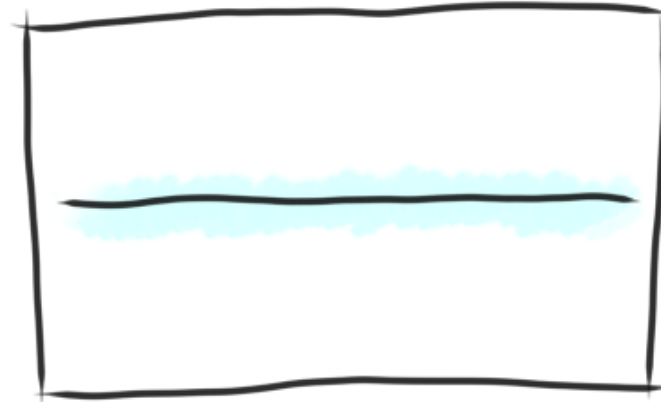
probability
stick



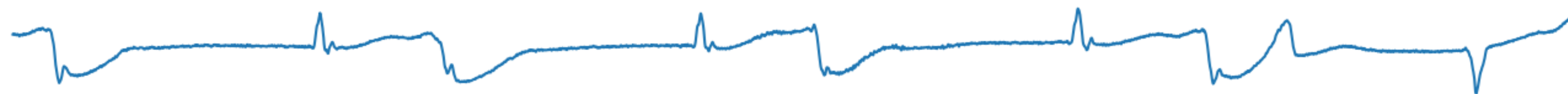
Hypothesis

My thesis in a napkin

probability
stick

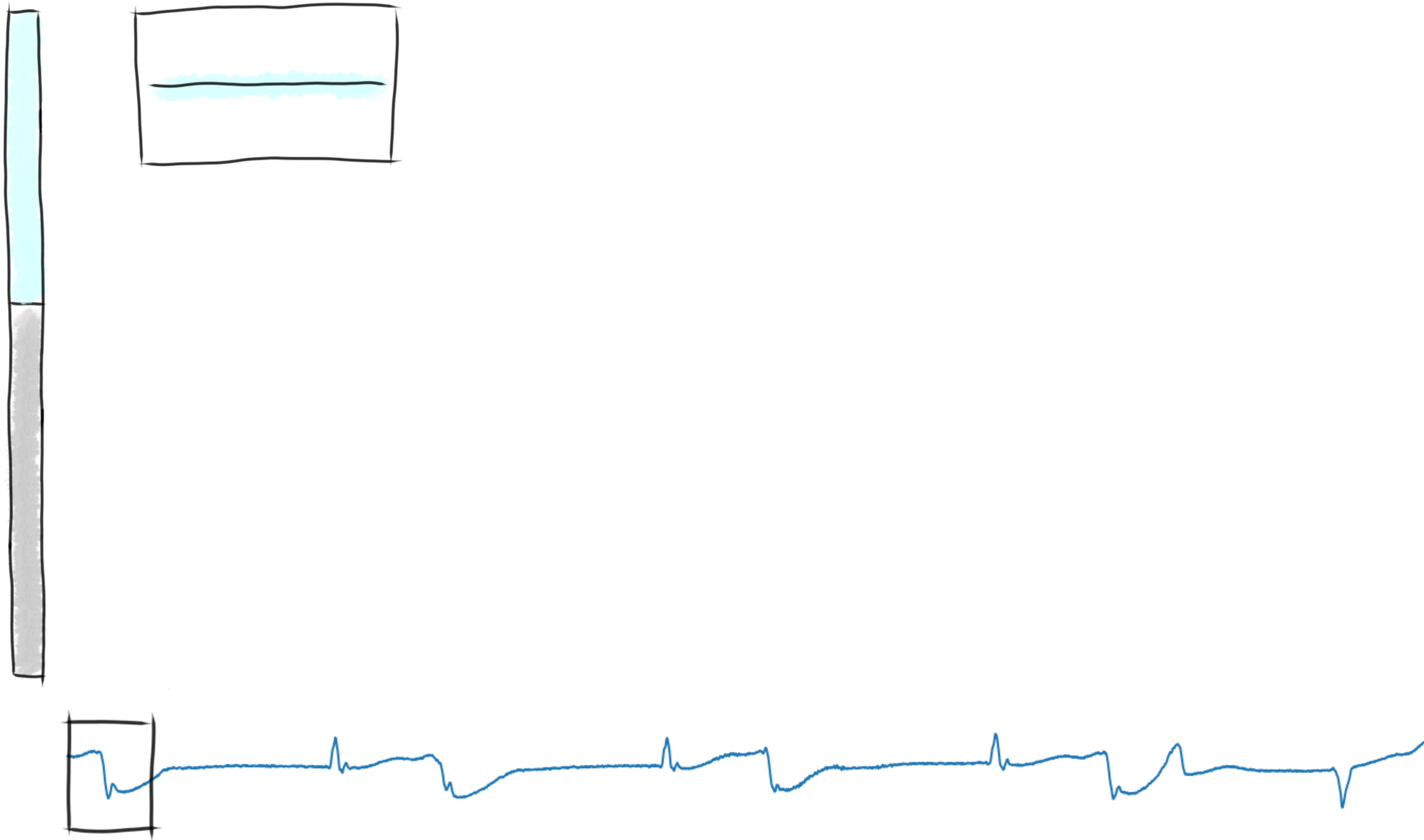


unused
probability



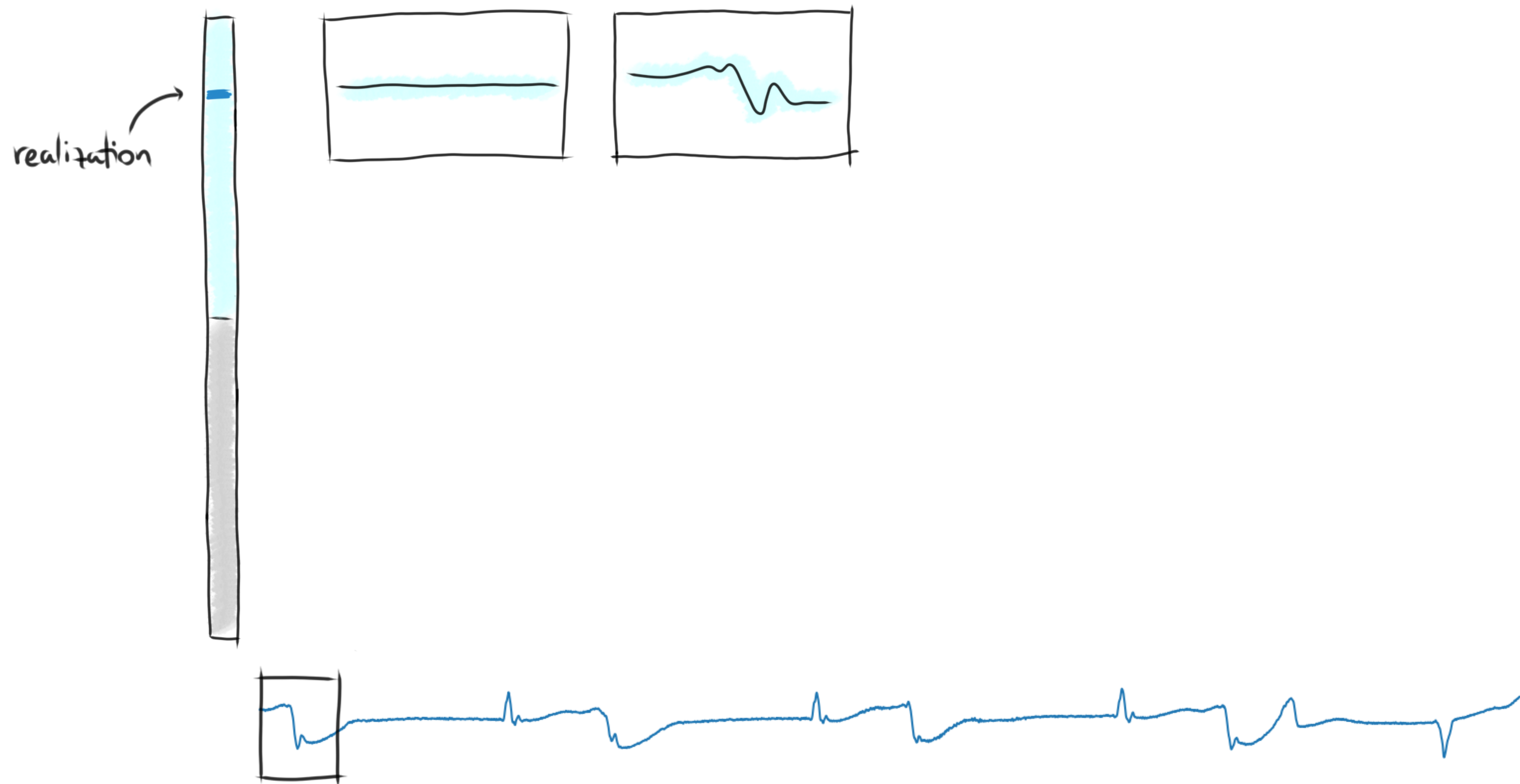
Hypothesis

My thesis in a napkin



Hypothesis

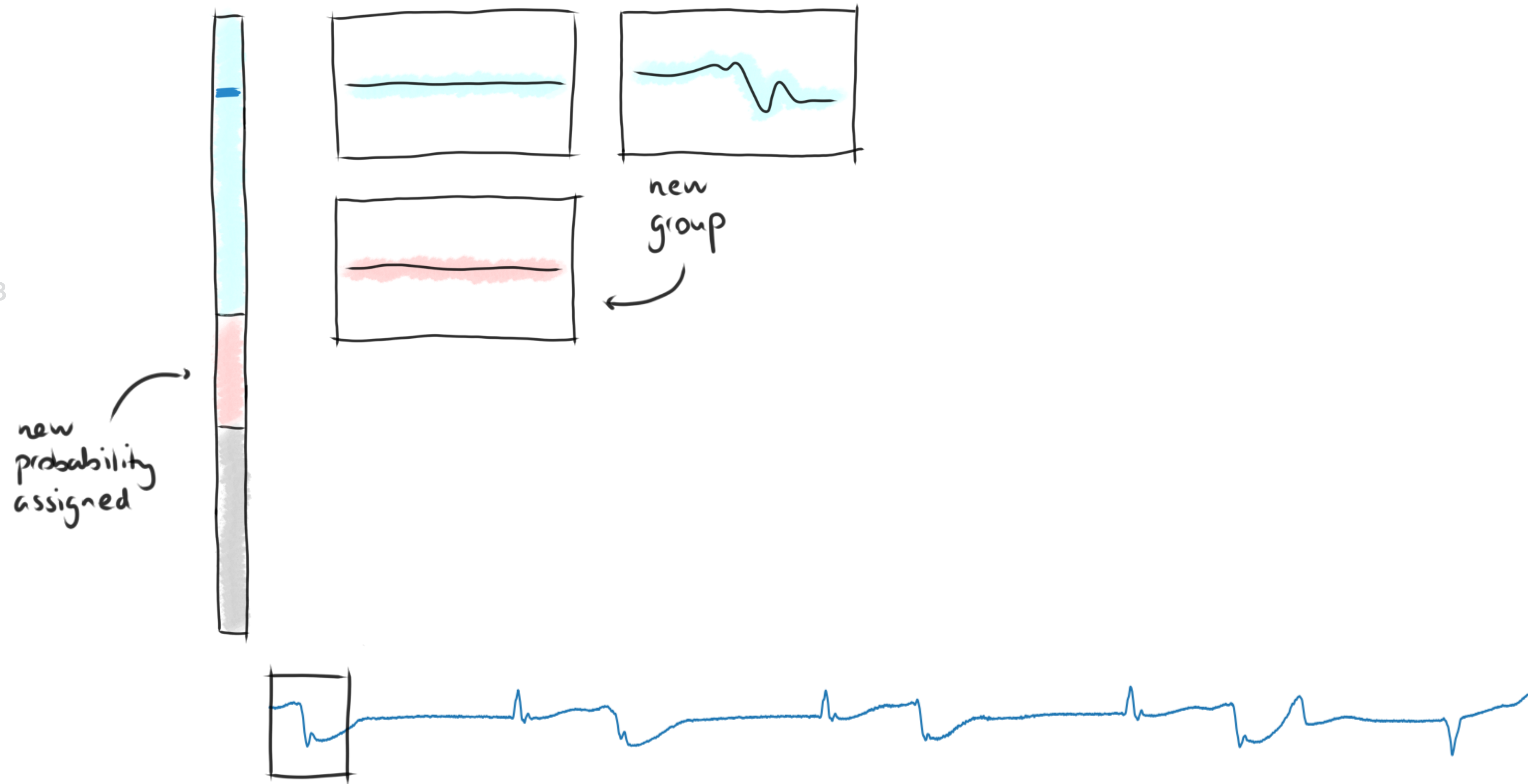
My thesis in a napkin



Hypothesis

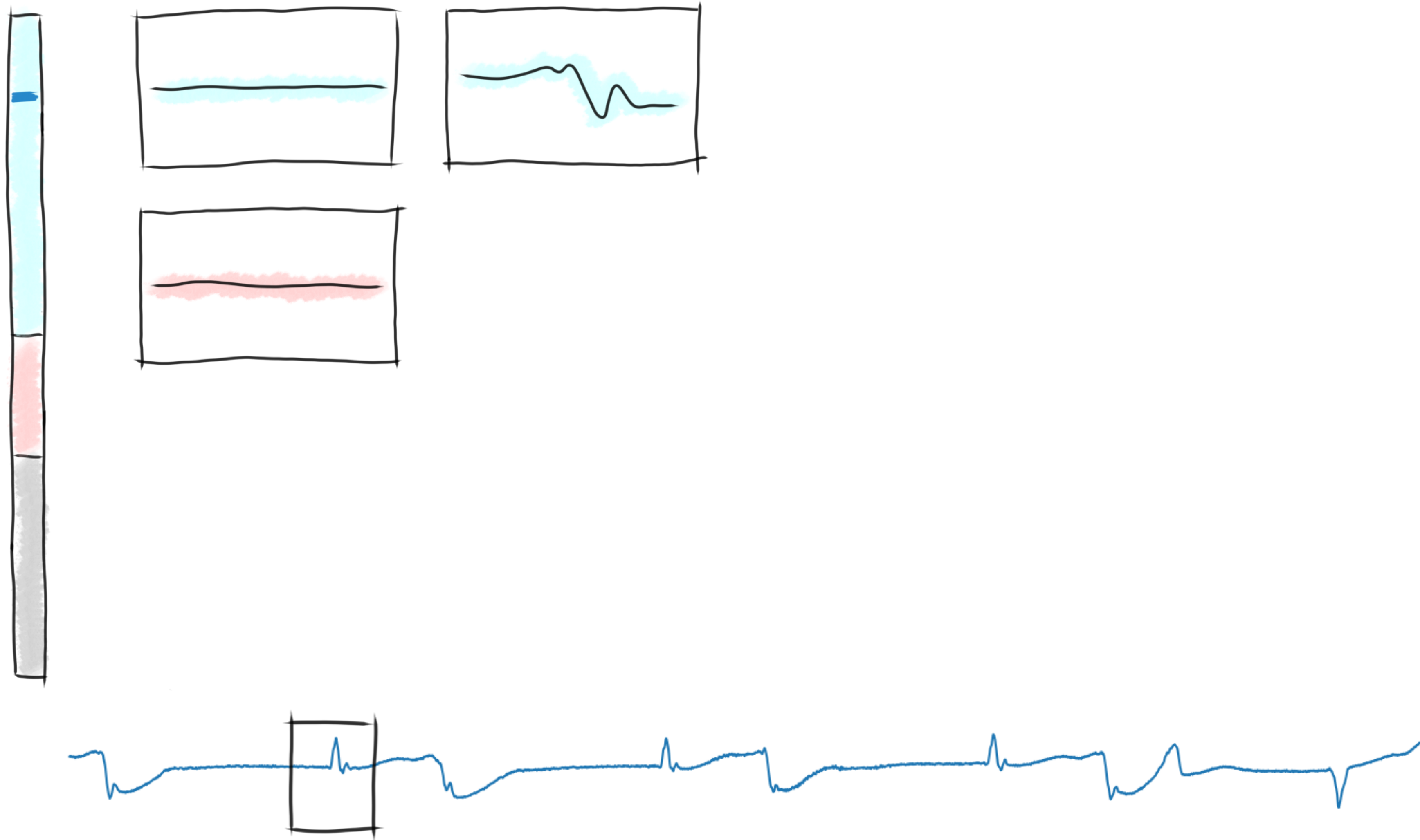
My thesis in a napkin

8



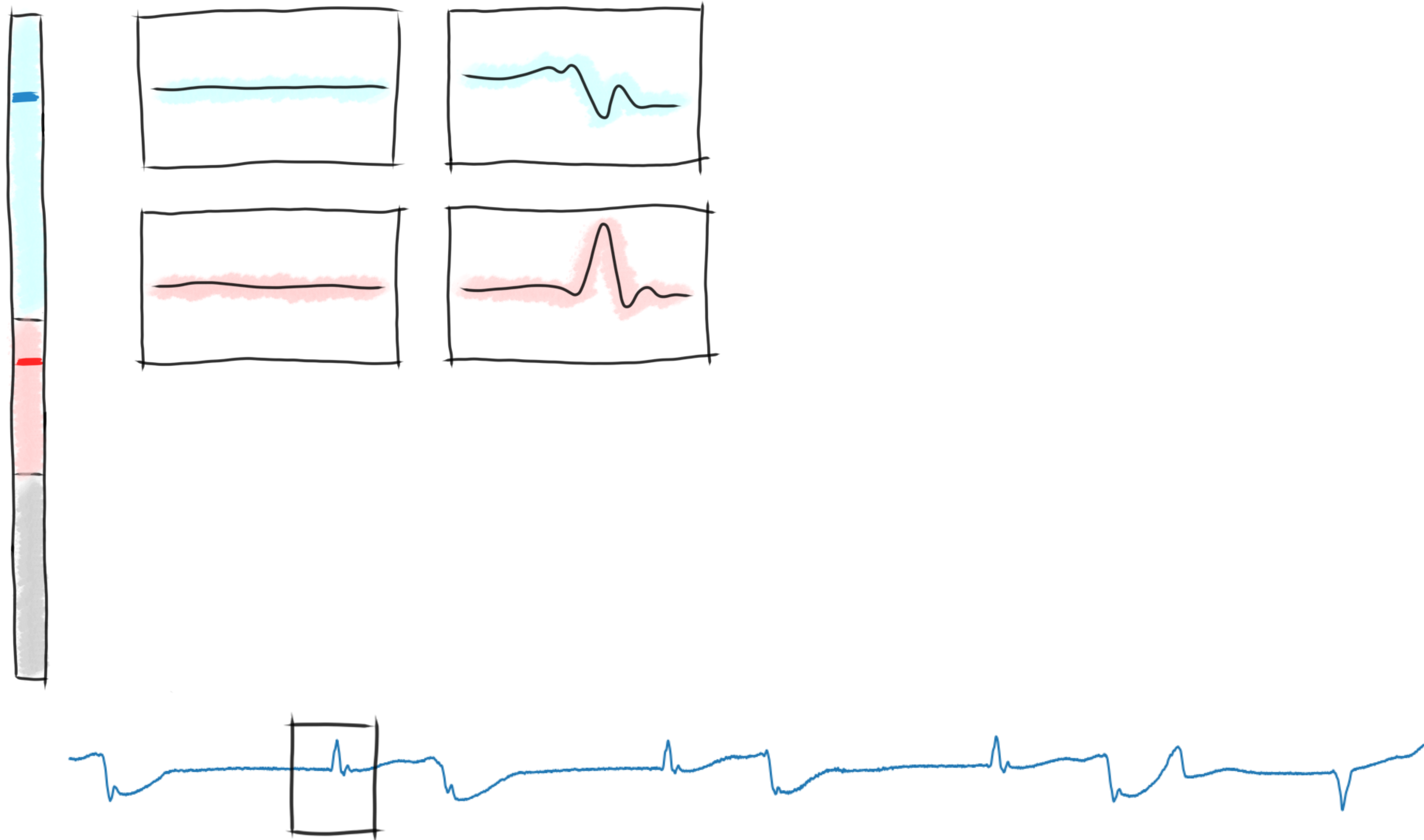
Hypothesis

My thesis in a napkin



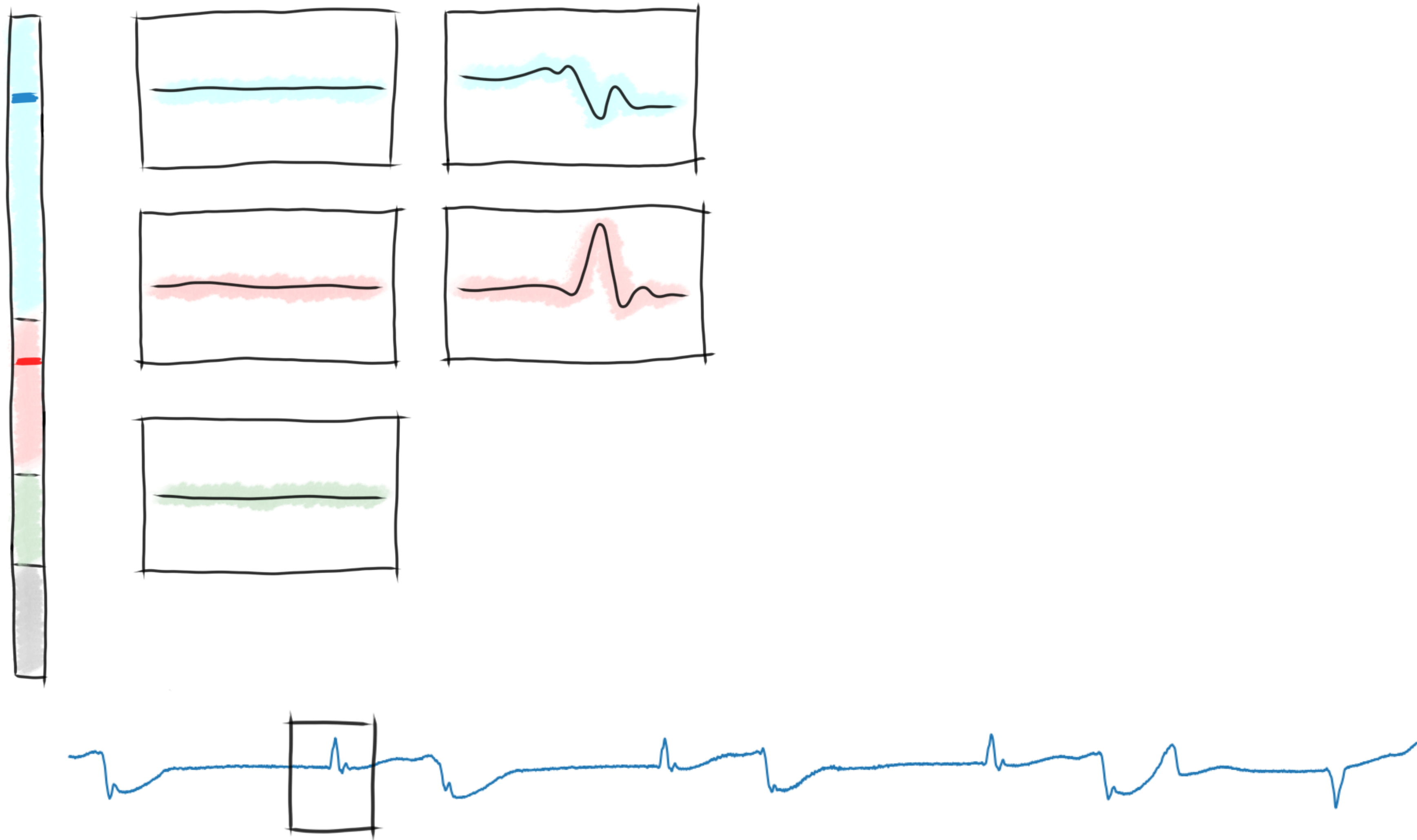
Hypothesis

My thesis in a napkin



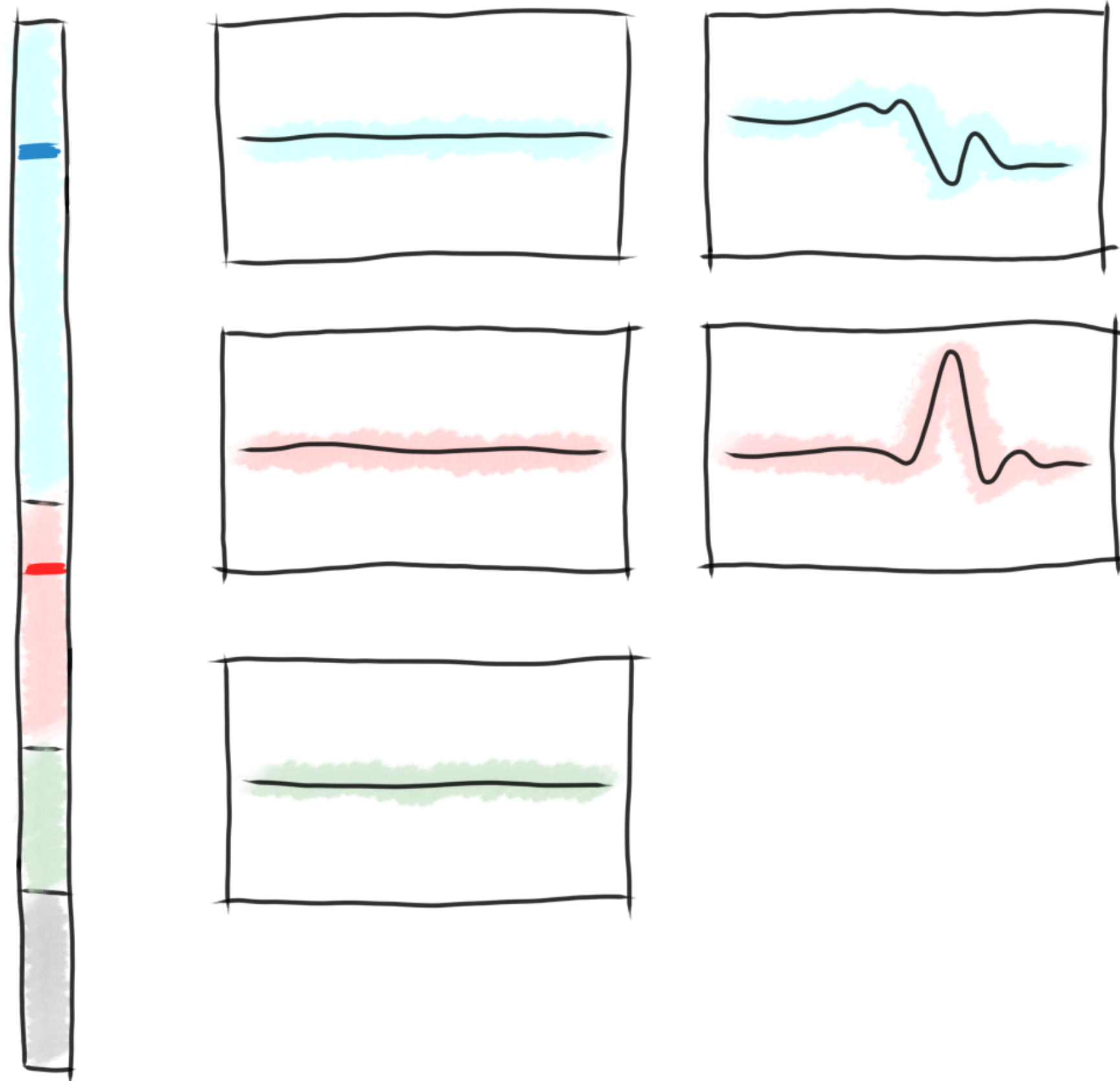
Hypothesis

My thesis in a napkin



Hypothesis

My thesis in a napkin

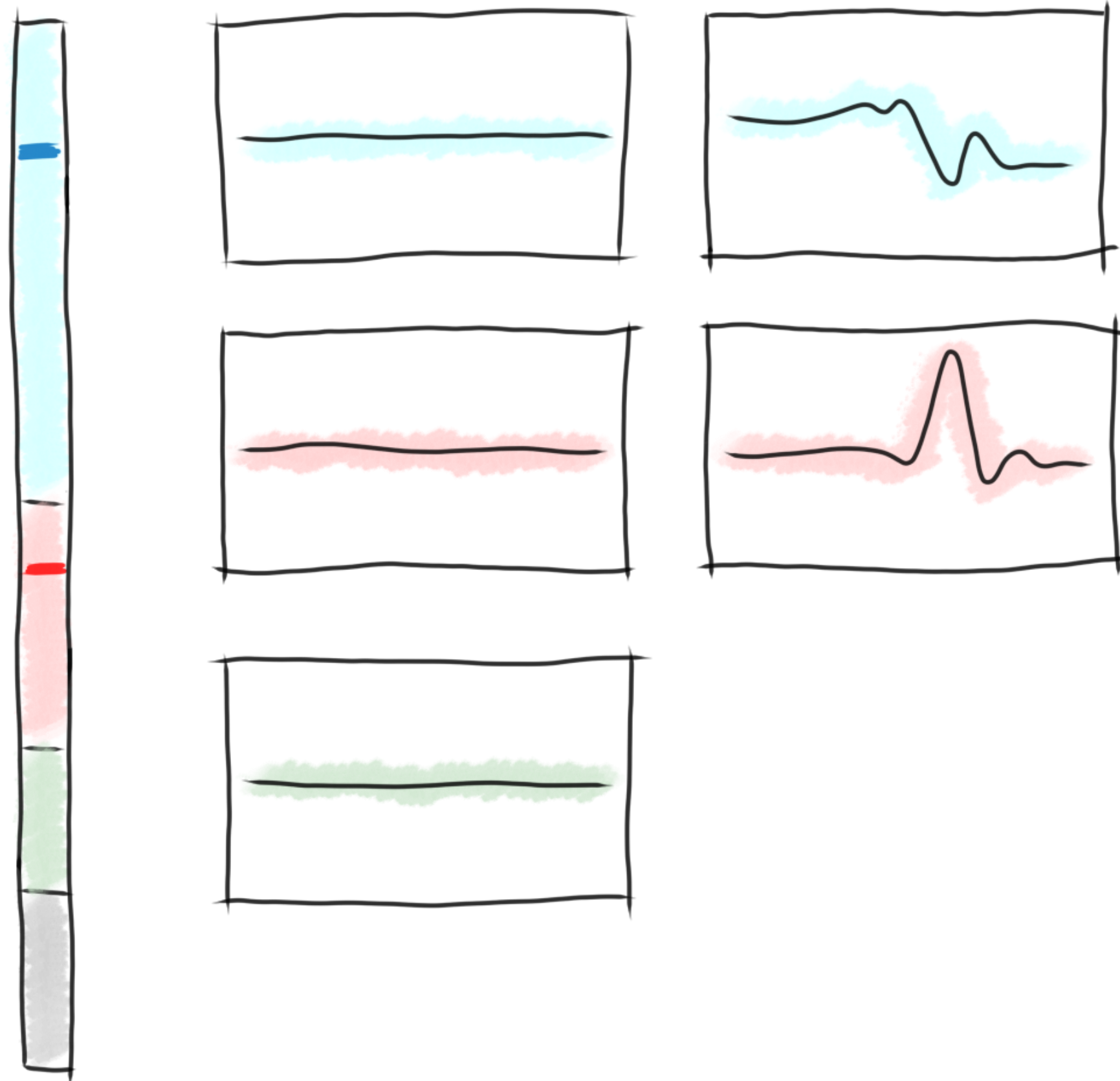


QUERY 1

GIVEN THE LAST ASSIGNMENT, WHAT SHOULD A FUTURE INSTANCE LOOK LIKE?

Hypothesis

My thesis in a napkin



QUERY 1

GIVEN THE LAST ASSIGNMENT, WHAT SHOULD A FUTURE INSTANCE LOOK LIKE?

ANSWER

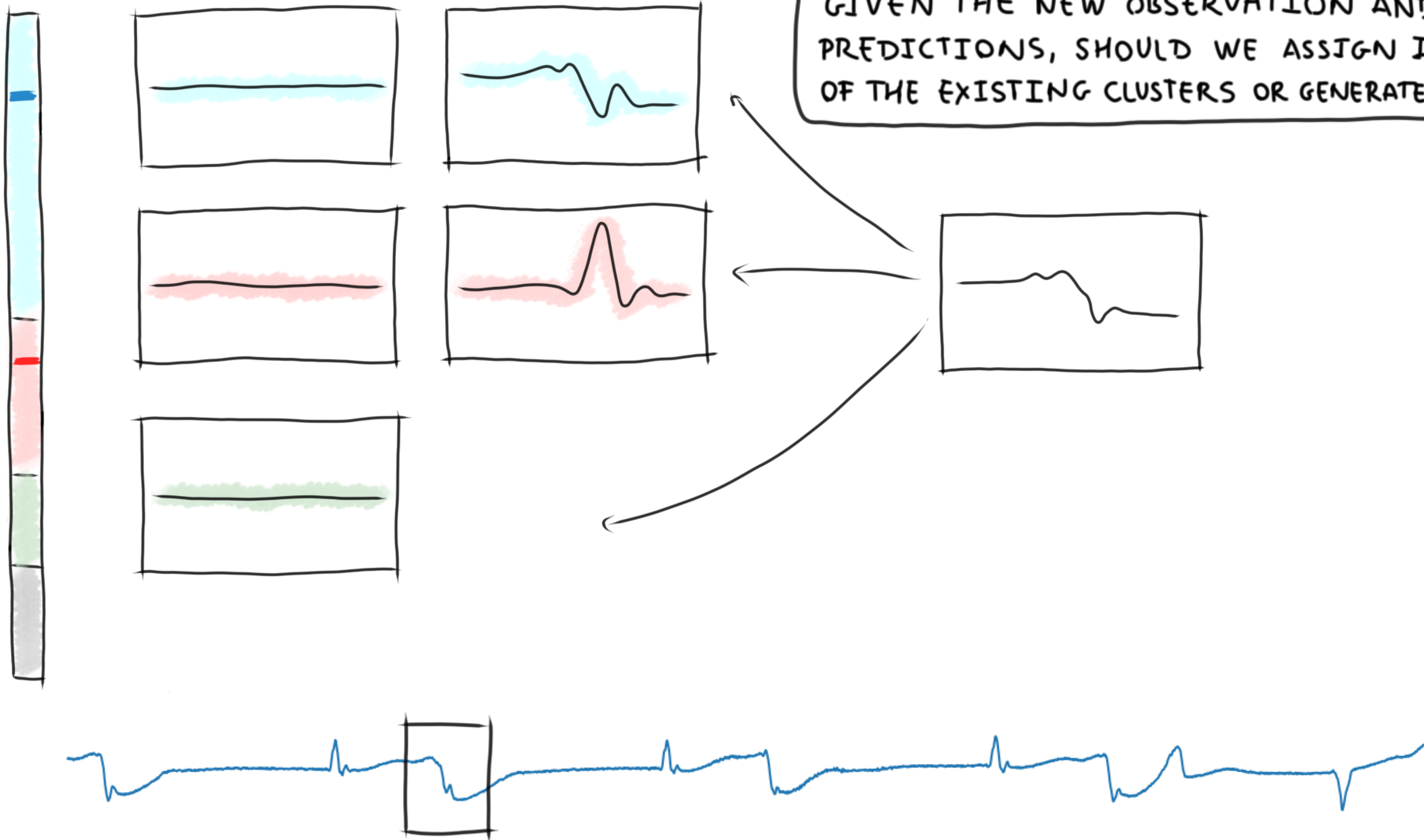
WE NEED THE ABILITY TO GENERATE PREDICTIONS AND MEASURE UNCERTAINTY TO KNOW WHAT IS EXPECTED AND NOVELTY

Hypothesis

My thesis in a napkin

QUERY 2

GIVEN THE NEW OBSERVATION AND OUR PREDICTIONS, SHOULD WE ASSIGN IT TO ONE OF THE EXISTING CLUSTERS OR GENERATE A NEW ONE?



Hypothesis

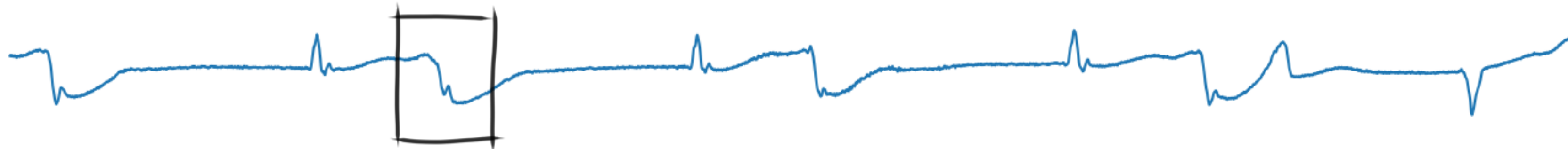
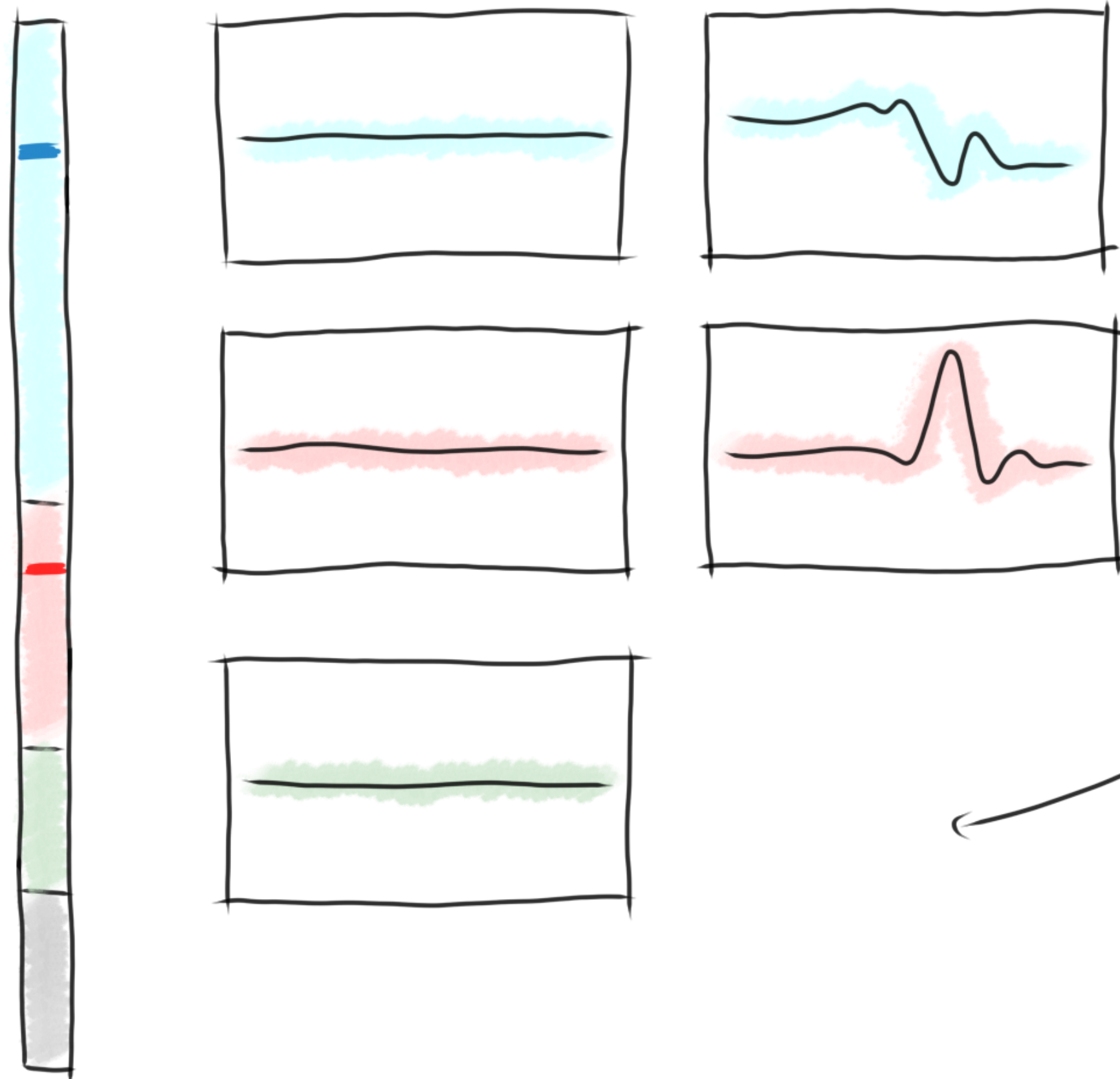
My thesis in a napkin

QUERY 2

GIVEN THE NEW OBSERVATION AND OUR PREDICTIONS, SHOULD WE ASSIGN IT TO ONE OF THE EXISTING CLUSTERS OR GENERATE A NEW ONE?

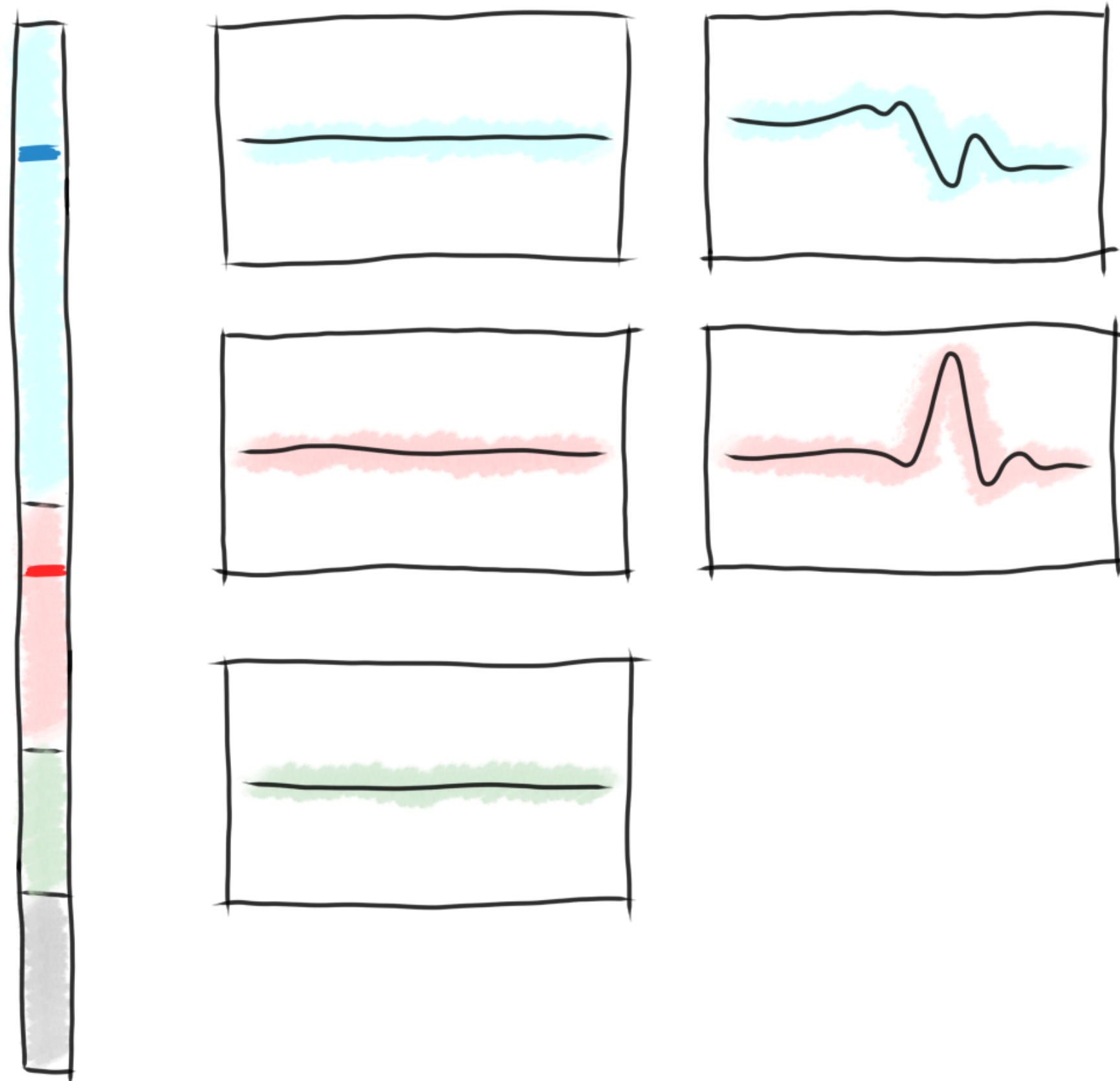
ANSWER

IF SO FLEXIBLE, WE COULD MODEL EVERY EXAMPLE TOGETHER; BUT IF SO RIGID THEN WE COULD GENERATE ONE CLUSTER PER SAMPLE



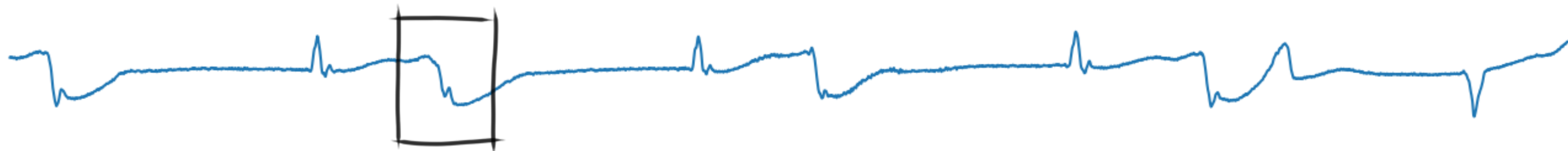
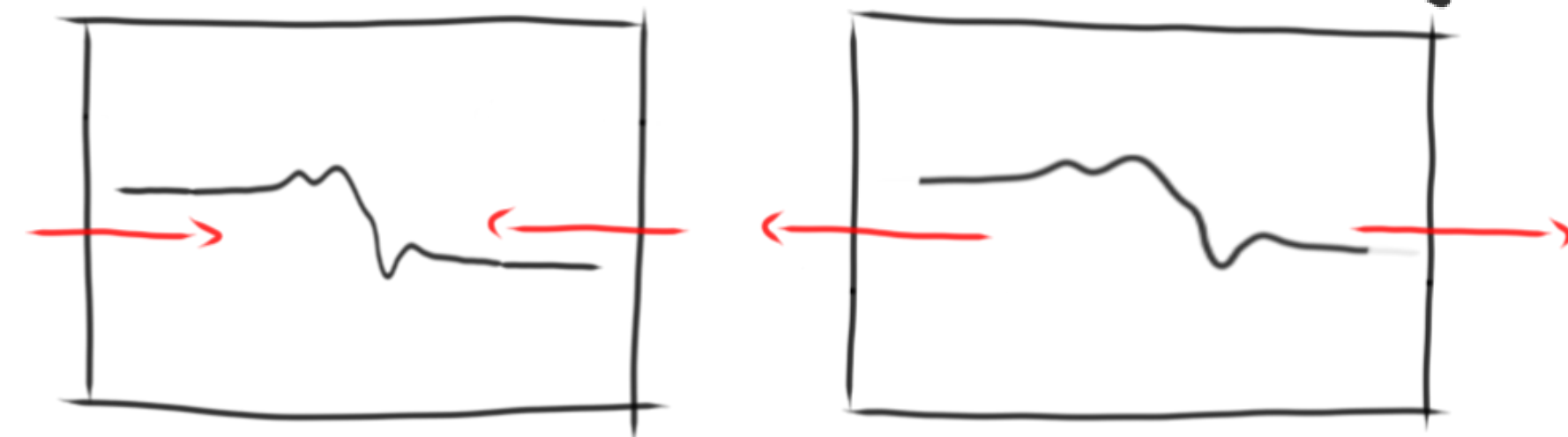
Hypothesis

My thesis in a napkin



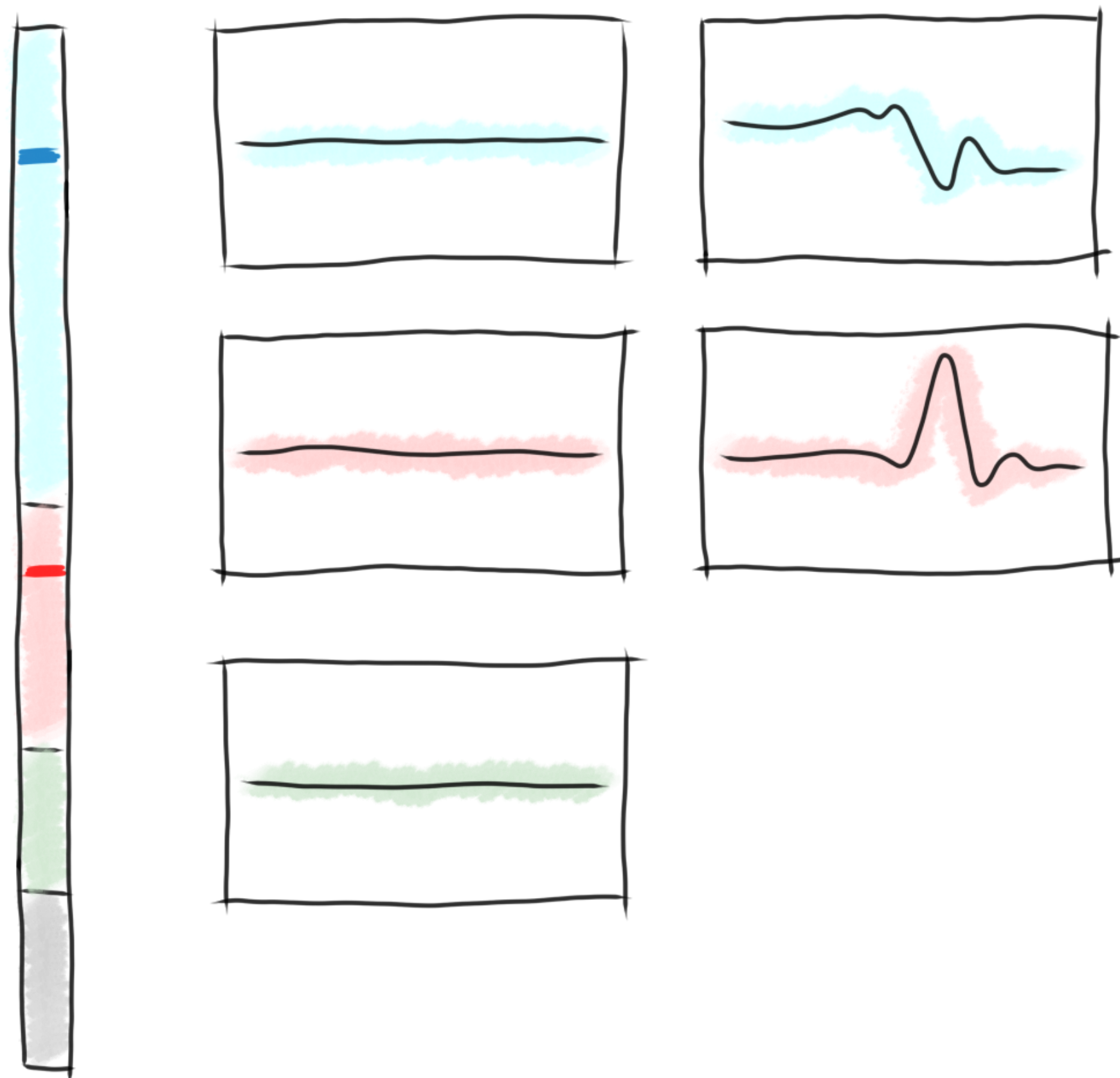
QUERY 3

CAN SOME DIFFERENCES ARISE
FROM AN ALIGNMENT DISTORTION?



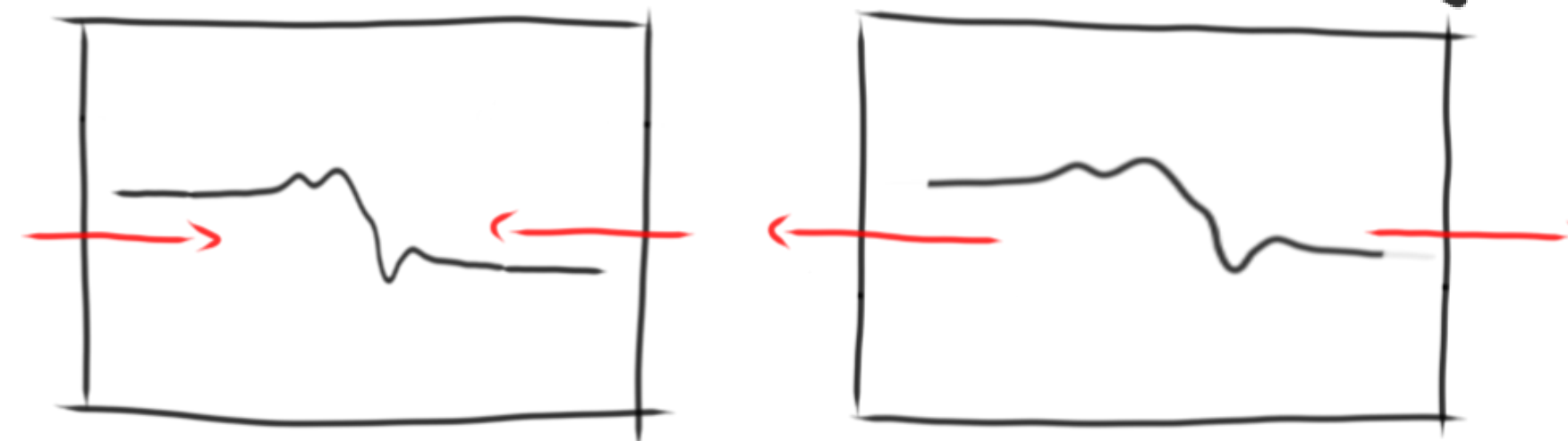
Hypothesis

My thesis in a napkin



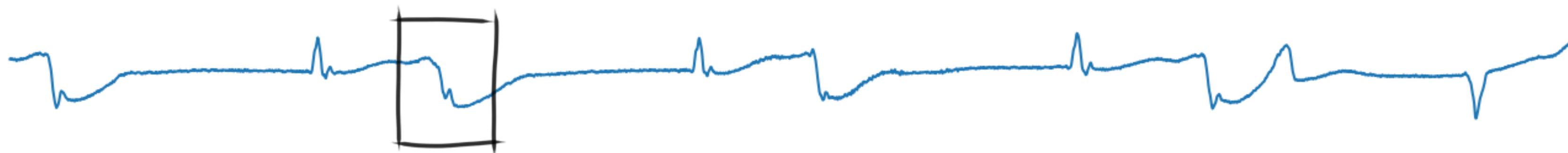
QUERY 3

CAN SOME DIFFERENCES ARISE FROM AN ALIGNMENT DISTORTION?



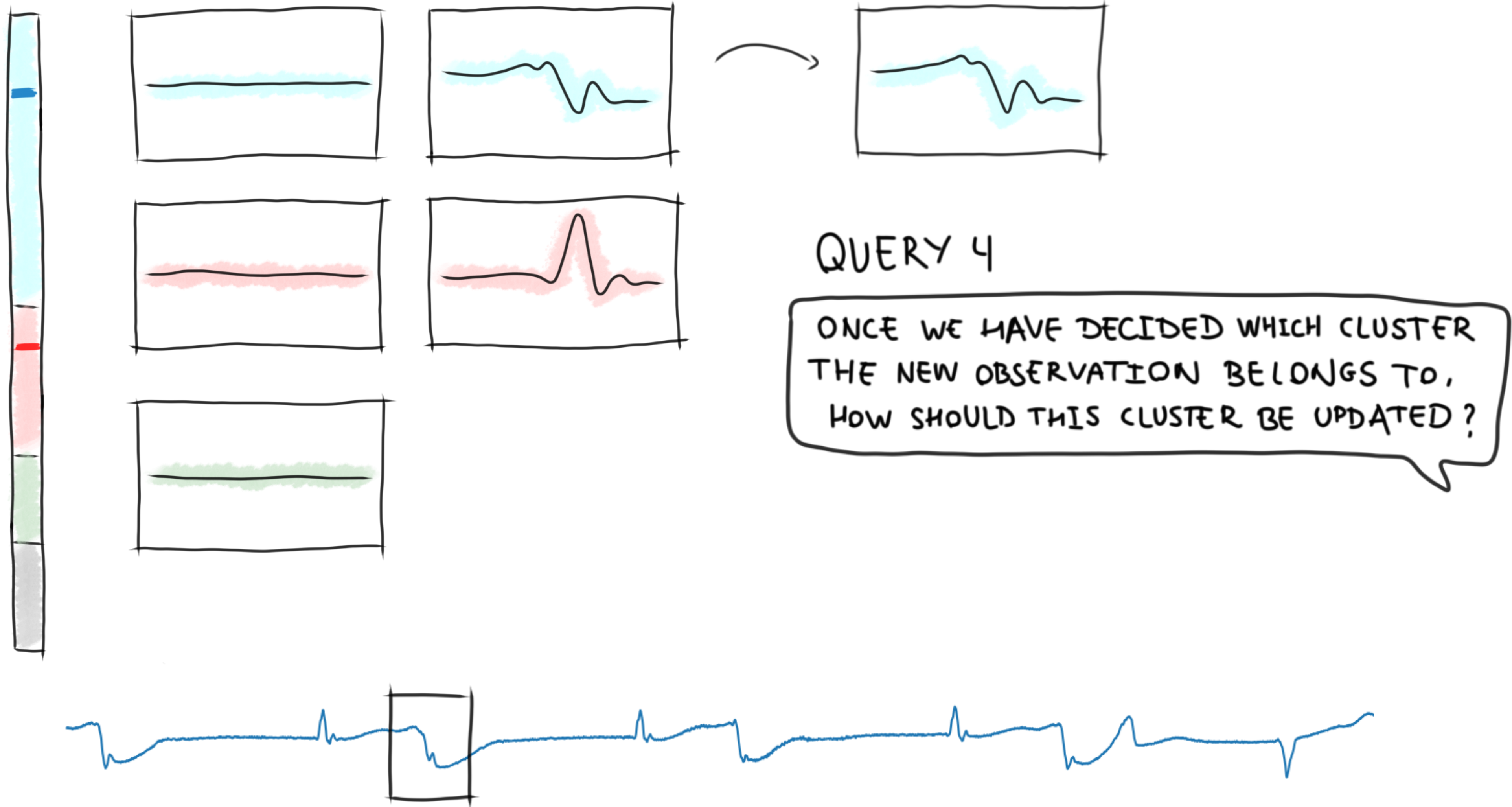
COMMENT

THE LATENT SAMPLES CAN SUFFER ELONGATION, CONTRACTION OR X-AXIS SHIFT BEFORE OBSERVATION



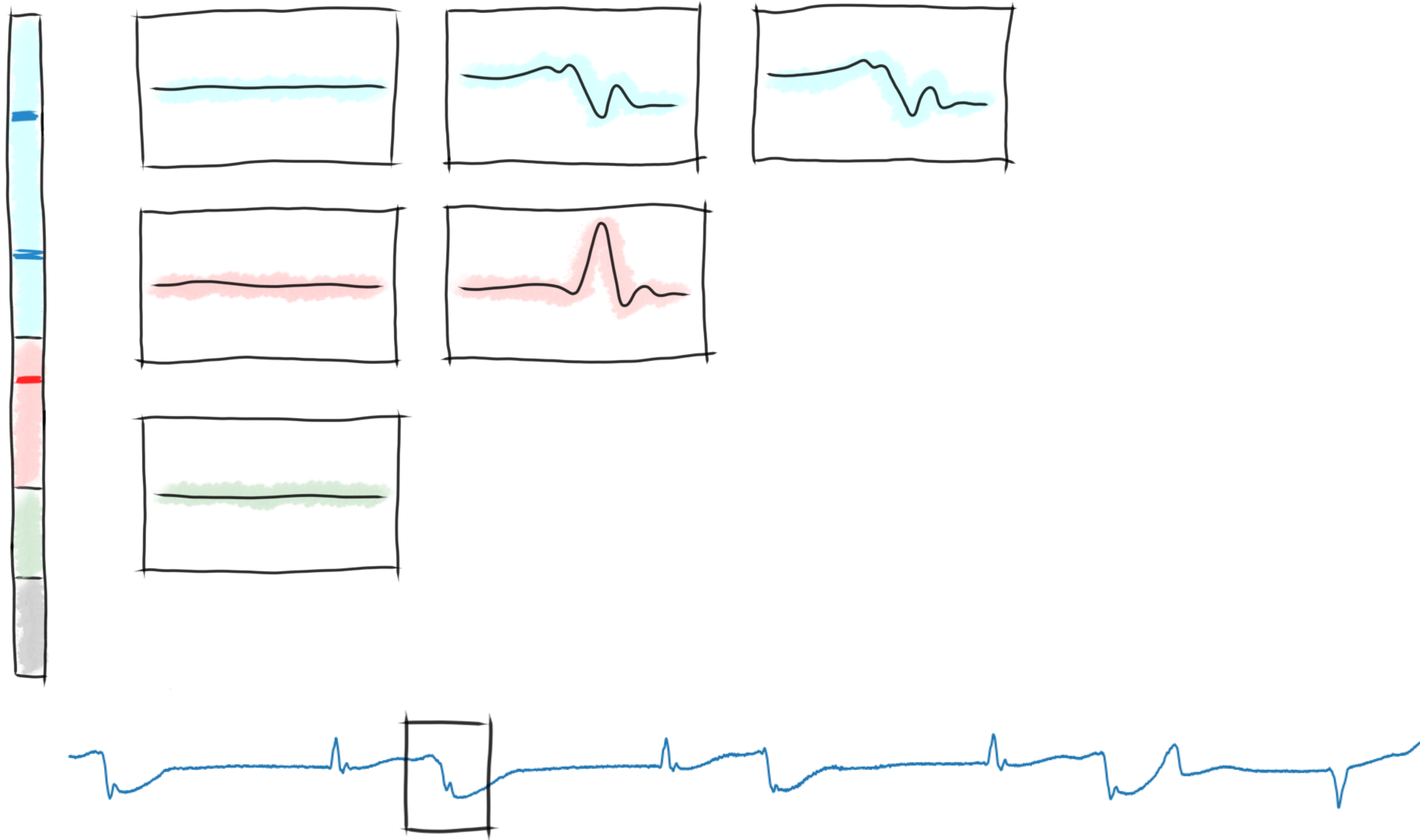
Hypothesis

My thesis in a napkin



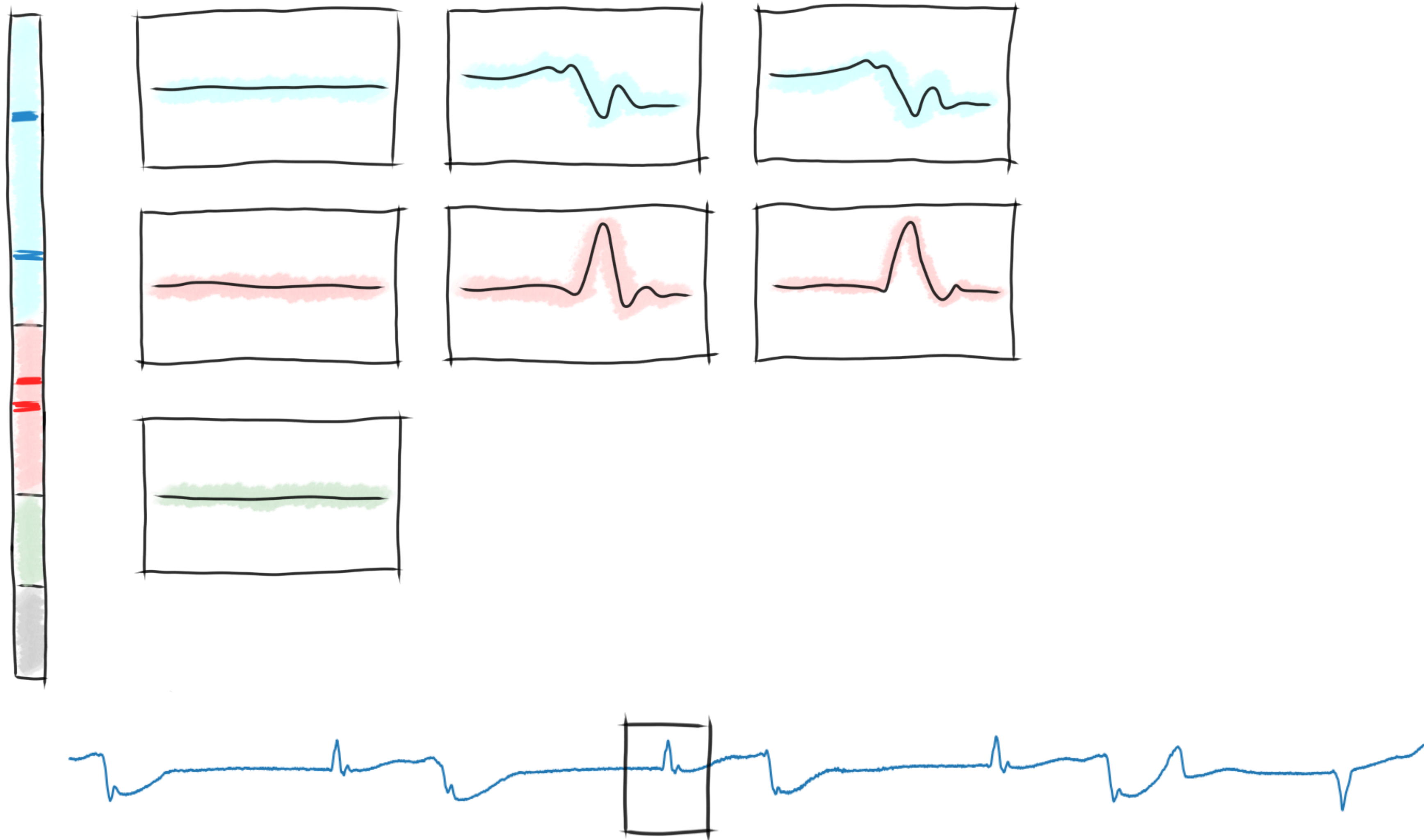
Hypothesis

My thesis in a napkin



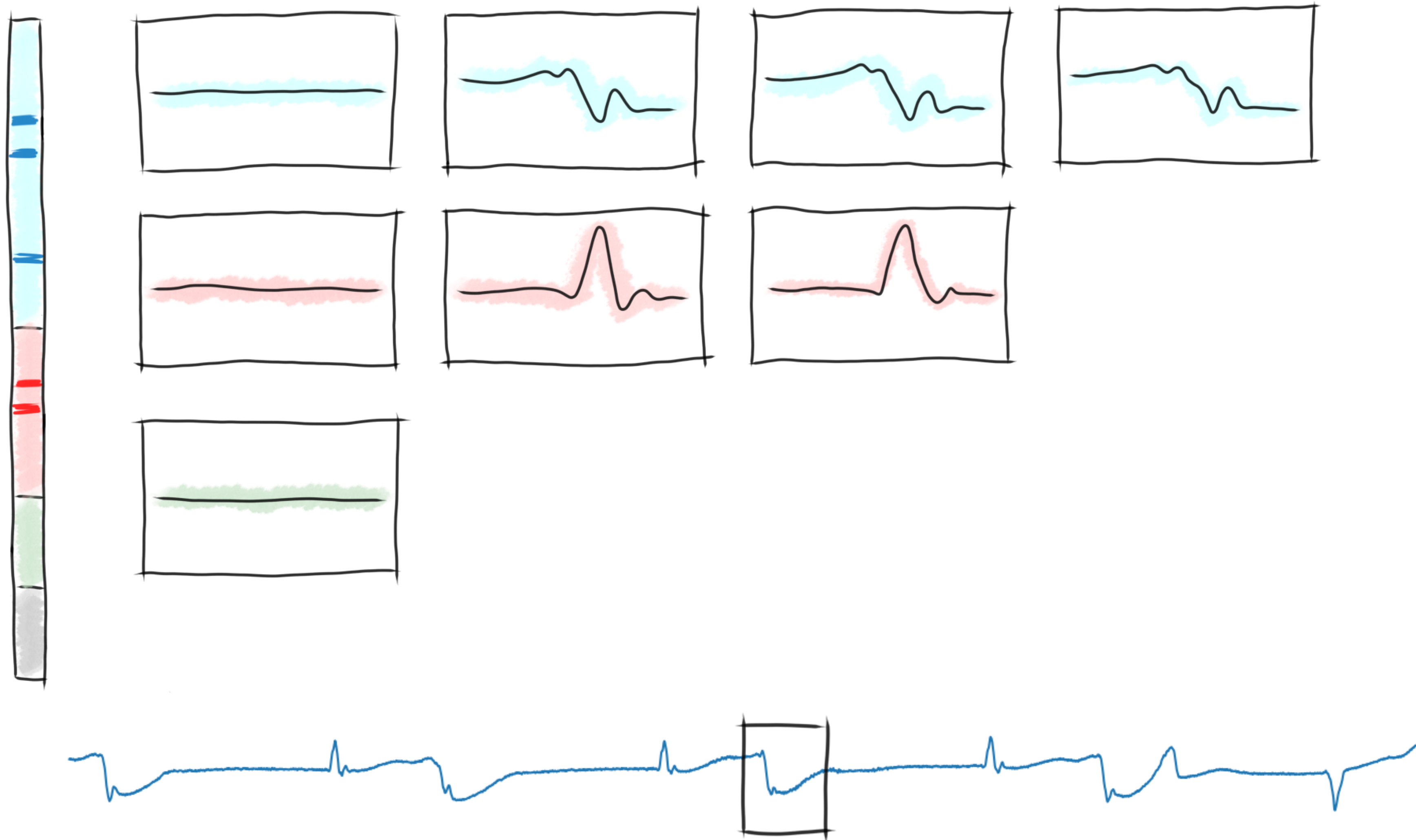
Hypothesis

My thesis in a napkin



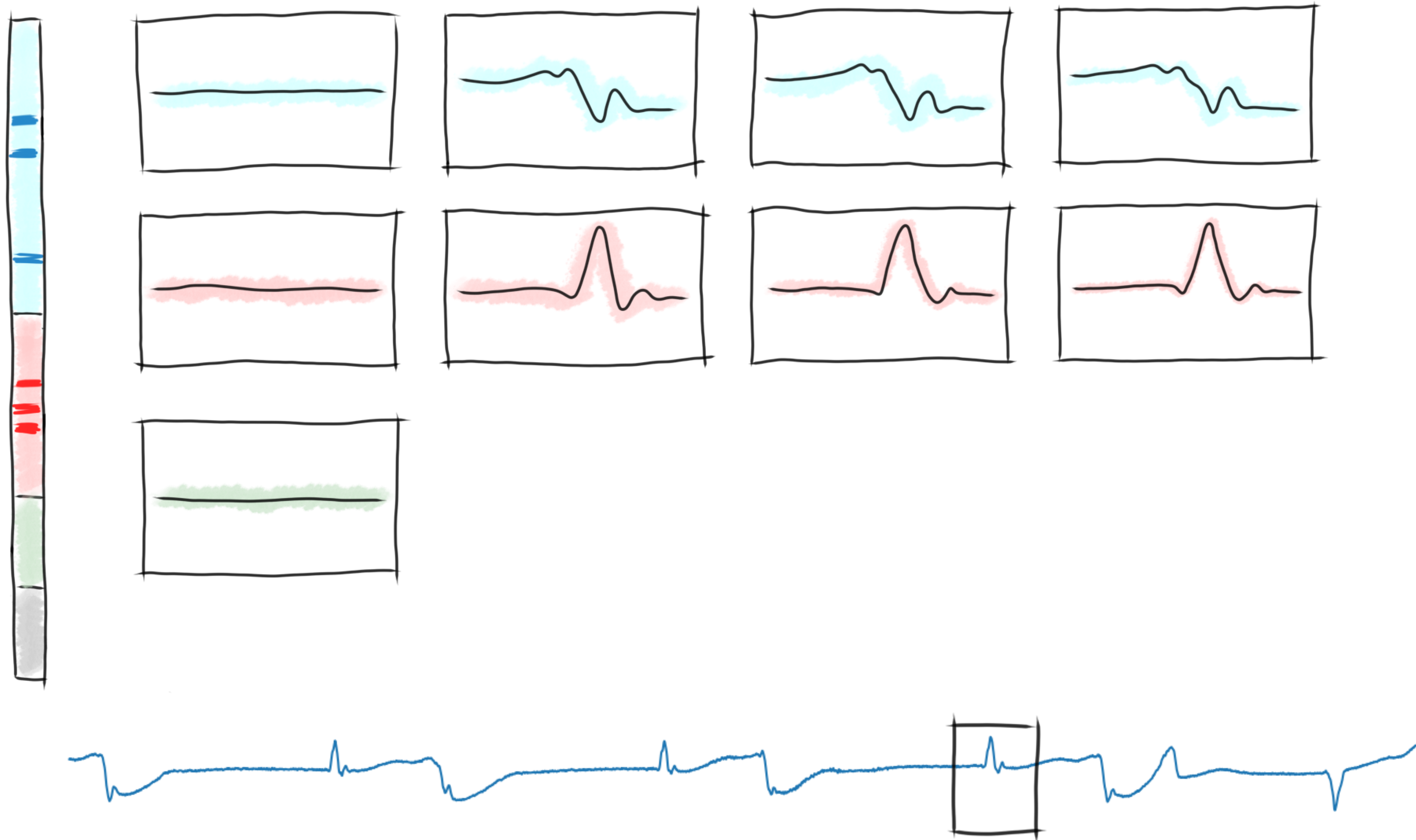
Hypothesis

My thesis in a napkin



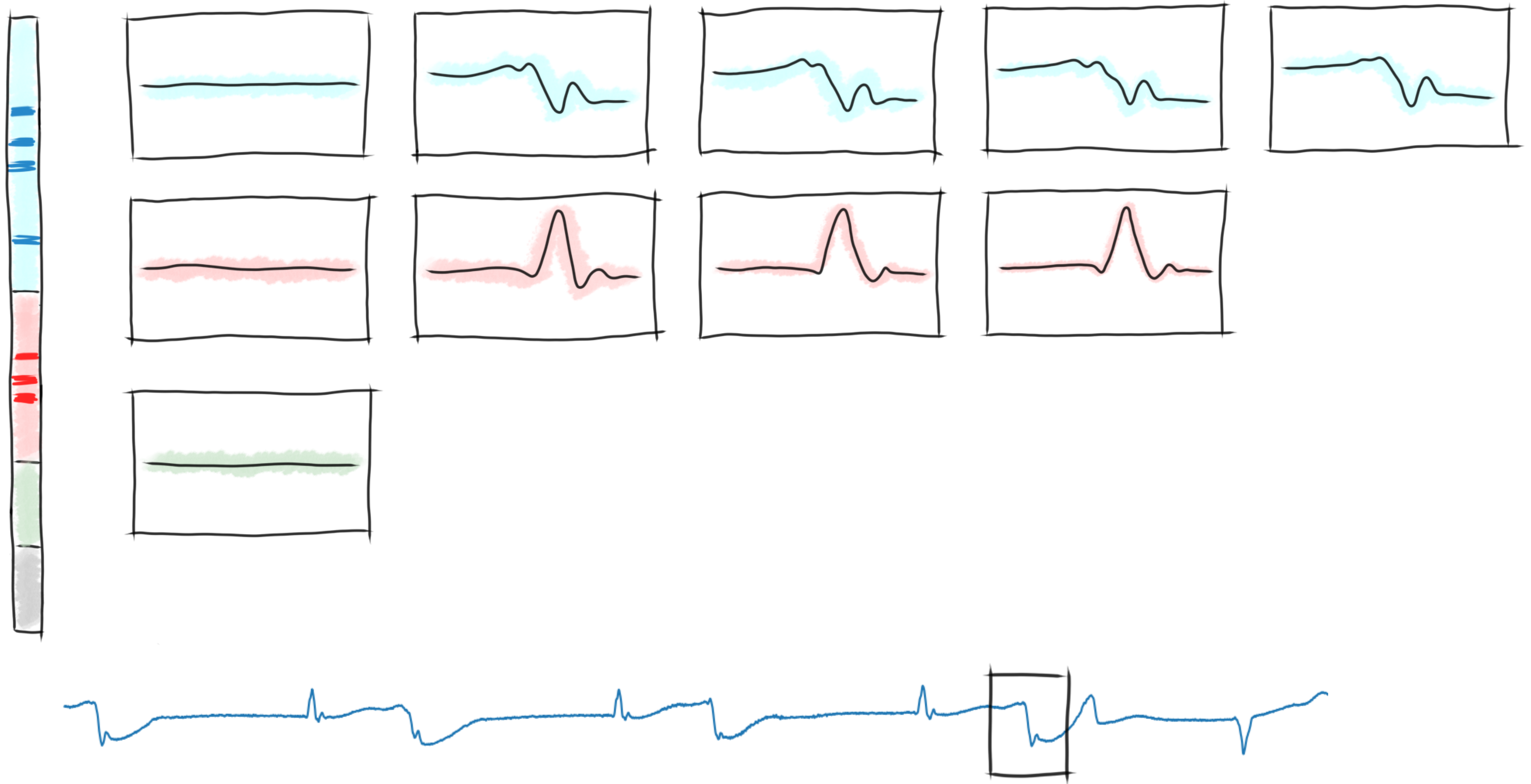
Hypothesis

My thesis in a napkin



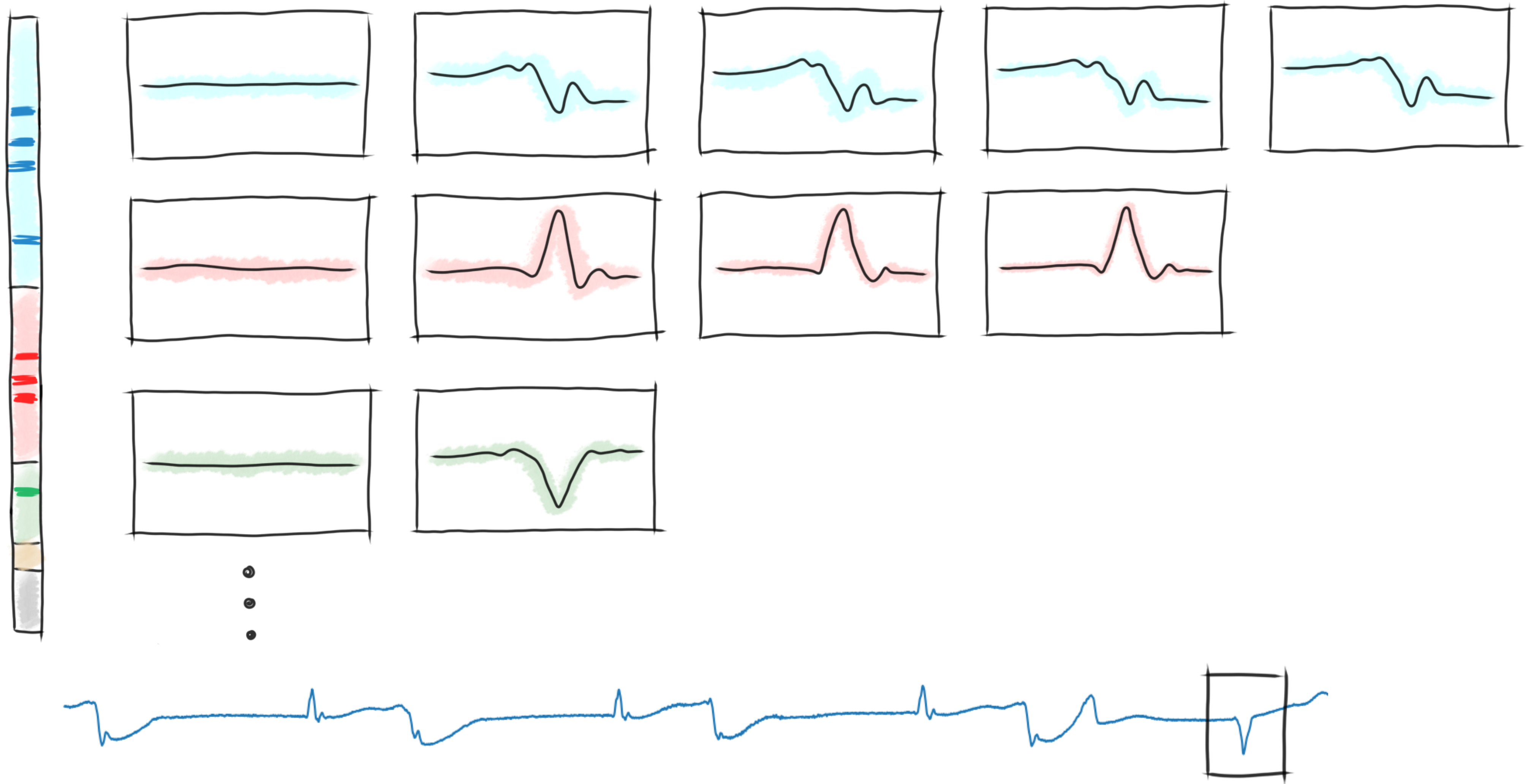
Hypothesis

My thesis in a napkin



Hypothesis

My thesis in a napkin



Hypothesis

Modelling

QUERY 1

GIVEN THE LAST ASSIGNMENT, WHAT SHOULD A FUTURE INSTANCE LOOK LIKE?

QUERY 2

GIVEN THE NEW OBSERVATION AND OUR PREDICTIONS, SHOULD WE ASSIGN IT TO ONE OF THE EXISTING CLUSTERS OR GENERATE A NEW ONE?

QUERY 3

CAN SOME DIFFERENCES ARISE FROM AN ALIGNMENT DISTORTION?

QUERY 4

ONCE WE HAVE DECIDED WHICH CLUSTER THE NEW OBSERVATION BELONGS TO, HOW SHOULD THIS CLUSTER BE UPDATED?

Hypothesis

Modelling

QUERY 1

GIVEN THE LAST ASSIGNMENT, WHAT SHOULD A FUTURE INSTANCE LOOK LIKE?

ALGORITHM 1

ALGORITHM 3

QUERY 3

CAN SOME DIFFERENCES ARISE FROM AN ALIGNMENT DISTORTION?

QUERY 2

GIVEN THE NEW OBSERVATION AND OUR PREDICTIONS, SHOULD WE ASSIGN IT TO ONE OF THE EXISTING CLUSTERS OR GENERATE A NEW ONE?

ALGORITHM 2

ALGORITHM 4

QUERY 4

ONCE WE HAVE DECIDED WHICH CLUSTER THE NEW OBSERVATION BELONGS TO, HOW SHOULD THIS CLUSTER BE UPDATED?

Hypothesis

Modelling

QUERY 1

GIVEN THE LAST ASSIGNMENT, WHAT SHOULD A FUTURE INSTANCE LOOK LIKE?

ALGORITHM 1

ALGORITHM 3

QUERY 3

CAN SOME DIFFERENCES ARISE FROM AN ALIGNMENT DISTORTION?

QUERY 2

GIVEN THE NEW OBSERVATION AND OUR PREDICTIONS, SHOULD WE ASSIGN IT TO ONE OF THE EXISTING CLUSTERS OR GENERATE A NEW ONE?

ALGORITHM 2

ALGORITHM 4

QUERY 4

ONCE WE HAVE DECIDED WHICH CLUSTER THE NEW OBSERVATION BELONGS TO, HOW SHOULD THIS CLUSTER BE UPDATED?

Hypothesis

Modelling

QUERY 1

GIVEN THE LAST ASSIGNMENT, WHAT SHOULD A FUTURE INSTANCE LOOK LIKE?

ALGORITHM 1

QUERY 2

GIVEN THE NEW OBSERVATION AND OUR PREDICTIONS, SHOULD WE ASSIGN IT TO ONE OF THE EXISTING CLUSTERS OR GENERATE A NEW ONE?

ALGORITHM 2

ALGORITHM 3

ALGORITHM 4

QUERY 3

CAN SOME DIFFERENCES ARISE FROM AN ALIGNMENT DISTORTION?

QUERY 4

ONCE WE HAVE DECIDED WHICH CLUSTER THE NEW OBSERVATION BELONGS TO, HOW SHOULD THIS CLUSTER BE UPDATED?

LOCALLY REASONABLE, GLOBALLY INCONSISTENT!

Hypothesis

Modelling

QUERY 1

GIVEN THE LAST ASSIGNMENT, WHAT SHOULD A FUTURE INSTANCE LOOK LIKE?

QUERY 2

GIVEN THE NEW OBSERVATION AND OUR PREDICTIONS, SHOULD WE ASSIGN IT TO ONE OF THE EXISTING CLUSTERS OR GENERATE A NEW ONE?

MODEL 1

QUERY 3

CAN SOME DIFFERENCES ARISE FROM AN ALIGNMENT DISTORTION?

QUERY 4

ONCE WE HAVE DECIDED WHICH CLUSTER THE NEW OBSERVATION BELONGS TO, HOW SHOULD THIS CLUSTER BE UPDATED?

Hypothesis

Modelling

QUERY 1

GIVEN THE LAST ASSIGNMENT, WHAT SHOULD A FUTURE INSTANCE LOOK LIKE?

QUERY 2

GIVEN THE NEW OBSERVATION AND OUR PREDICTIONS, SHOULD WE ASSIGN IT TO ONE OF THE EXISTING CLUSTERS OR GENERATE A NEW ONE?

GENERATIVE
MODELLING
PERSPECTIVE

QUERY 3

CAN SOME DIFFERENCES ARISE FROM AN ALIGNMENT DISTORTION?

QUERY 4

ONCE WE HAVE DECIDED WHICH CLUSTER THE NEW OBSERVATION BELONGS TO, HOW SHOULD THIS CLUSTER BE UPDATED?

Scientific Hypothesis

10

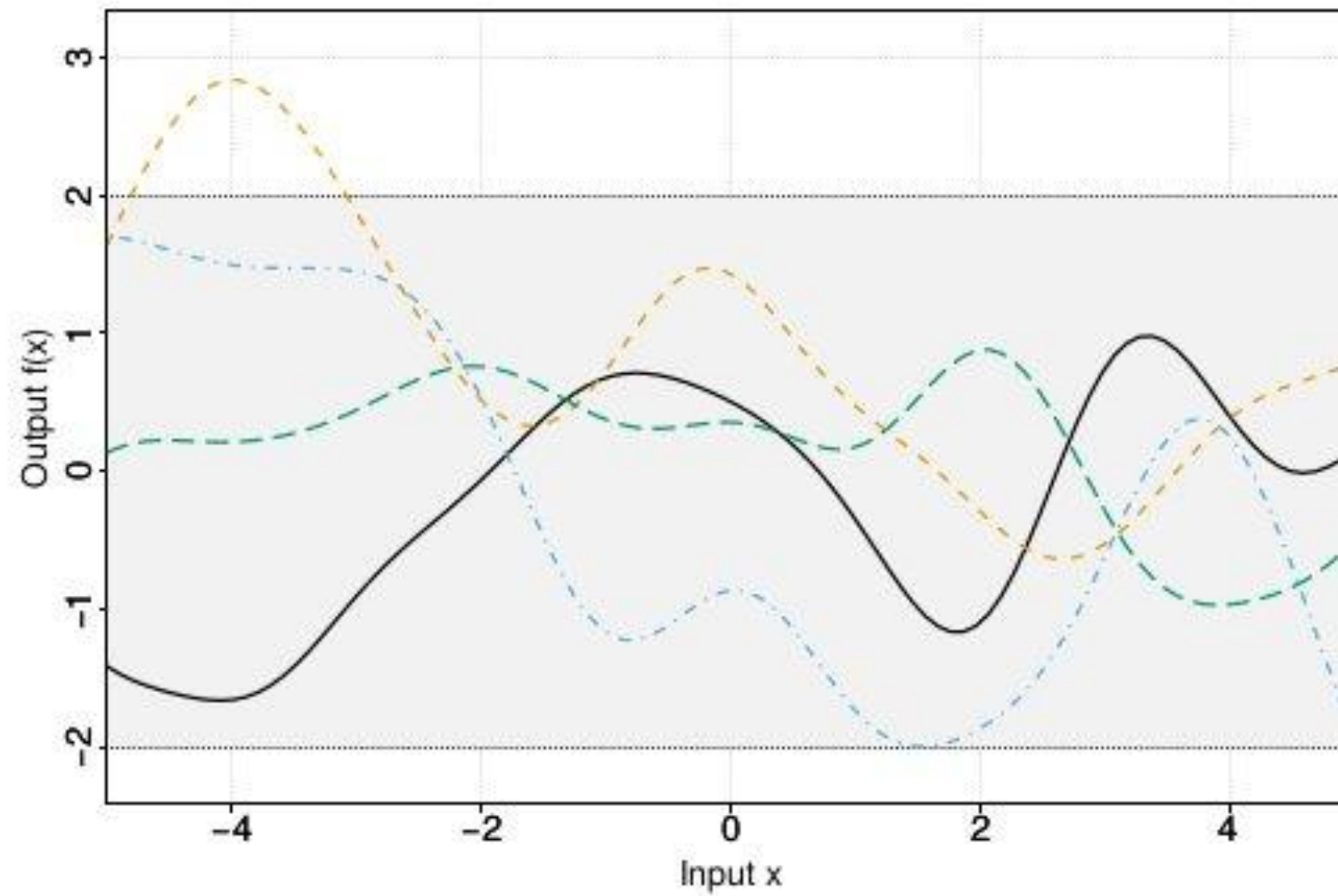
We hypothesize that a **Bayesian non-parametric** ($k=\infty$) approach with a **hierarchical structure**, where emission distributions are represented by **Gaussian processes**, and the latent space is characterised by **Markovian linear dynamics**, can adequately model the evolution of a previously unknown number of morphologies in time series, by splitting them based on similarity grounds.

Background

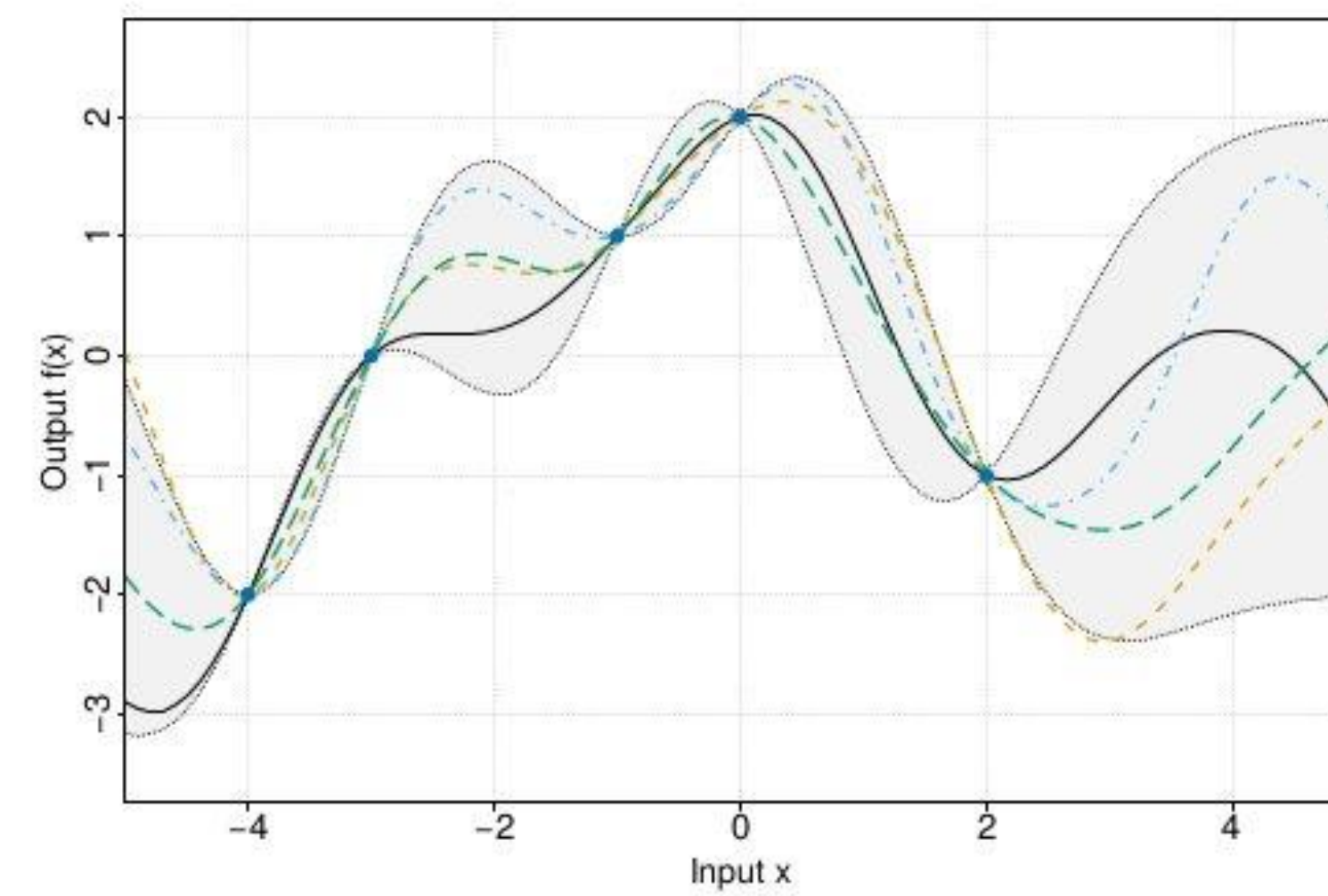
Background

Gaussian Processes (Rasmussen & Williams, 2006)

$$f(t) \sim \mathcal{GP}(m(t), k(t, t'))$$



Prior

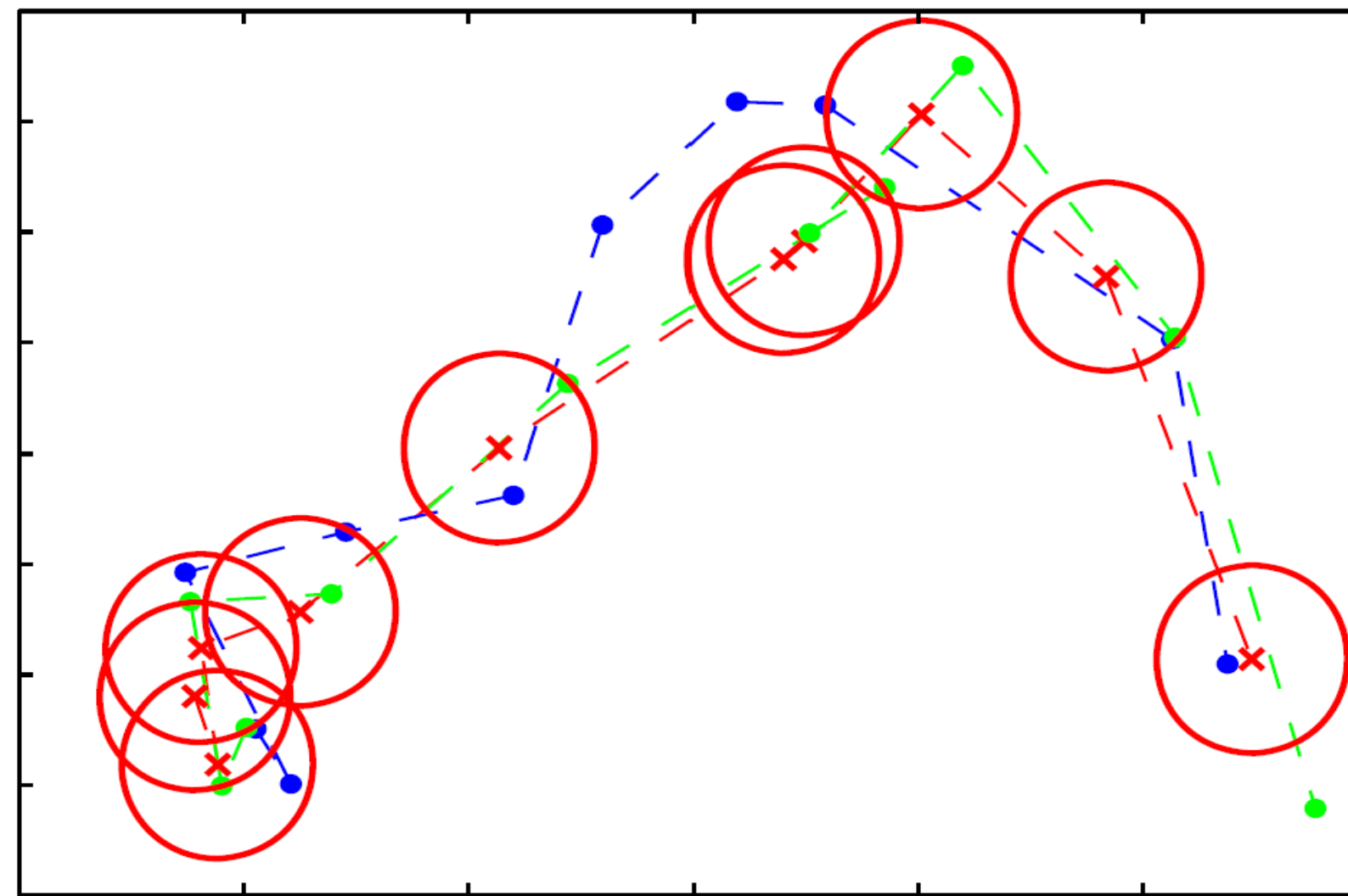


Posterior

Background

Linear Dynamical System (Kalman, 1960)

$$\begin{aligned}\mathbf{x}_n &= \mathbf{A}\mathbf{x}_{n-1} + \omega_n & \omega_n &\sim \mathcal{N}(0, \Sigma_\omega) \\ \mathbf{y}_n &= \mathbf{C}\mathbf{x}_n + \epsilon_n & \epsilon_n &\sim \mathcal{N}(0, \Sigma_\epsilon)\end{aligned}$$



Background

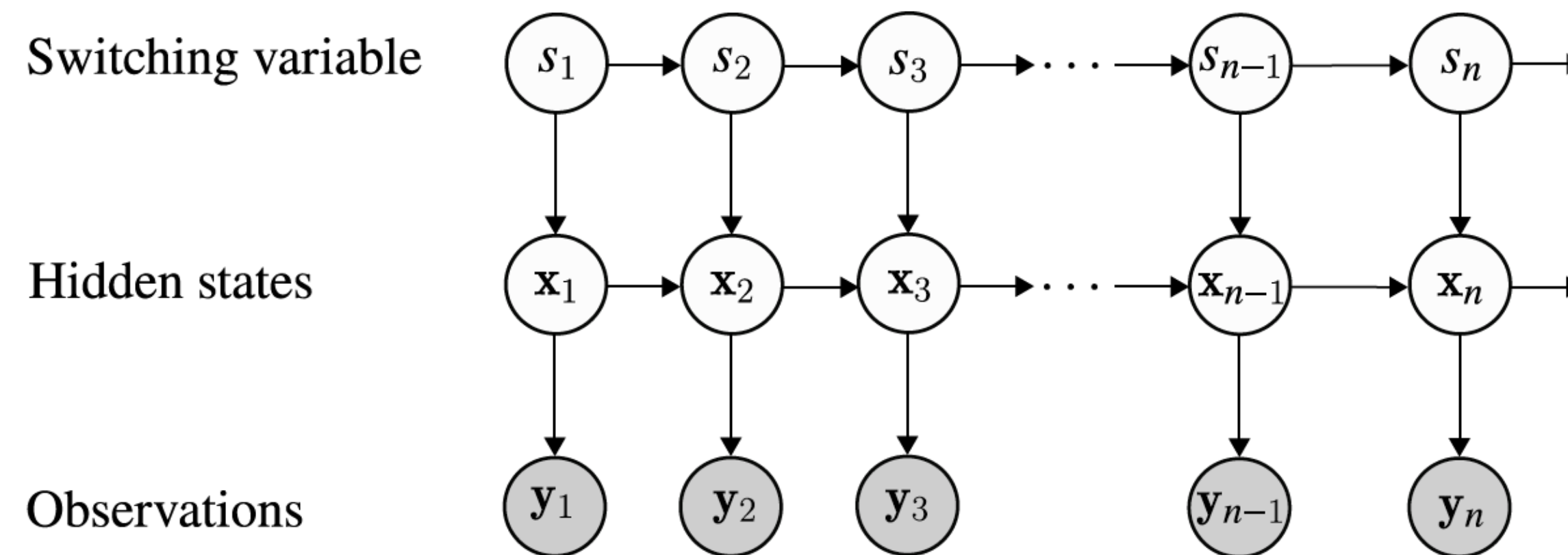
Switching Linear Dynamical System (Ghahramani & Hinton, 2000)

$$s_n \sim \pi^{(s_{n-1})}$$

$$\mathbf{x}_n = \mathbf{A}^{(s_n)} \mathbf{x}_{n-1} + \omega_n^{(s_n)} \quad \omega_n^{(s_n)} \sim \mathcal{N}(0, \Sigma_{\omega}^{(s_n)})$$

$$\mathbf{y}_n = \mathbf{C}^{(s_n)} \mathbf{x}_n + \epsilon_n^{(s_n)} \quad \epsilon_n^{(s_n)} \sim \mathcal{N}(0, \Sigma_{\epsilon}^{(s_n)})$$

13



Background

Dirichlet Processes (Ferguson, 1973)

$$\text{DP} \quad | \quad \beta \sim \text{GEM}(\gamma),$$



Background

Hierarchical Dirichlet Processes (Teh et al., 2006)

HDP

$$\begin{aligned}\beta &\sim \text{GEM}(\gamma), \\ \pi^{(j)} &\stackrel{iid}{\sim} \text{DP}(\alpha\beta) \\ s_n &\sim \pi^{(s_{n-1})}\end{aligned}$$

β



$\pi^{(1)}$



$\pi^{(2)}$



$\pi^{(3)}$



$\vdots$

The HDP-GPC model

The HDP-GPC model

The hierarchical Dirichlet process – Gaussian process clustering model (HDP-GPC).

$$\beta \sim \text{GEM}(\gamma),$$

$$\pi^{(j)} \stackrel{iid}{\sim} \text{DP}(\alpha\beta)$$

$$\phi^{(j)} \stackrel{iid}{\sim} H_{\text{LDS-D}}$$

$$s_n \sim \pi^{(s_{n-1})}$$

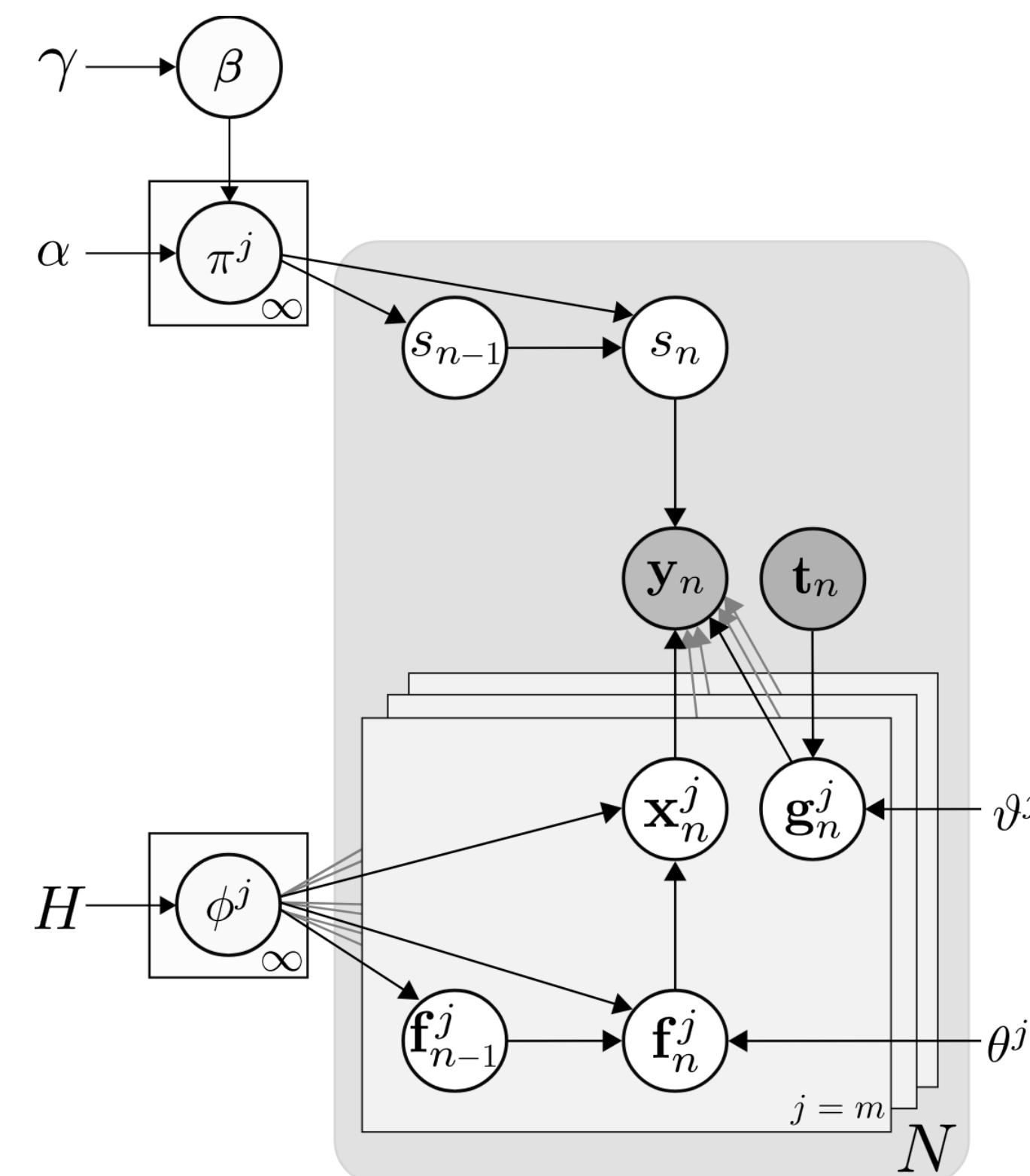
$$\mathbf{f}_1^{(s_n)} \sim \mathcal{GP}(m_{\theta(s_n)}(\mathbf{t}_1), k_{\theta(s_n)}(\mathbf{t}_1, \mathbf{t}_1))$$

$$\mathbf{f}_n^{(s_n)} = A^{(s_n)} \mathbf{f}_{n-1}^{(s_n)} + \omega_n^{(s_n)}$$

$$\mathbf{x}_{n_i}^{(s_n)} = C^{(s_n)} \mathbf{f}_{n_i}^{(s_n)} + \epsilon_{n_i}^{(s_n)}$$

$$\mathbf{t}_n^w = \mathbf{g}_n^{(s_n)}(\mathbf{t}_n) \sim \mathcal{GP}(m_{\mathcal{Y}^k}(\mathbf{t}_n), k_{\mathcal{Y}^k}(\mathbf{t}_n, \mathbf{t}_n))$$

$$\mathbf{y}_n \sim \mathcal{N}(m_{x^{(s_n)}}(\mathbf{t}_n^w), K_{\mathbf{t}_n^w, \mathbf{t}_n^w}^{x^{(s_n)}})$$



The HDP-GPC model

Generative model

HDP

$$\beta \sim \text{GEM}(\gamma),$$

$$\pi^{(j)} \stackrel{iid}{\sim} \text{DP}(\alpha\beta)$$

$$\phi^{(j)} \stackrel{iid}{\sim} H_{\text{LDS-D}}$$

$$s_n \sim \pi^{(s_{n-1})}$$

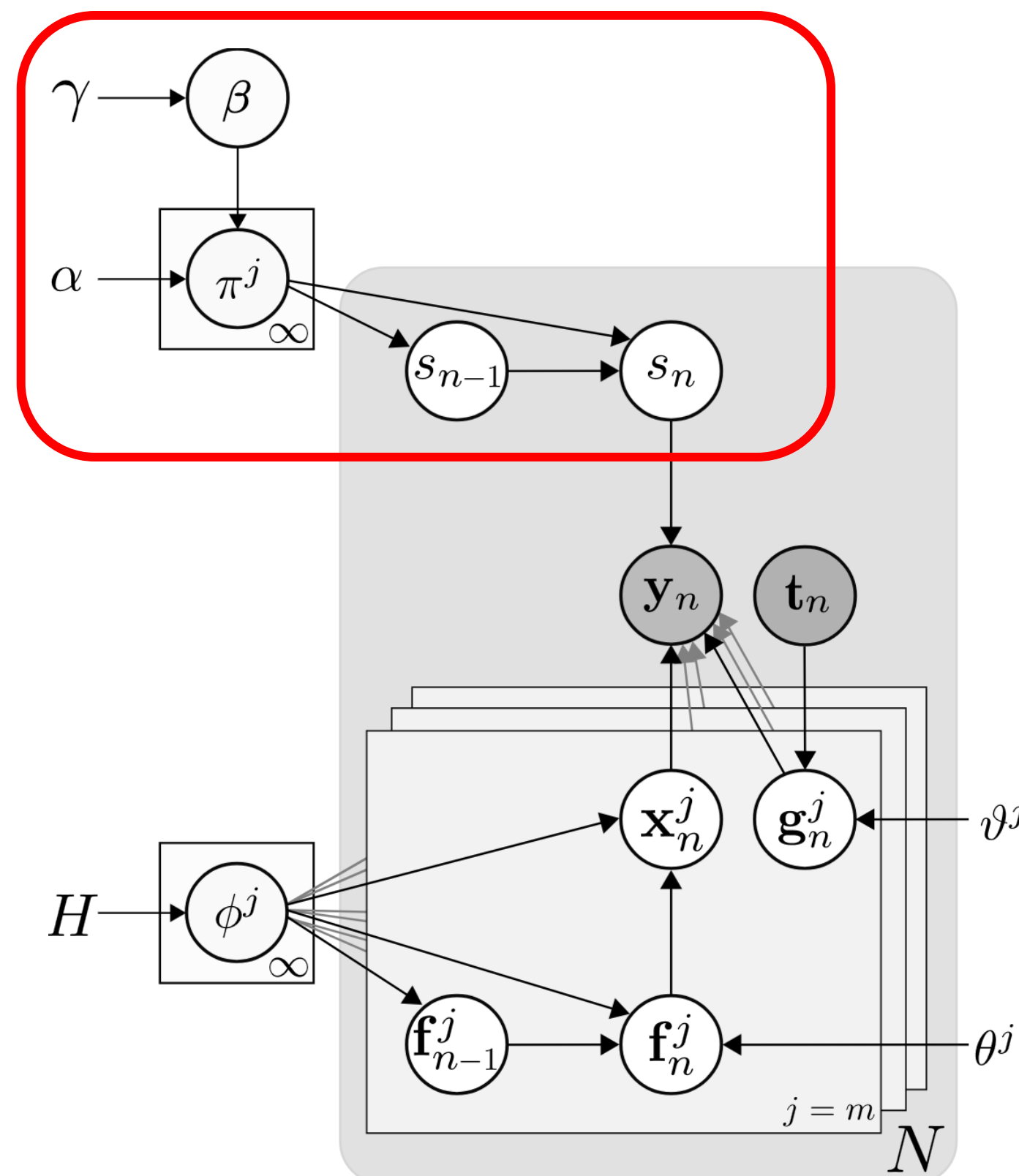
$$\mathbf{f}_1^{(s_n)} \sim \mathcal{GP}(m_{\theta^{(s_n)}}(\mathbf{t}_1), k_{\theta^{(s_n)}}(\mathbf{t}_1, \mathbf{t}_1))$$

$$\mathbf{f}_n^{(s_n)} = A^{(s_n)} \mathbf{f}_{n-1}^{(s_n)} + \omega_n^{(s_n)}$$

$$\mathbf{x}_n^{(s_n)} = C^{(s_n)} \mathbf{f}_n^{(s_n)} + \epsilon_n^{(s_n)}$$

$$\mathbf{t}_n^w = \mathbf{g}_n^{(s_n)}(\mathbf{t}_n) \sim \mathcal{GP}(m_{\vartheta^k}(\mathbf{t}_n), k_{\vartheta^k}(\mathbf{t}_n, \mathbf{t}_n))$$

$$\mathbf{y}_n \sim \mathcal{N}(m_{x^{(s_n)}}(\mathbf{t}_n^w), K_{\mathbf{t}_n^w, \mathbf{t}_n^w}^{x^{(s_n)}})$$



The HDP-GPC model

Generative model

$$\beta \sim \text{GEM}(\gamma),$$

$$\pi^{(j)} \stackrel{iid}{\sim} \text{DP}(\alpha\beta)$$

$$\phi^{(j)} \stackrel{iid}{\sim} H_{\text{LDS-D}}$$

$$s_n \sim \pi^{(s_{n-1})}$$

GP

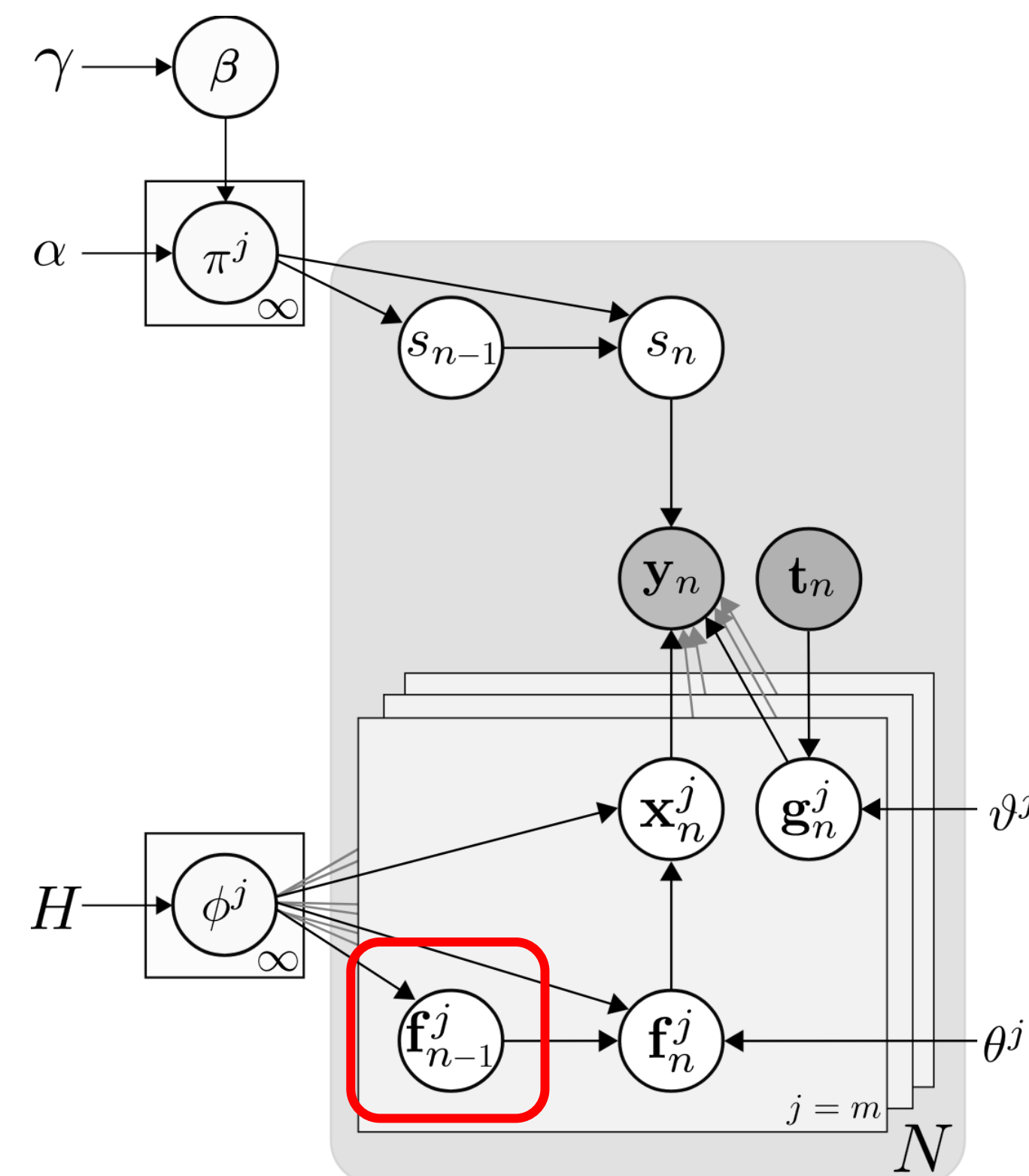
$$\mathbf{f}_1^{(s_n)} \sim \mathcal{GP}(m_{\theta^{(s_n)}}(\mathbf{t}_1), k_{\theta^{(s_n)}}(\mathbf{t}_1, \mathbf{t}_1))$$

$$\mathbf{f}_n^{(s_n)} = A^{(s_n)} \mathbf{f}_{n-1}^{(s_n)} + \omega_n^{(s_n)}$$

$$\mathbf{x}_n^{(s_n)} = C^{(s_n)} \mathbf{f}_n^{(s_n)} + \epsilon_n^{(s_n)}$$

$$\mathbf{t}_n^w = \mathbf{g}_n^{(s_n)}(\mathbf{t}_n) \sim \mathcal{GP}(m_{\vartheta^k}(\mathbf{t}_n), k_{\vartheta^k}(\mathbf{t}_n, \mathbf{t}_n))$$

$$\mathbf{y}_n \sim \mathcal{N}(m_{x^{(s_n)}}(\mathbf{t}_n^w), K_{\mathbf{t}_n^w, \mathbf{t}_n^w}^{x^{(s_n)}})$$



The HDP-GPC model

Generative model

$$\beta \sim \text{GEM}(\gamma),$$

$$\pi^{(j)} \stackrel{iid}{\sim} \text{DP}(\alpha\beta)$$

$$\phi^{(j)} \stackrel{iid}{\sim} H_{\text{LDS-D}}$$

$$s_n \sim \pi^{(s_{n-1})}$$

$$\mathbf{f}_1^{(s_n)} \sim \mathcal{GP}(m_{\theta^{(s_n)}}(\mathbf{t}_1), k_{\theta^{(s_n)}}(\mathbf{t}_1, \mathbf{t}_1))$$

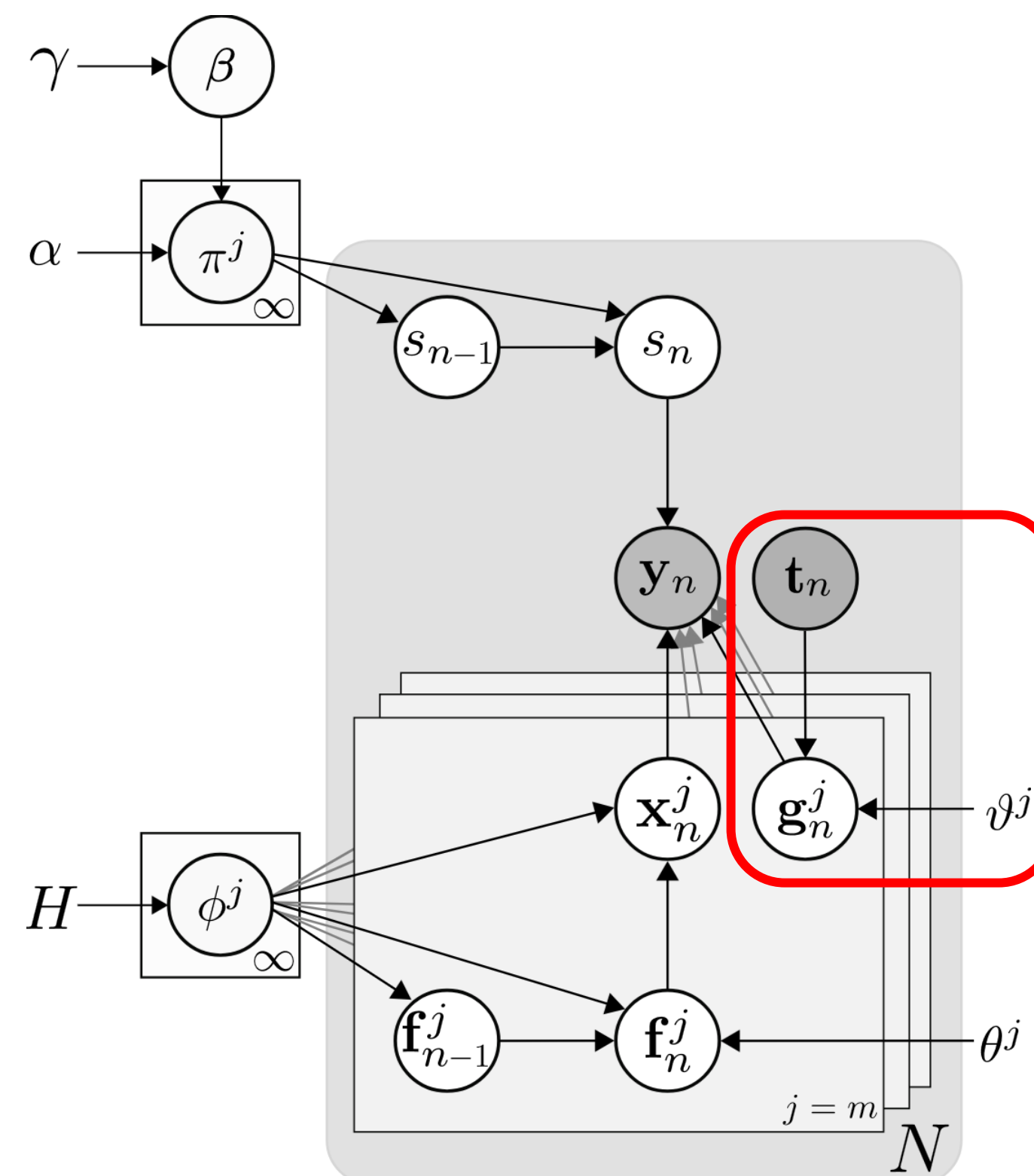
$$\mathbf{f}_n^{(s_n)} = A^{(s_n)} \mathbf{f}_{n-1}^{(s_n)} + \omega_n^{(s_n)}$$

$$\mathbf{x}_n^{(s_n)} = C^{(s_n)} \mathbf{f}_n^{(s_n)} + \epsilon_n^{(s_n)}$$

Alignment GP

$$\mathbf{t}_n^w = \mathbf{g}_n^{(s_n)}(\mathbf{t}_n) \sim \mathcal{GP}(m_{\vartheta^k}(\mathbf{t}_n), k_{\vartheta^k}(\mathbf{t}_n, \mathbf{t}_n))$$

$$\mathbf{y}_n \sim \mathcal{N}(m_{x^{(s_n)}}(\mathbf{t}_n^w), K_{\mathbf{t}_n^w, \mathbf{t}_n^w}^{x^{(s_n)}})$$



The HDP-GPC model

Generative model

$$\beta \sim \text{GEM}(\gamma),$$

$$\pi^{(j)} \stackrel{iid}{\sim} \text{DP}(\alpha\beta)$$

$$\phi^{(j)} \stackrel{iid}{\sim} H_{\text{LDS-D}}$$

$$s_n \sim \pi^{(s_{n-1})}$$

$$\mathbf{f}_1^{(s_n)} \sim \mathcal{GP}(m_{\theta^{(s_n)}}(\mathbf{t}_1), k_{\theta^{(s_n)}}(\mathbf{t}_1, \mathbf{t}_1))$$

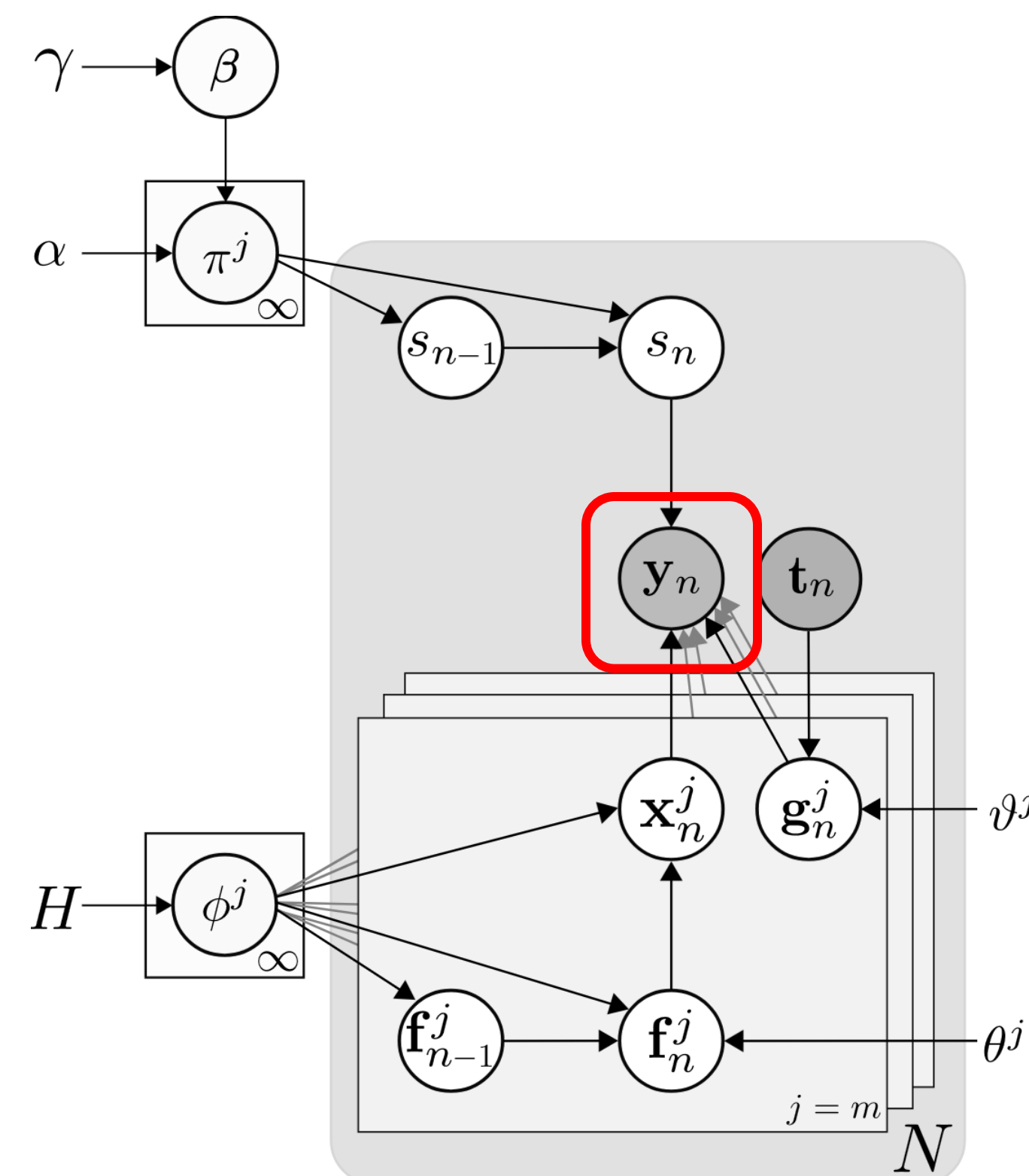
$$\mathbf{f}_n^{(s_n)} = A^{(s_n)} \mathbf{f}_{n-1}^{(s_n)} + \omega_n^{(s_n)}$$

$$\mathbf{x}_n^{(s_n)} = C^{(s_n)} \mathbf{f}_n^{(s_n)} + \epsilon_n^{(s_n)}$$

$$\mathbf{t}_n^w = \mathbf{g}_n^{(s_n)}(\mathbf{t}_n) \sim \mathcal{GP}(m_{\vartheta^k}(\mathbf{t}_n), k_{\vartheta^k}(\mathbf{t}_n, \mathbf{t}_n))$$

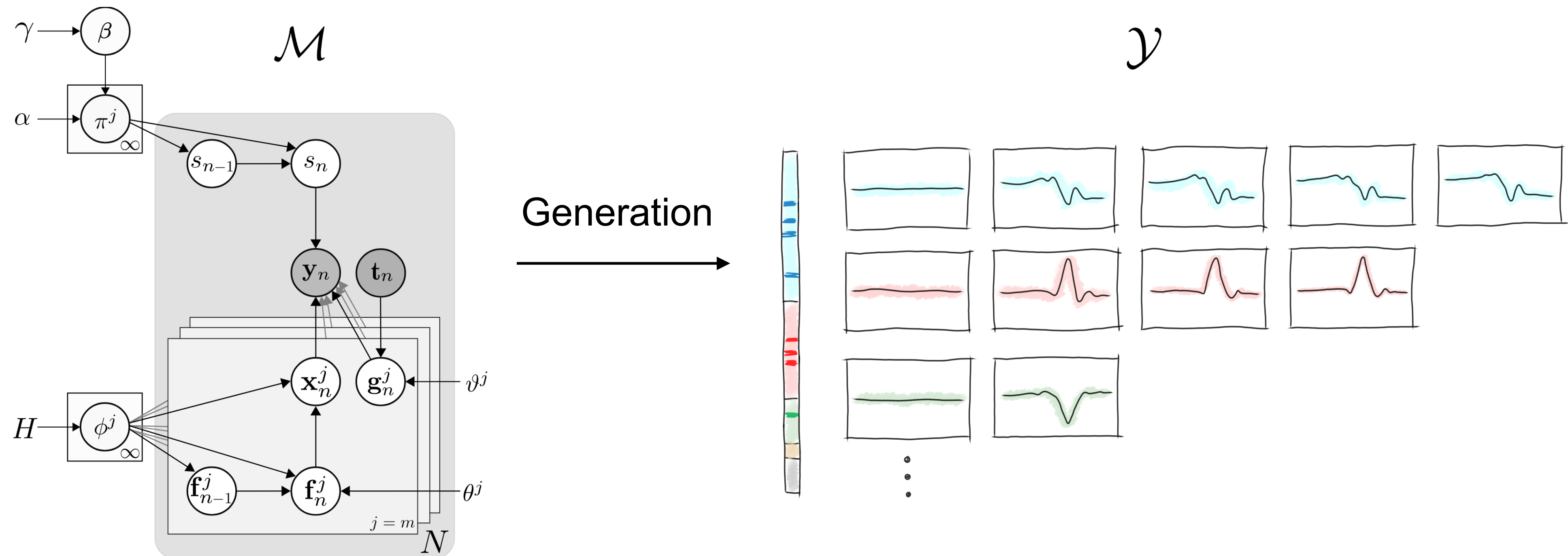
Observation

$$\mathbf{y}_n \sim \mathcal{N}(m_{x^{(s_n)}}(\mathbf{t}_n^w), K_{\mathbf{t}_n^w, \mathbf{t}_n^w}^{x^{(s_n)}})$$



The HDP-GPC model

Generative model



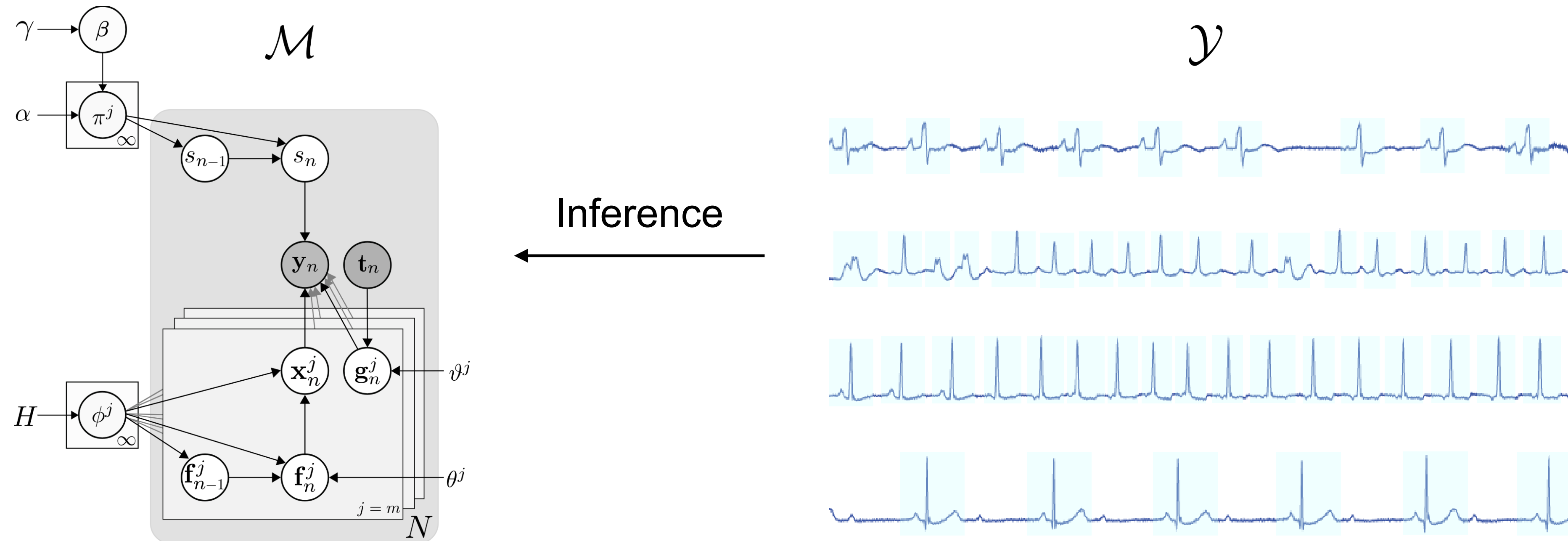
Inference

Inference

What is inference?

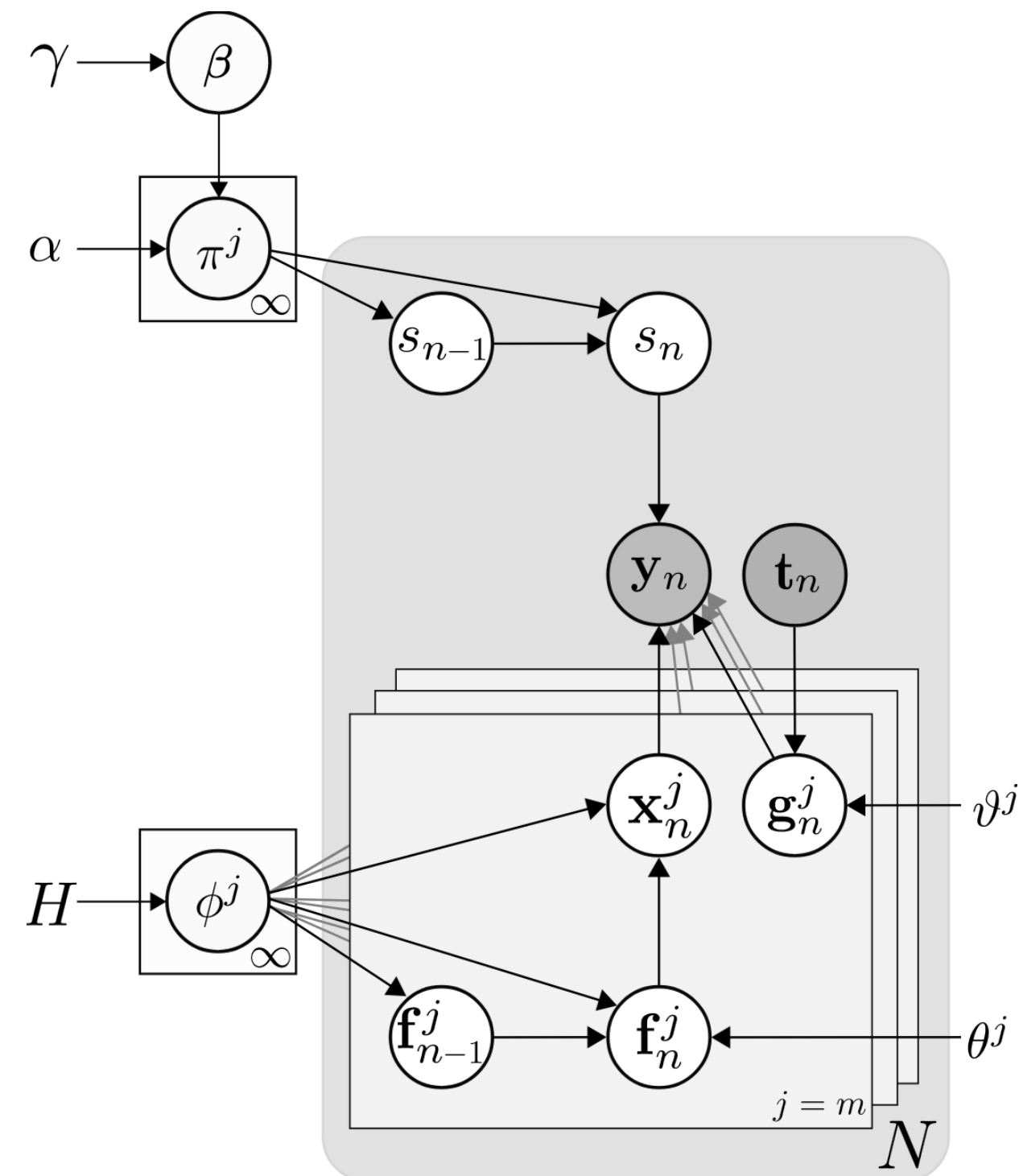
Our goal is to infer the posterior probability of the hidden variables conditioned on the time sequences.

$$p(\mathcal{M}|\mathcal{Y}) \propto p(\mathcal{Y}|\mathcal{M})p(\mathcal{M})$$



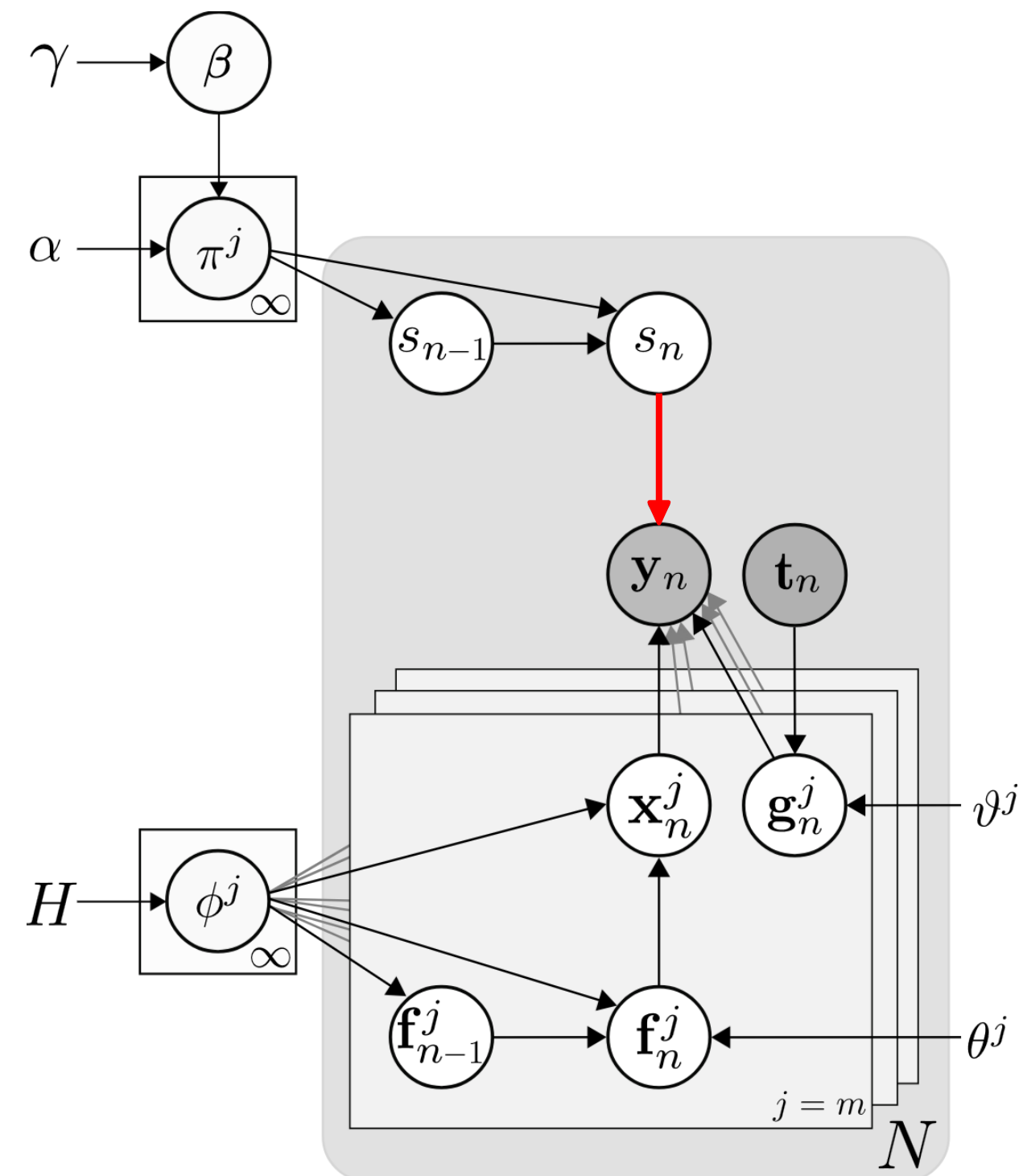
Variational Inference

Exact inference is intractable



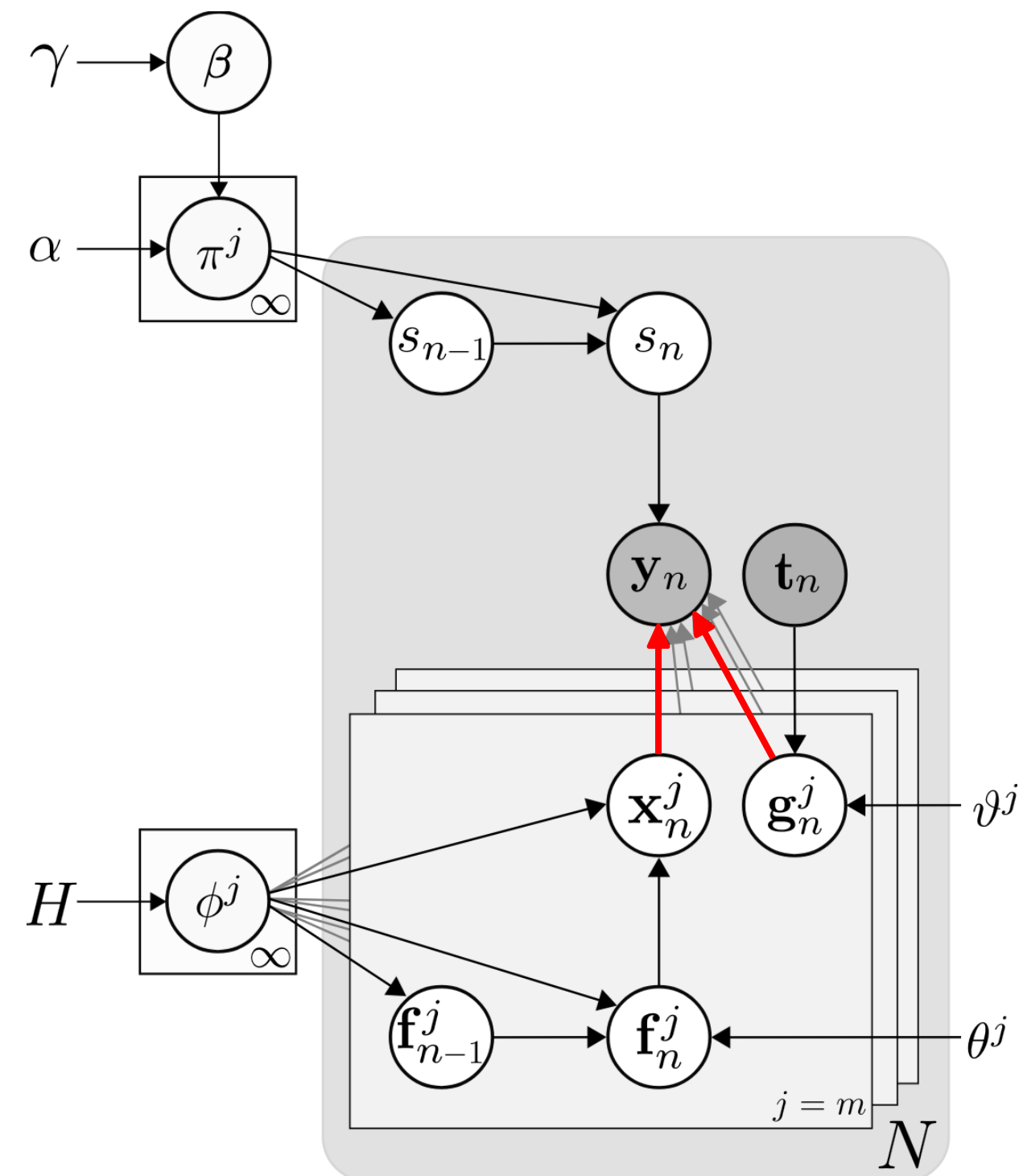
Variational Inference

Exact inference is intractable: switching variable couples the dynamics



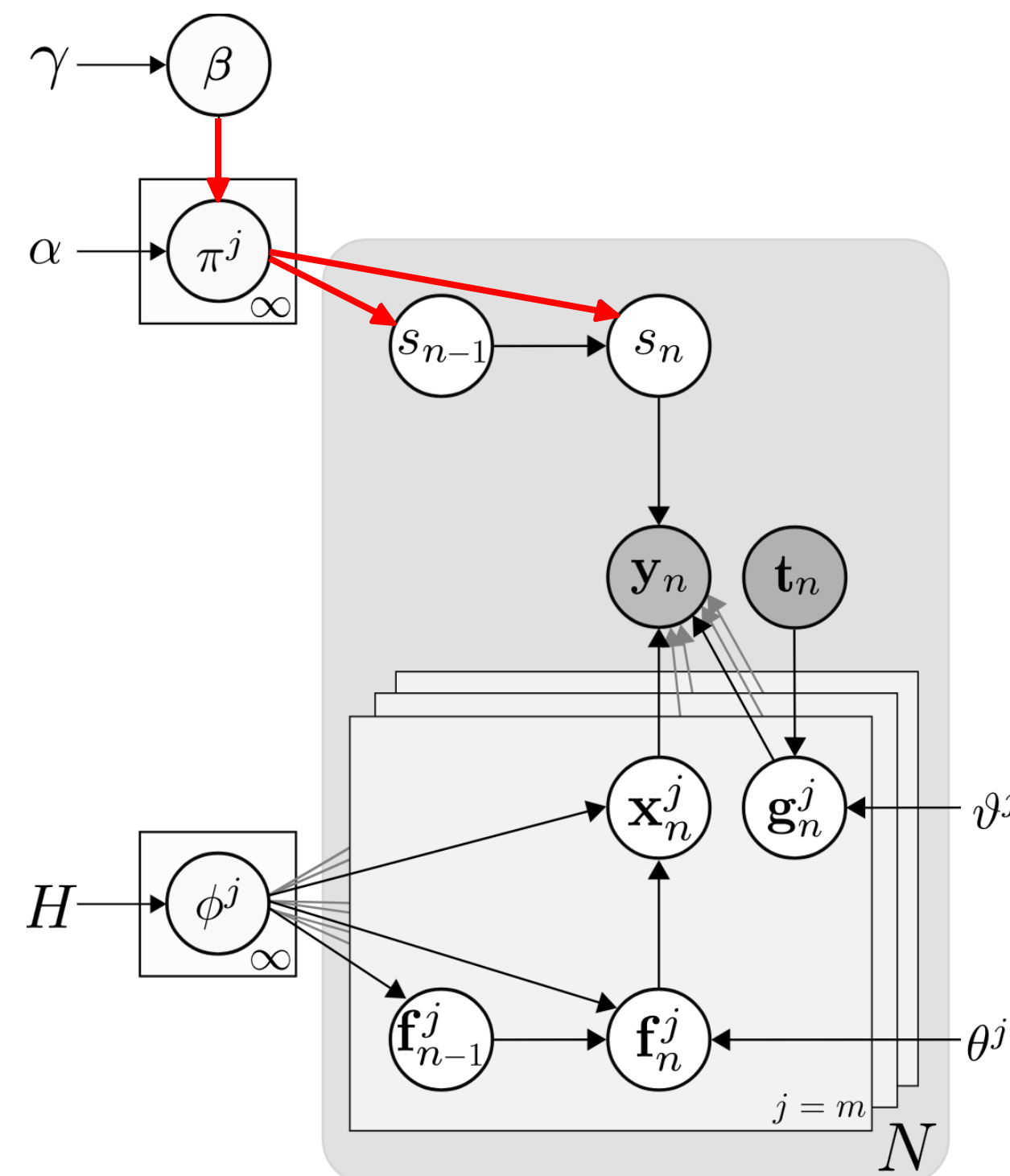
Variational Inference

Exact inference is intractable: Gaussian Process composition



Variational Inference

Exact inference is intractable: HDP posterior is not available in closed form



Variational Inference

Inference method

Exact inference is intractable, methods as Gibbs sampling or stochastic variational inference could be considered. But in our case, we have sequential structure to exploit, so we adopt a scalable variational inference approach which looks for minimizing the KL divergence:

$$\log p(\mathcal{Y}) = \mathcal{L}(q) + \text{KL}(q||p)$$

Equivalent to maximize:

$$\mathcal{L}(q) = \mathbb{E}_q [\log p(\mathcal{M}, \mathcal{Y}) - \log q(\mathcal{M})]$$

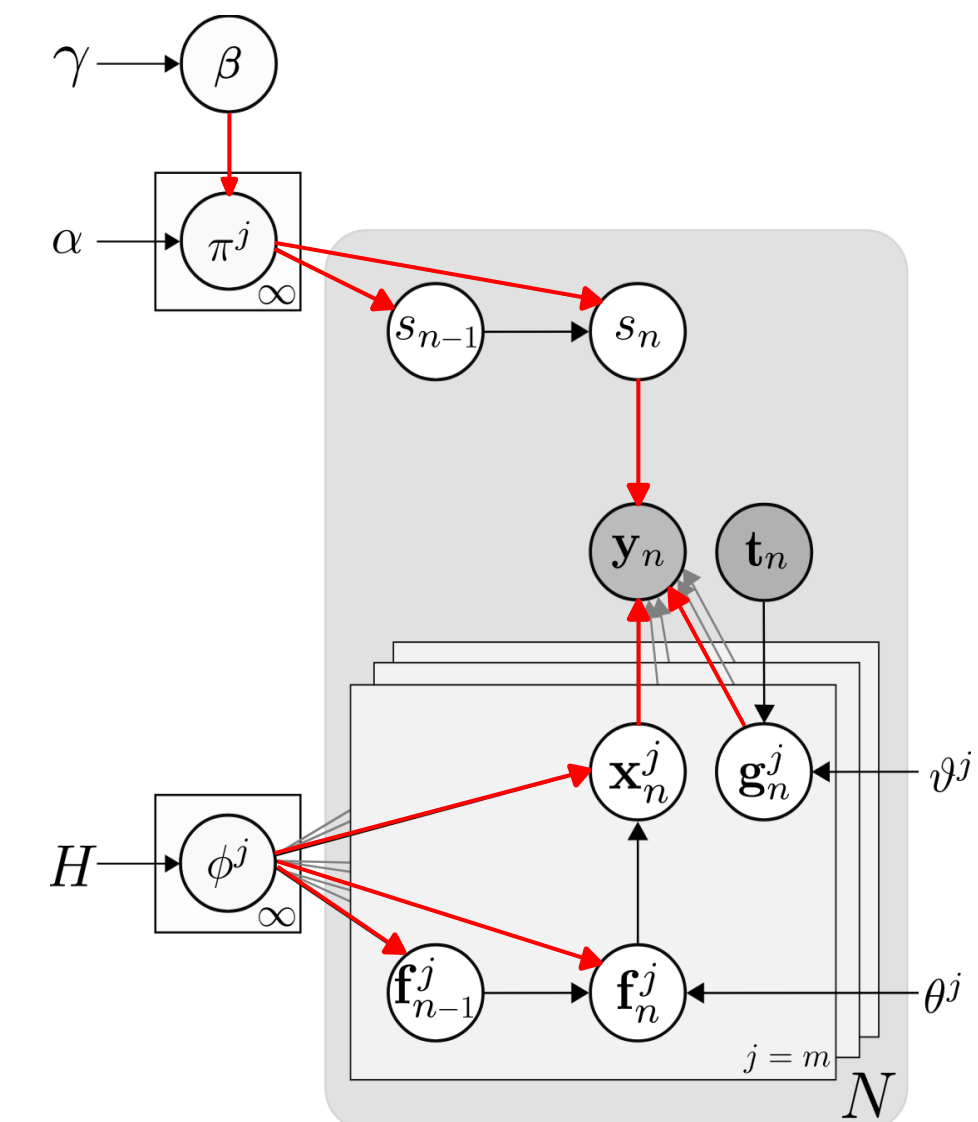
The mean-field approximation is:

$$q(\cdot) = q(\beta)q(\pi)q(\phi)q(\mathcal{S})q(\mathcal{F})q(\mathcal{X})q(\mathcal{G})$$

With $\mathcal{S} = \{s_1, \dots, s_N\}$, $\mathcal{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_N\}$, $\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ and
and $\mathcal{G} = \{\mathbf{g}_1, \dots, \mathbf{g}_N\}$.

And the optimum family is computed as:

$$q_j^*(\mathbf{z}_j) \propto \exp\{\mathbb{E}_{i \neq j} [\log p(\mathcal{M}, \mathcal{Y})]\}$$



Variational Inference

Scalable variational inference approach.

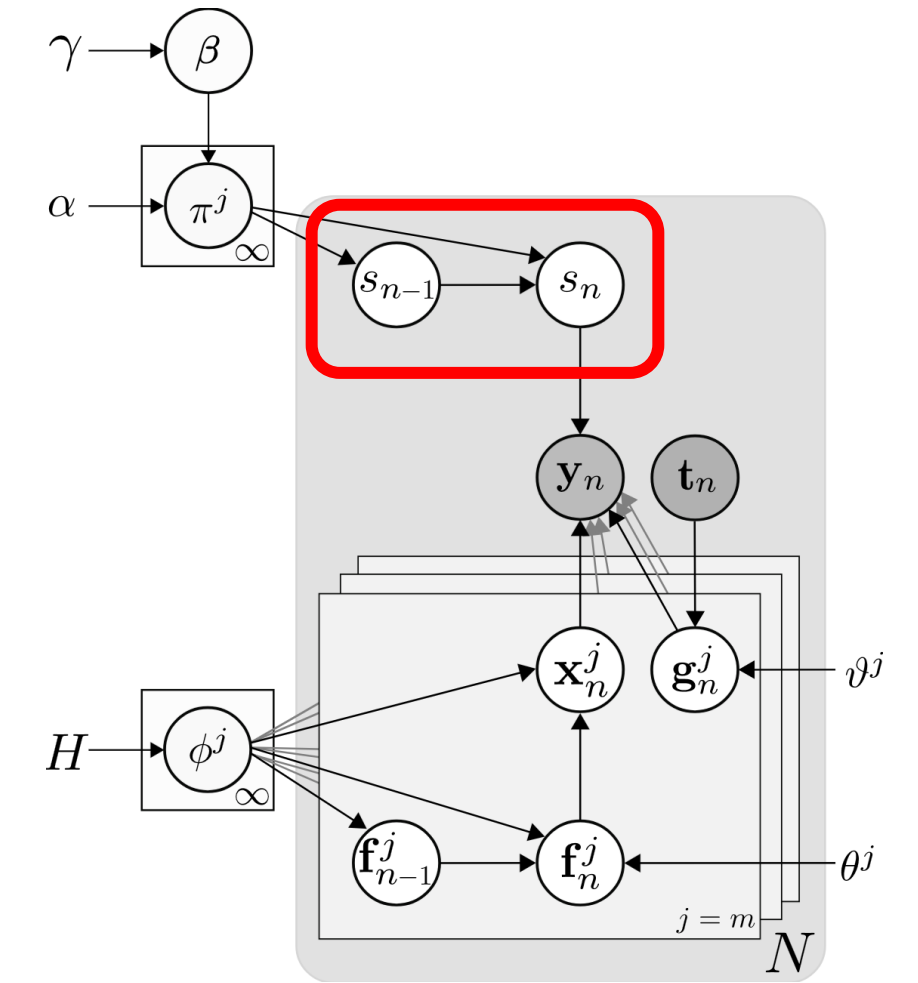
Factor	Latent variables	Variational family
$q(\beta)$	Stick-breaking variables	$\prod_{k=1}^K \text{Beta}(v_k \lambda_k \eta_k, (1 - \lambda_k) \eta_k)$
$q(\pi)$	Transition distributions	$\prod_{j=0}^K \text{Dir}(\pi^{(j)} \kappa_{j1}, \dots, \kappa_{j,K+1})$
$q(\mathcal{S})$	Switching states	$\prod_{n=1}^N \prod_{j=1}^K \prod_{k=1}^K \xi_{njk}^{s_{n-1,j} s_{nk}}$
$q(\mathcal{G})$	Warping functions	$\prod_{n=1}^N \prod_{k=1}^K \delta_{\hat{\mathbf{g}}_n^{(k)}}(\mathbf{g}_n^{(k)})$
$q(\phi)$	LDS parameters	$\prod_{k=1}^K q(\mathbf{A}^{(k)}, \Sigma_{\omega}^{(k)}) q(\mathbf{C}^{(k)}, \Sigma_{\epsilon}^{(k)})$
$q(\mathbf{A}^{(k)}, \Sigma_{\omega}^{(k)})$	Latent transition parameters	$\mathcal{MNIW}(\tilde{\mathbf{M}}_A^{(k)}, \tilde{\mathbf{W}}_A^{(k)}, \tilde{\mathbf{V}}_A^{(k)}, \tilde{v}_A^{(k)}, \tilde{\mathbf{S}}_{\omega}^{(k)})$
$q(\mathbf{C}^{(k)}, \Sigma_{\epsilon}^{(k)})$	Emission parameters	$\mathcal{MNIW}(\tilde{\mathbf{M}}_C^{(k)}, \tilde{\mathbf{W}}_C^{(k)}, \tilde{\mathbf{V}}_C^{(k)}, \tilde{v}_C^{(k)}, \tilde{\mathbf{S}}_{\epsilon}^{(k)})$
$q(\mathcal{F}, \mathcal{X})$	Latent and emission processes	$\prod_{k=1}^K q(\mathcal{X}^{(k)} \mathcal{F}^{(k)}) q(\mathcal{F}^{(k)})$

Variational Inference

Switching variable

Follows a Markov chain factorization:

$$q^*(\mathcal{S}) = \prod_{n=1}^N \prod_{j=1}^K \prod_{k=1}^K \xi_{njk}^{s_{n-1}, j, s_{nk}} \quad \text{where} \quad \xi_{njk} = q(s_{n-1} = j, s_n = k)$$



EXTRA

We compute

$$\log \zeta_{njk} = \mathbb{E}[\log \pi_{jk}] - \frac{1}{2} \mathbb{E}[\log |K_{\mathbf{g}_n^k, \mathbf{g}_n^k}^{x_n^k}|] - \frac{|\mathbf{y}_n|}{2} \log(2\pi) - \frac{1}{2} \mathbb{E}_{x_n^k, \mathbf{g}_n^k}[(\mathbf{y}_n - m_{x_n^k}(\mathbf{g}_n^k))^{\top} K_{\mathbf{g}_n^k, \mathbf{g}_n^k}^{x_n^k-1} (\mathbf{y}_n - m_{x_n^k}(\mathbf{g}_n^k))]$$

Then use ζ_{njk} to perform **forward-backward recursion** to compute ξ_{njk}

$$q(\mathbf{f}_n^k) = \mathcal{N}(\mathbf{f}_n^k | \mu_{\mathbf{f}_n^k}, \Sigma_{\mathbf{f}_n^k})$$

Variational Inference

Transition matrix

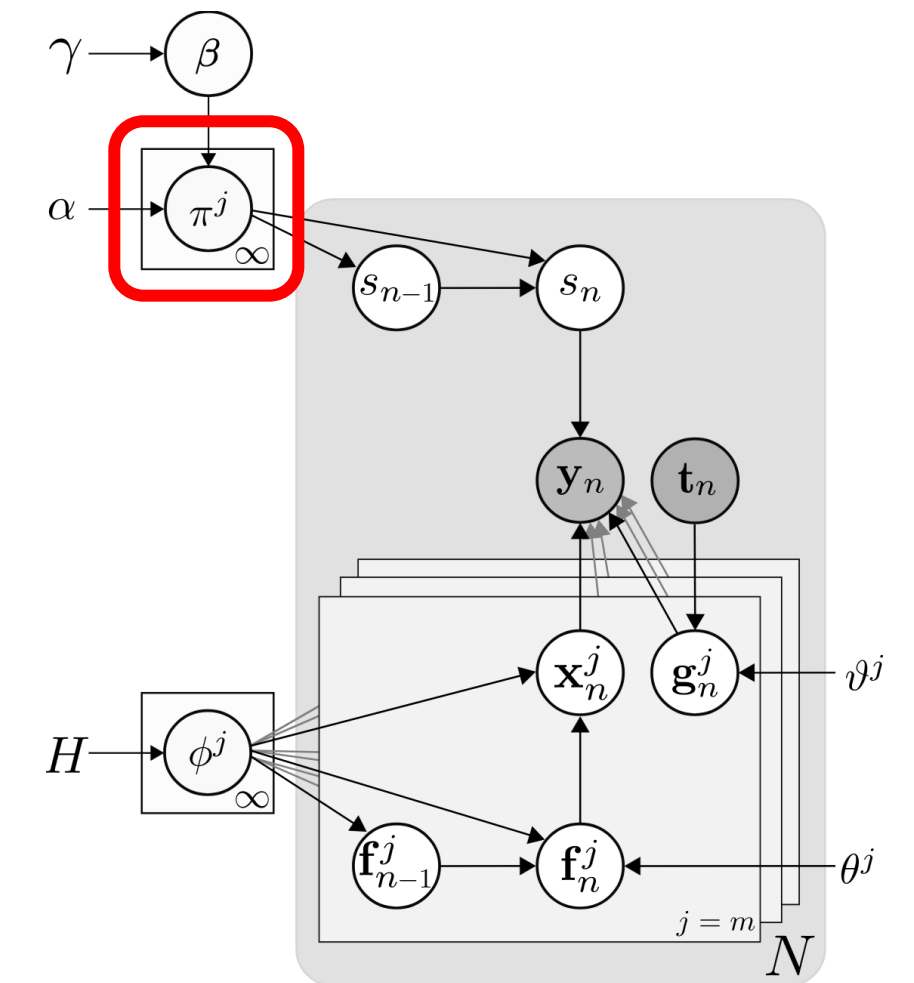
We recognize a product of Dirichlet distributions

$$q^*(\pi) = \prod_{j=1}^{K+1} \text{Dir}(\pi_j | \kappa_{j1}, \dots, \kappa_{jK+1})$$

with updated parameters

$$\kappa_{jk} = \alpha \mathbb{E}_v[\beta_k] + N_{jk}$$

$$N_{jk} = \sum_{n=1}^N \xi_{njk}$$



EXTRA

Variational Inference

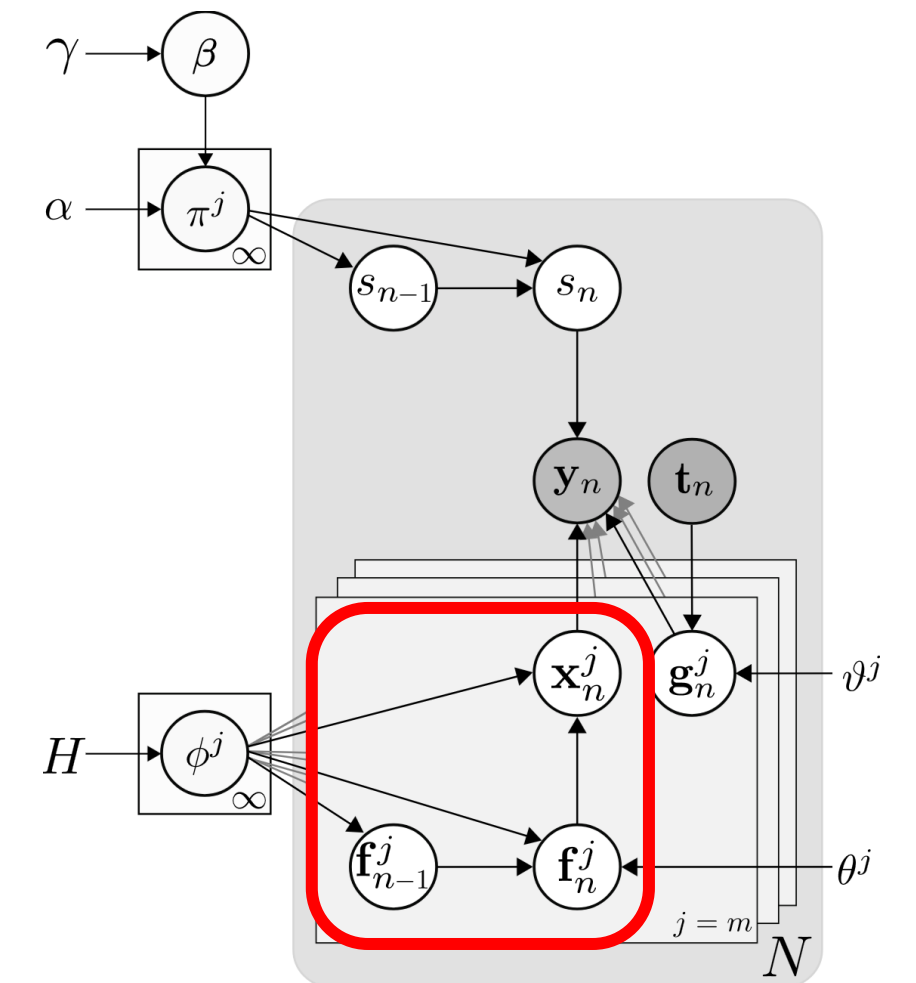
Latent and emission GPs

As the LDS transformations remain linear

$$f_n^{(s_n)}(t) = \mathcal{L}_A^{(s_n)} f_{n-1}^{(s_n)}(t) + v_n^{(s_n)}(t)$$

$$x_n^{(s_n)}(t) = \mathcal{L}_C^{(s_n)} f_n^{(s_n)}(t) + e_n^{(s_n)}(t)$$

the GP structure of the latent process and the emission is preserved.



Variational Inference

Latent and emission GPs

A message passing algorithm can be used, with forward step

$$q(\mathbf{f}_n^k) = \mathcal{N}(\mathbf{f}_n^k | \mu_{\mathbf{f}_n^k}, \Sigma_{\mathbf{f}_n^k})$$

where $\mu_{\mathbf{f}_n^k} \triangleq \mathbb{E}[\mathbf{f}_n^k] = \widetilde{M}_A^k \mu_{\mathbf{f}_{n-1}^k} + H_n^k (\mathbf{y}_n - \hat{K}_n^k \widetilde{M}_C^k \widetilde{M}_A^k \mu_{\mathbf{f}_{n-1}^k})$

$$\Sigma_{\mathbf{f}_n^k} \triangleq \text{cov}(\mathbf{f}_n^k) = (I - H_n^k \hat{K}_n^k \widetilde{M}_C^k) P_{n-1}^k,$$

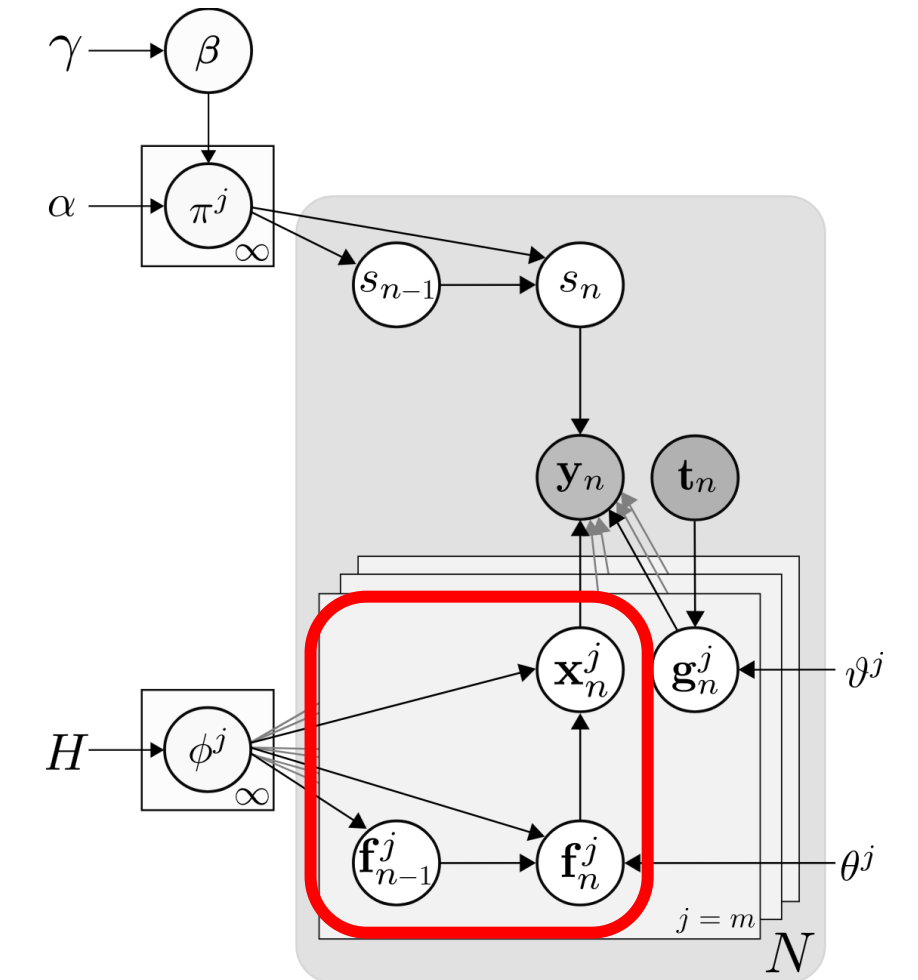
where H_n^k is the **Kalman gain matrix** weighted by responsibilities, and backwards step

$$q^*(\mathbf{f}_n^k) = \mathcal{N}(\mathbf{f}_n^k | \tilde{\mu}_{\mathbf{f}_n^k}, \tilde{\Sigma}_{\mathbf{f}_n^k}) \quad \text{where} \quad \tilde{\mu}_{\mathbf{f}_n^k} \triangleq \mu_{\mathbf{f}_n^k} + J_n^k (\tilde{\mu}_{\mathbf{f}_{n+1}^k} - \widetilde{M}_A^k \mu_{\mathbf{f}_n^k})$$

$$\tilde{\Sigma}_{\mathbf{f}_n^k} \triangleq \Sigma_{\mathbf{f}_n^k} + J_n^k (\tilde{\Sigma}_{\mathbf{f}_{n+1}^k} - P_n^k) J_n^{kT}$$

Finally, the emission GP is defined as

$$q^*(\mathbf{x}_n^k | \tilde{\mathbf{f}}_n^k) = \mathcal{N}(\mathbf{x}_n^k | \widetilde{M}_C^k \tilde{\mathbf{f}}_n^k, \tilde{S}_\epsilon^k)$$



EXTRA

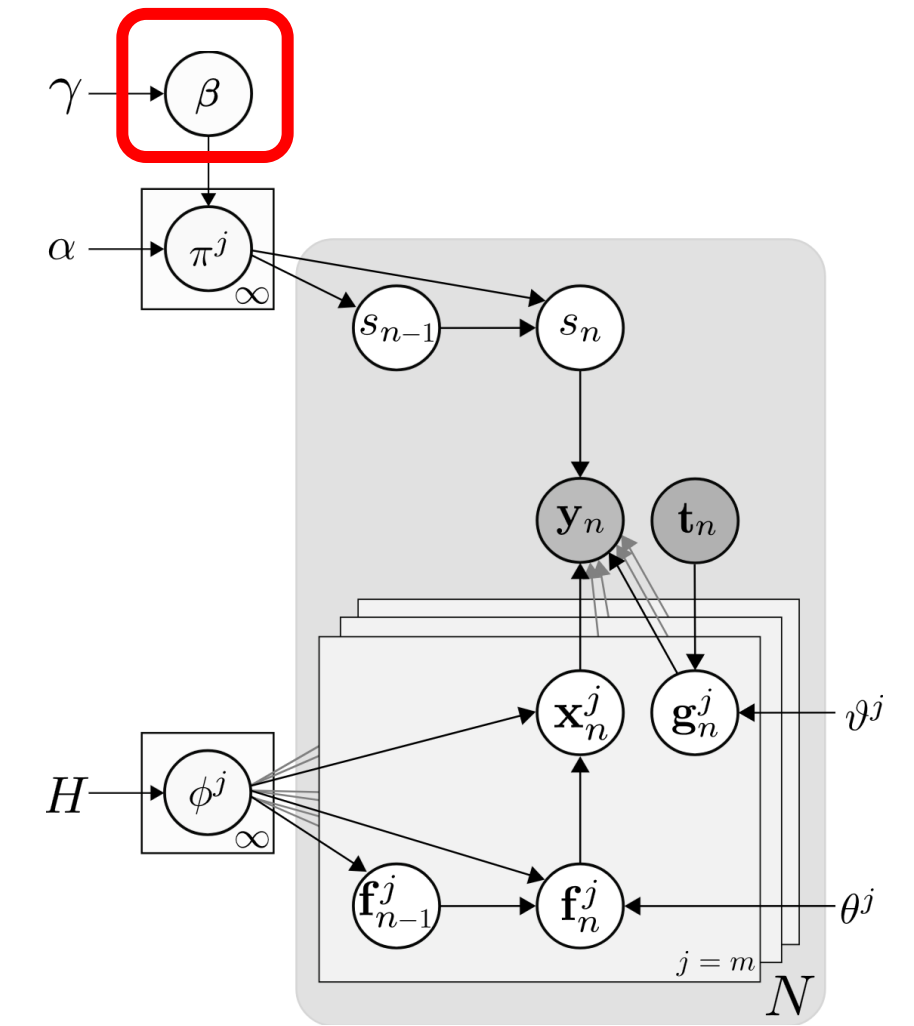
Variational Inference

Stick-breaking weights

The update to $q(\beta)$ is not conjugate given the other factors. The usual approach, a point estimate

$$q(\beta) = \delta_{\beta^*}(\beta)$$

Does not penalize empty clusters, then is suggested a **proper Beta distribution** (Hughes et al., 2015)



$$q^*(v|\lambda, \eta) = \prod_{k=1}^{\infty} \text{Beta}(v_k | \lambda_k \eta_k, (1 - \lambda_k) \eta_k)$$

then

$$\beta_k = v_k \prod_{j=1}^{k-1} (1 - v_j), \quad v_k \sim \text{Beta}(1, \gamma)$$

Variational Inference

GP alignment function

The log of the optimized factor is given by

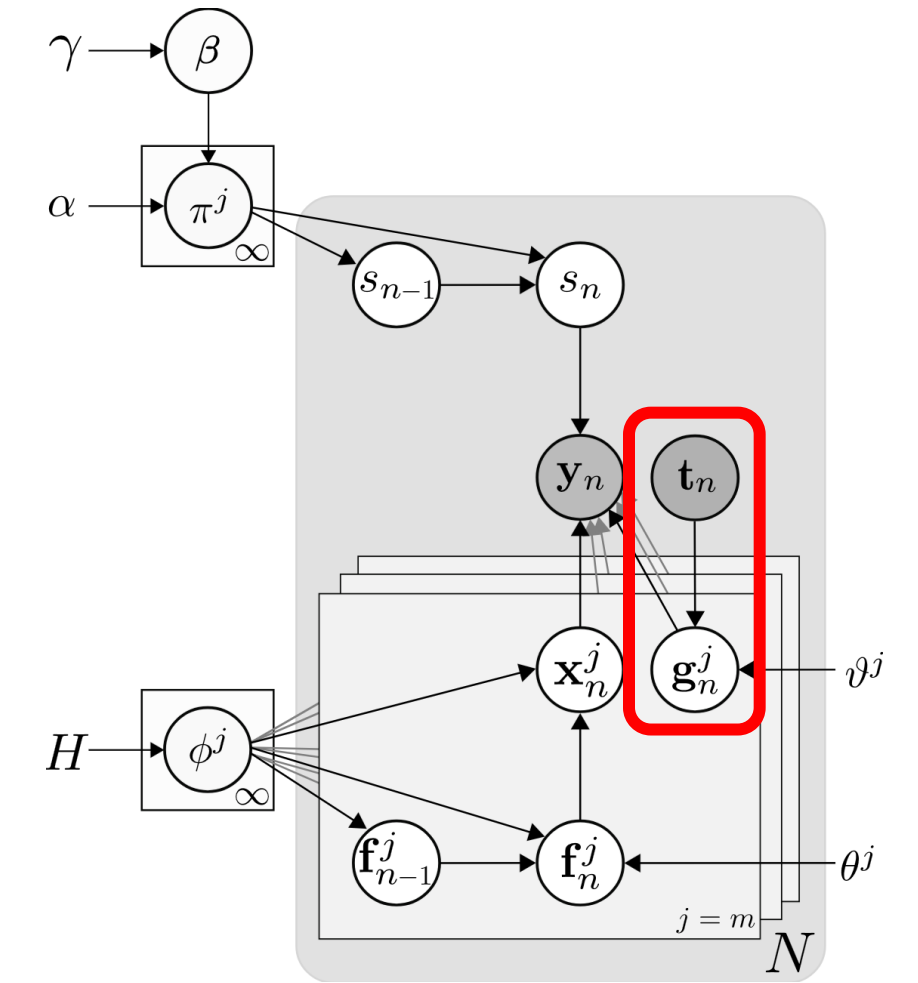
$$\log q^*(\mathcal{G}) = \log p(\mathcal{G}) + \mathbb{E}_{\mathcal{S}, \mathcal{X}} [\log p(\mathcal{Y} | \mathcal{S}, \mathcal{X}, \mathcal{G})] + \text{const.}$$

Which does not hold a closed-form solution, we then use a **maximum a posteriori** (MAP) approach, maximizing

$$\log q^*(\mathbf{g}_n^k | \mathbf{y}_n, \mathbf{t}_n) = \log p(\mathbf{y}_n | m_{x_n^k}(\mathbf{g}_n^k), K_{\mathbf{g}_n^k, \mathbf{g}_n^k}^{x_n^k}) + \log p(\mathbf{g}_n^k | \mathbf{t}_n) + \text{const.}$$

And to ensure monotonicity, we parametrize the warping function as (Kazlauskaite et al., 2019)

$$\mathbf{a}_n^k \in \mathbb{R}^Q \text{ such that } g_{n,q}^k = Q \sum_{i=1}^q [\text{softmax}(\mathbf{a}_n^k)]_i, \text{ with } \mathbf{a}_n^k \sim \mathcal{N}(\mathbf{0}, I),$$



EXTRA

Variational Inference

Dynamic parameters

The matrix-normal inverse Wishart prior

$$q(A^k, \Sigma_\omega^k) = q(A^k | \Sigma_\omega^k) q(\Sigma_\omega^k) = \mathcal{MN}(M_A^k, W_A^k, V_A^k) \mathcal{IW}(\nu_0, S_\omega^k)$$

provides conjugates to the likelihood equations, then the optimized factor is given by

$$q^*(A^k, \Sigma_\omega^k) = \mathcal{MN}(\Psi_{m,m-1} \Psi_{m-1,m-1}^{-1}, \Sigma_\omega^k, \Psi_{m-1,m-1}) \\ \mathcal{IW}(N_k + \nu_0, \Psi_{m|m-1} + S_\omega^k)$$

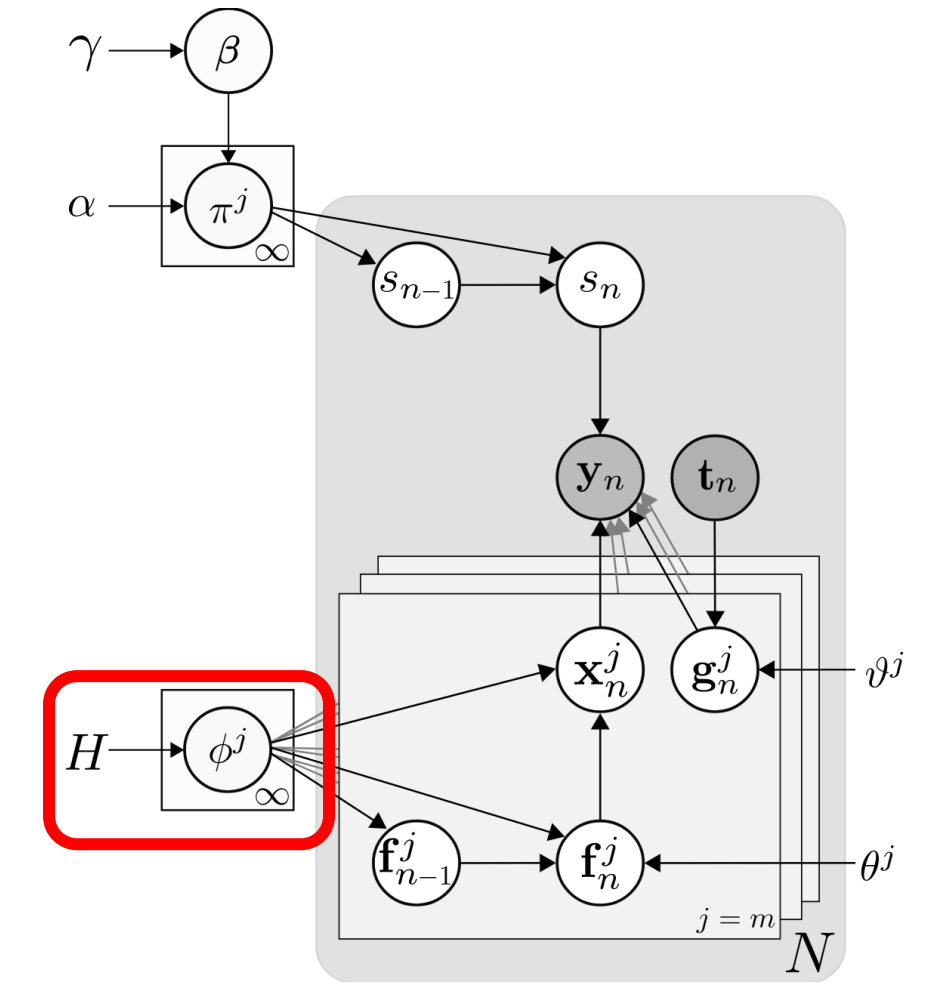
where

$$\Psi_{m-1,m-1} = \mathbb{E}[\mathbf{f}_{m-1}^k \mathbf{f}_{m-1}^{kT}] + V_A^k,$$

$$\Psi_{m,m-1} = \mathbb{E}[\mathbf{f}_m^k \mathbf{f}_{m-1}^{kT}] + M_A^k V_A^k,$$

$$\Psi_{m,m} = \mathbb{E}[\mathbf{f}_m^k \mathbf{f}_m^{kT}] + M_A^k V_A^k M_A^{kT},$$

$$\Psi_{m|m-1} = \Psi_{m,m} - \Psi_{m,m-1} \Psi_{m-1,m-1}^{-1} \Psi_{m,m-1}^T$$



EXTRA

Variational Inference

Variational objective function

Maximizing the objective function can be decomposed as

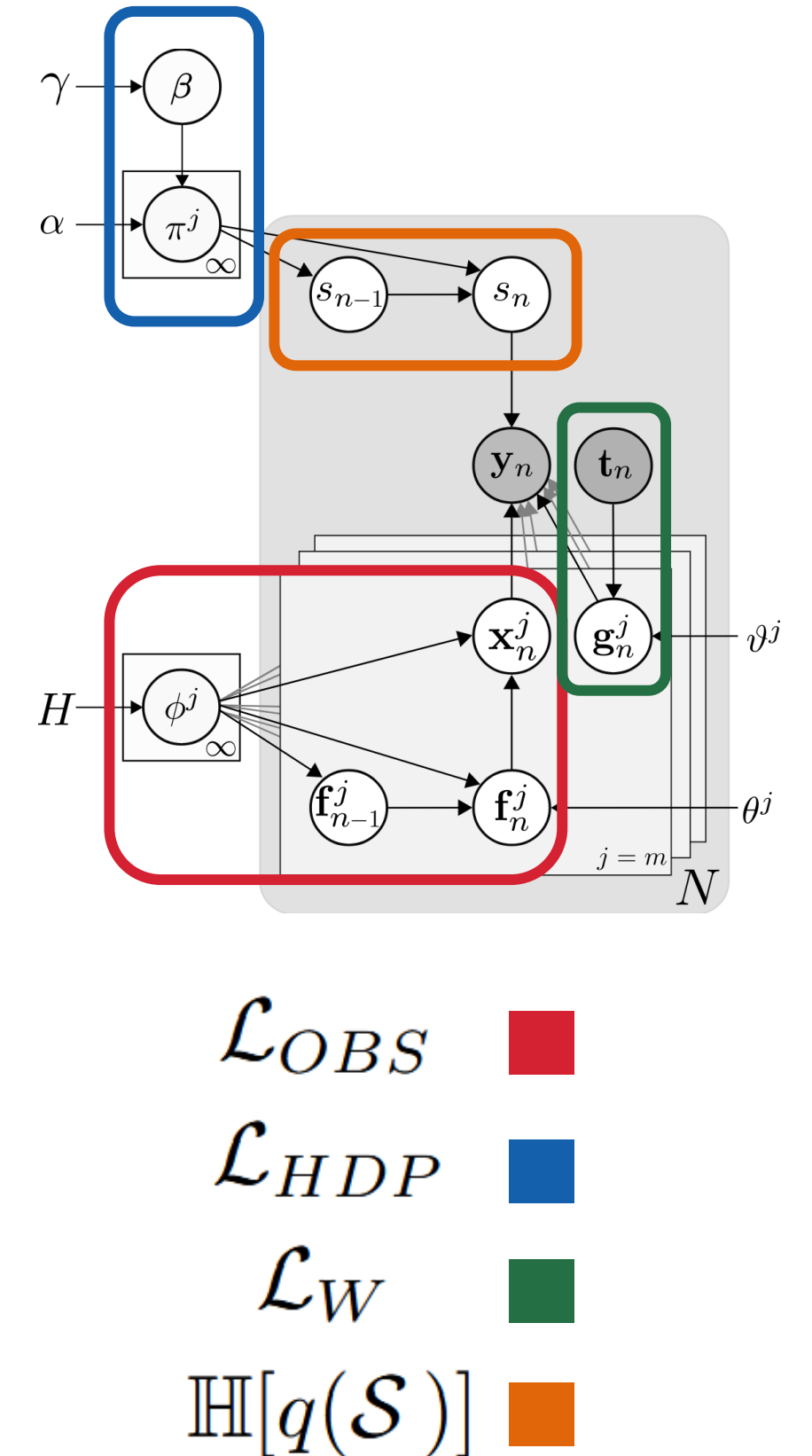
$$\begin{aligned}\mathcal{L}(q) &= \mathbb{E}_q [\log p(\mathcal{M}, \mathcal{Y}) - \log q(\mathcal{M})] \\ &= \mathcal{L}_{OBS} + \mathcal{L}_{HDP} + \mathcal{L}_W + \mathbb{H}[q(\mathcal{S})]\end{aligned}$$

where

$$\mathcal{L}_{OBS} = \mathbb{E}_q \left[\log p(\mathcal{Y} | \mathcal{X}, \mathcal{G}, \mathcal{S}) + \log \frac{p(\mathcal{X} | \mathcal{F}, \phi) p(\mathcal{F} | \phi)}{q(\mathcal{X}, \mathcal{F})} + \log \frac{p(\phi)}{q(\phi)} \right]$$

$$\mathcal{L}_{HDP} = \mathbb{E}_q \left[\log \frac{p(\mathcal{S} | \pi) p(\pi | v, \alpha)}{q(\pi)} + \log \frac{p(v)}{q(v)} \right]$$

and $\mathcal{L}_W$ is absorbed into the constant terms as **the MAP estimate** $\delta_{\mathcal{G}^*}(\mathcal{G})$ is always degenerate.



Variational Inference

Algorithm 2 Online variational inference

Input: A dataset $\mathcal{Y} = \{(\mathbf{t}_n, \mathbf{y}_n)\}_{n=1}^N$ of time sequences

Input: A model $p(\mathcal{M}, \mathcal{Y})$

Output: The variational density $q(\mathcal{M}|\mathcal{Y})$

```
1: procedure ONLINEVARIATIONALINFERENCE( $\mathcal{Y}, \mathcal{M}$ )
2:   for all  $(\mathbf{t}_n, \mathbf{y}_n) \in \mathcal{Y}$  do
3:     while  $\mathcal{L}$  has not converged do
4:       for all  $k = 1, 2, \dots$  do
5:         compute  $q(\mathbf{g}_n^k)$  numerically using eq. (3.22) //local update
6:         compute  $q(s_{nk})$  using eq. (3.14)
7:         for all  $m = 1, \dots, n_k$  do
8:           compute  $q(\mathbf{f}_m^k)$  and  $q(\mathbf{x}_m^k)$  using eq. (3.31)
9:         end for
10:        compute  $q(\pi_k)$  using eq. (3.19) //global update
11:        compute  $q(v_k)$  numerically using eq. (3.38)
12:        compute  $q(\phi^k)$  using eq. (3.26)
13:      end for
14:      compute  $\mathcal{L} = \mathcal{L}_{OBS} + \mathcal{L}_{HDP} + \mathcal{L}_W + \mathbb{H}[q(S)]$ 
15:    end while
16:  end for
17:  return  $q(\mathcal{M}|\mathcal{Y})$ 
18: end procedure
```

Variational Inference

Multi-output

To deal with multi-output signals $\mathbf{y}_n^{(l)}$ with $l = \{1, \dots, L\}$ a MOGP could be proposed, but the LDS update entangles output channels through latent function requiring block-structured LDS parameters.

$$\begin{bmatrix} f_n^1 \\ f_n^2 \\ \vdots \\ f_n^l \end{bmatrix} \sim \mathcal{GP} \left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \underbrace{\begin{bmatrix} K_n^{1,1} & K_n^{1,2} & \dots & K_n^{1,l} \\ K_n^{2,1} & K_n^{2,2} & \dots & K_n^{2,l} \\ \vdots & \vdots & \ddots & \vdots \\ K_n^{l,1} & K_n^{l,2} & \dots & K_n^{l,l} \end{bmatrix}}_{\mathcal{K}_n} \right) \rightarrow \begin{bmatrix} C_j^{1,1} & C_j^{1,2} & \dots & C_j^{1,l} \\ A_j^{1,1} & A_j^{1,2} & \dots & A_j^{1,l} \\ A_j^{2,1} & A_j^{2,2} & \dots & A_j^{2,l} \\ \vdots & \vdots & \ddots & \vdots \\ A_j^{l,1} & A_j^{l,2} & \dots & A_j^{l,l} \end{bmatrix} \begin{bmatrix} C_j^{2,l} \\ \vdots \\ C_j^{l,l} \end{bmatrix}$$

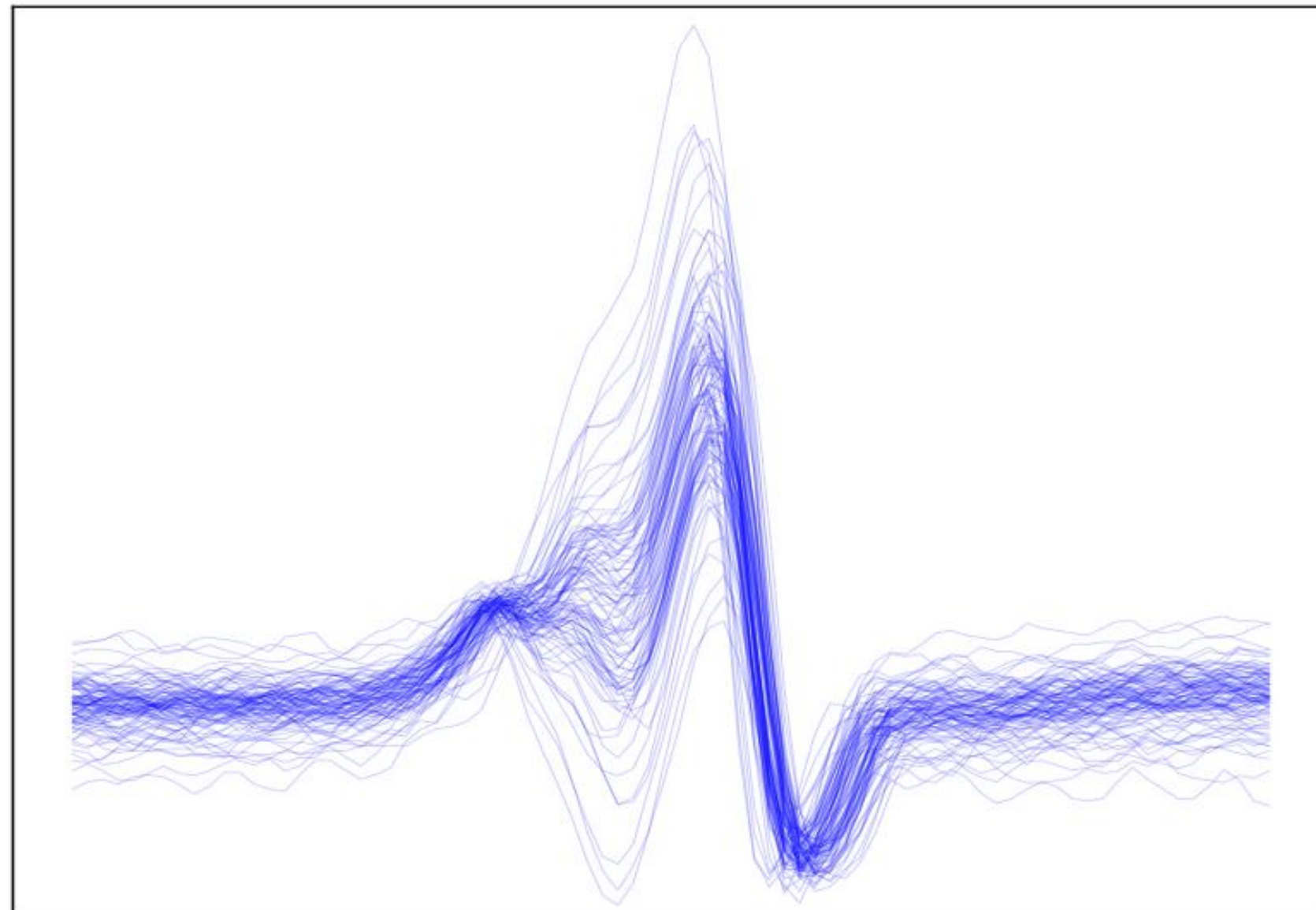
We assume a pragmatic simple approach: conditional independence across outputs given the discrete regime, defining an observation contribution as a geometric mean.

$$p(\mathbf{y}_n | s_n, \mathcal{X}, \mathcal{G}) \propto \prod_{l=1}^L \left[p(\mathbf{y}_n^{(l)} | s_n, \mathcal{X}^{(l)}, \mathcal{G}^{(l)}) \right]^{w_n^{(l)}} \quad \log \zeta_{njk} = \mathbb{E}[\log \pi_{jk}] + \sum_{l=1}^L w_n^{(l)} \log \zeta_{njk}^{(l)}$$

Where $w_n^{(l)}$ is chosen depending on the application.

Stability-Plasticity dilemma

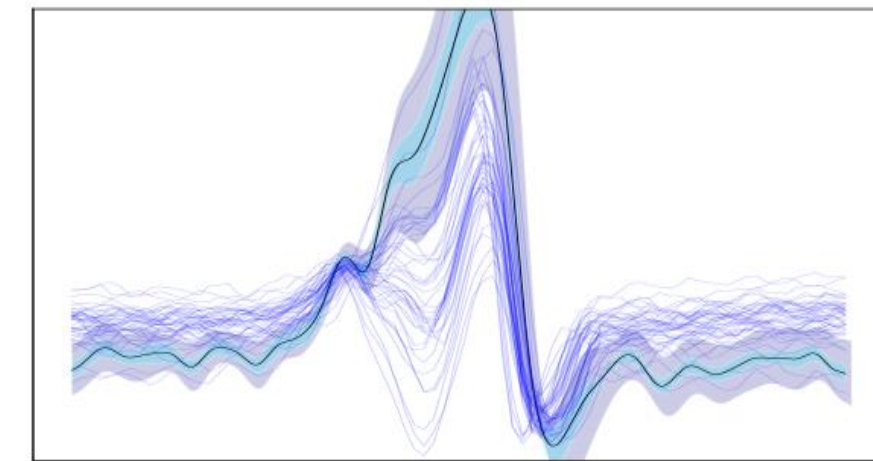
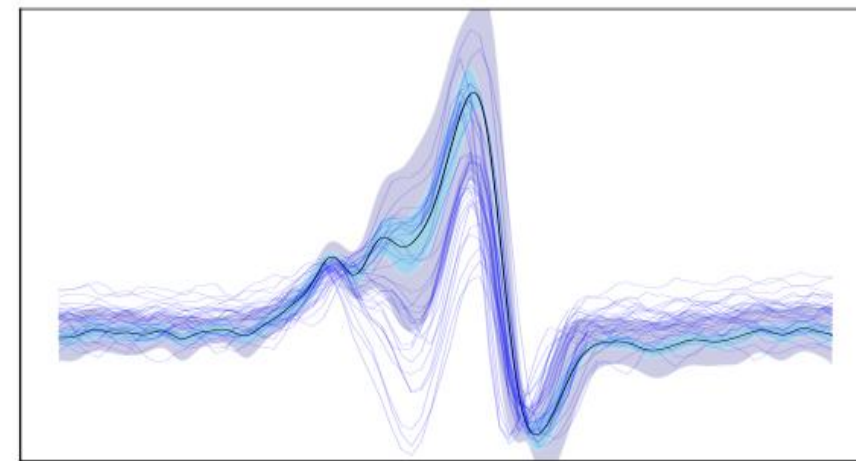
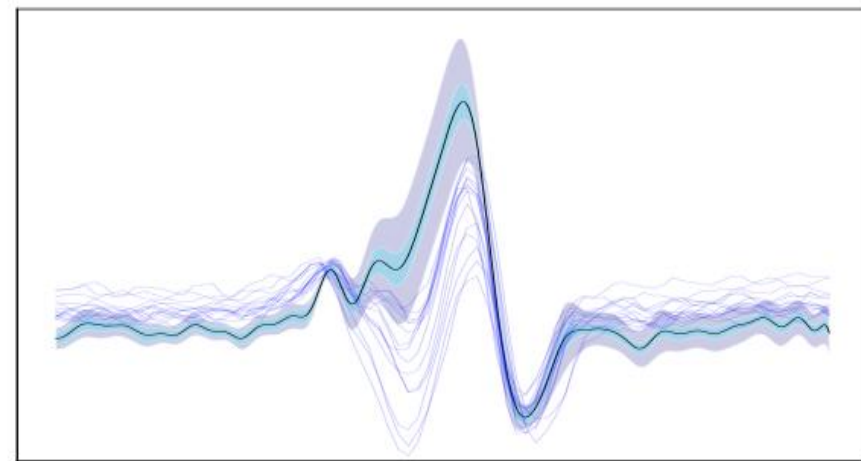
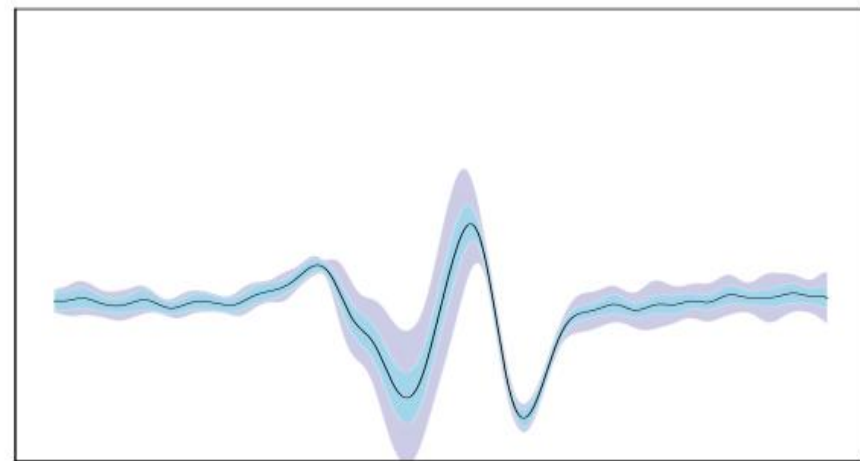
Where should variability go?



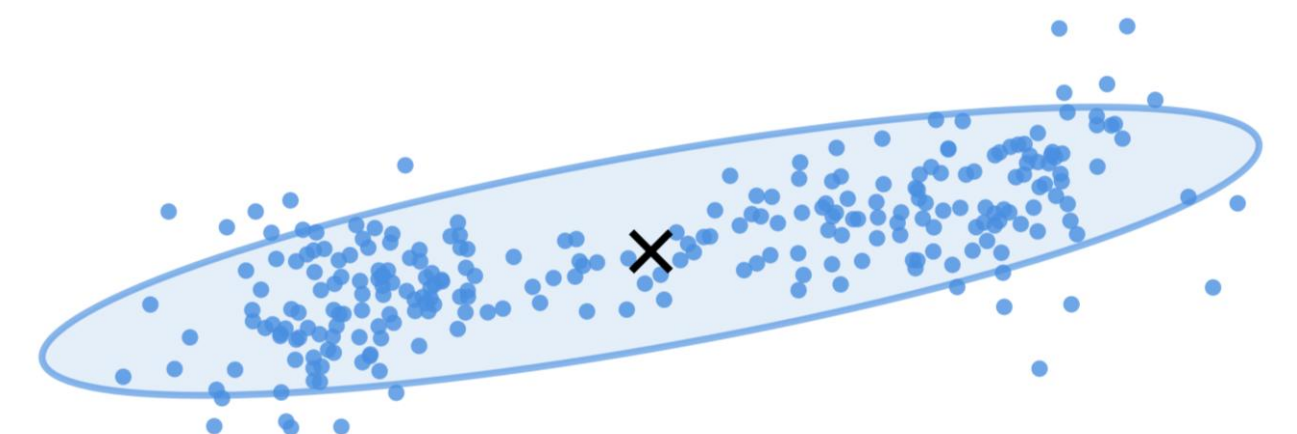
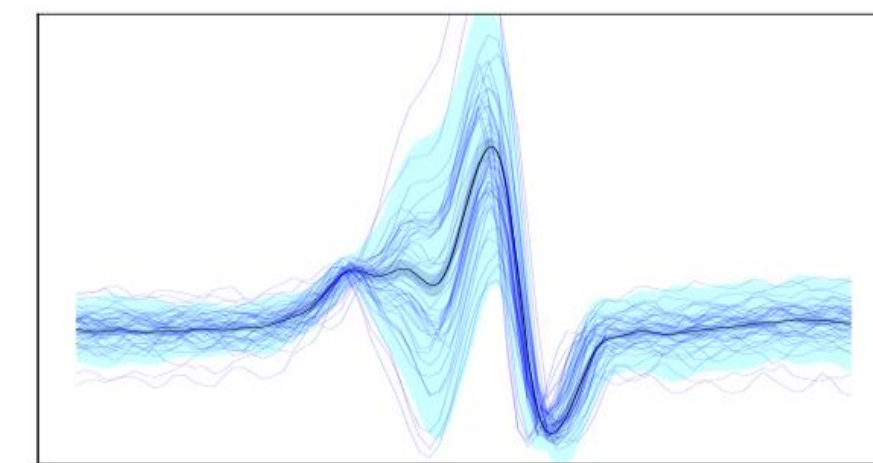
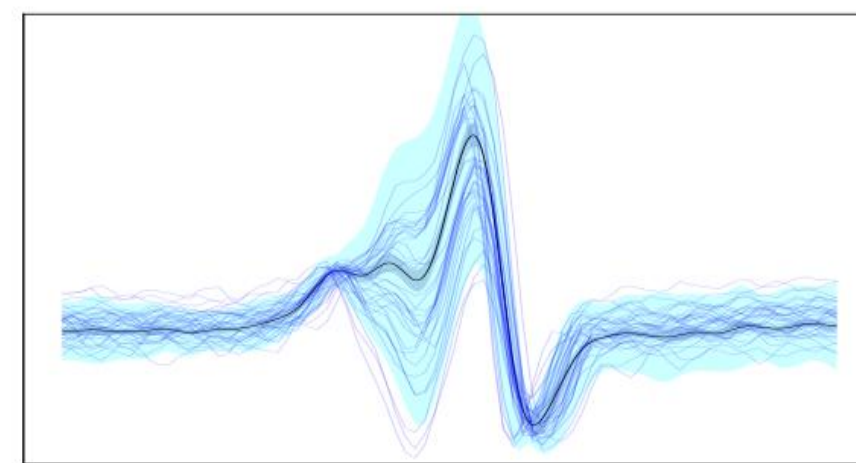
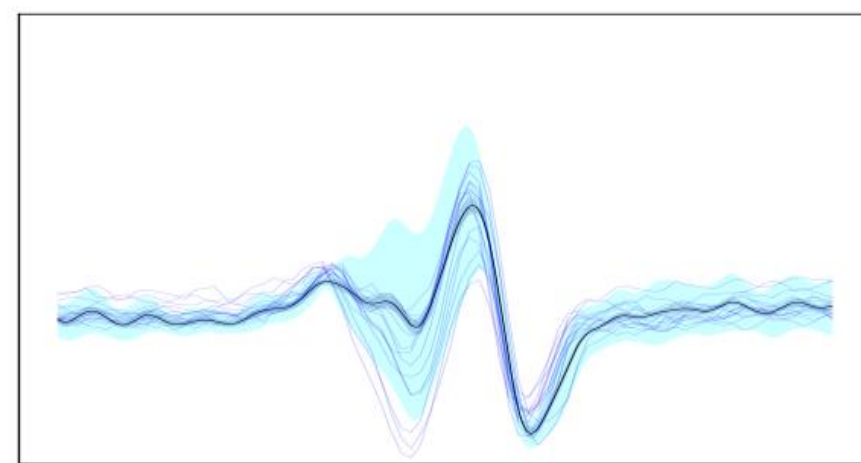
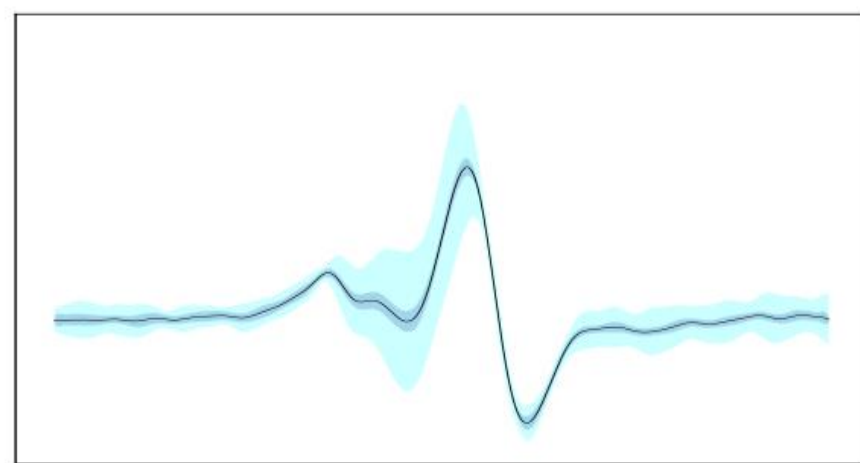
Stability-Plasticity dilemma

Hyperparameters of the model condition the behaviour of the clustering.

Plasticity $\Sigma_{\omega} > \Sigma_{\epsilon}$



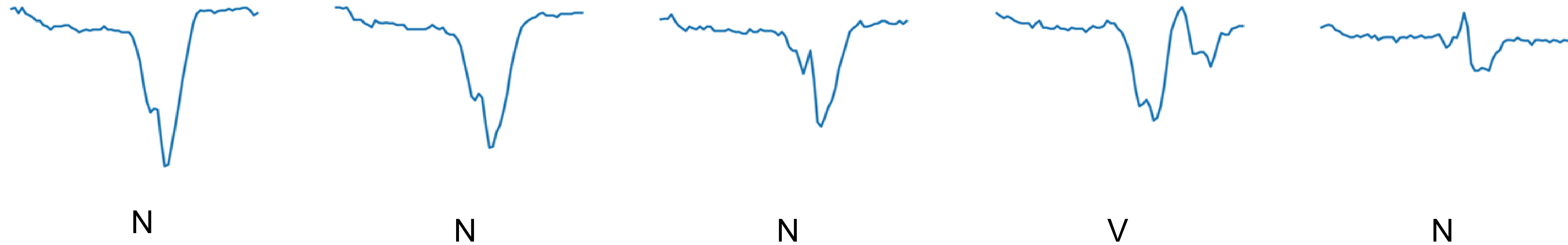
Stability $\Sigma_{\epsilon} > \Sigma_{\omega}$



Stability-Plasticity dilemma

Motivation example of dynamic clustering

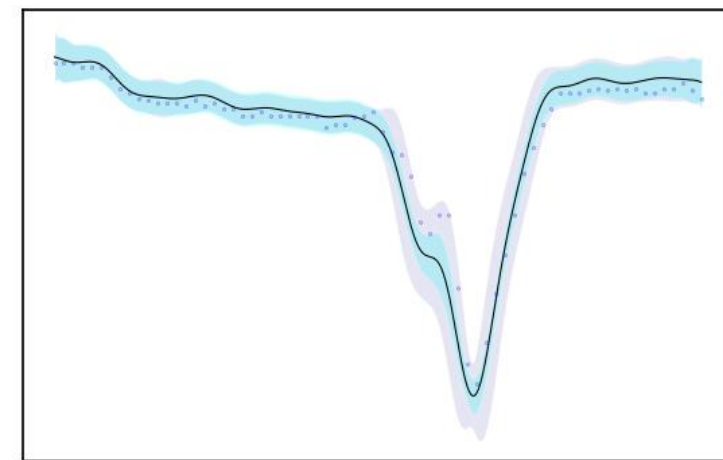
27



Stability-Plasticity dilemma

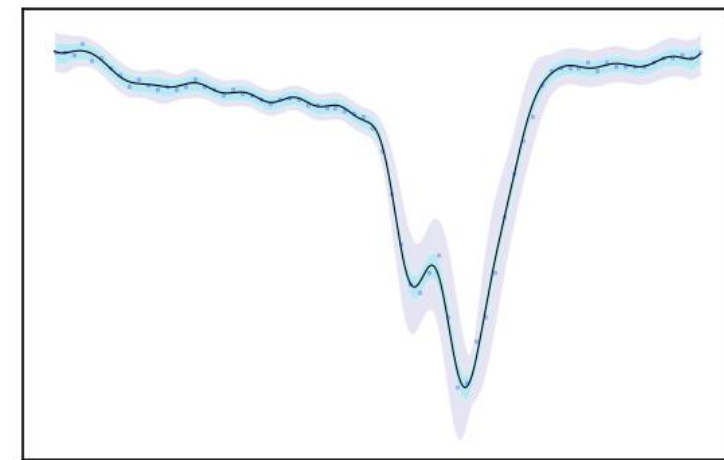
Our model captures correctly the nature of evolving morphologies

27



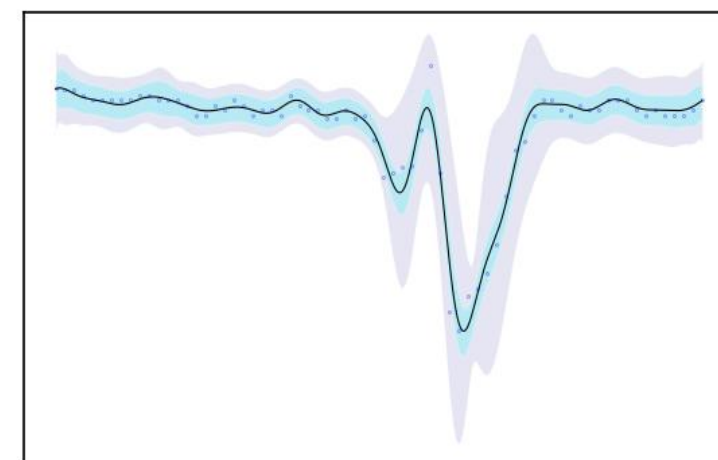
Model on $t = 2$

N



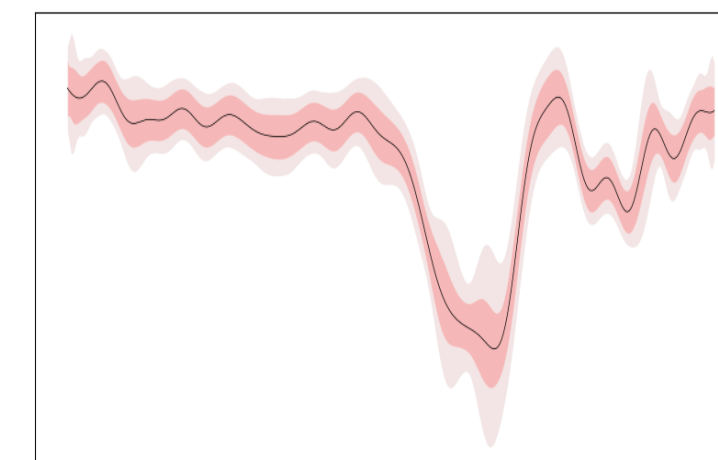
Model on $t = 100$

N



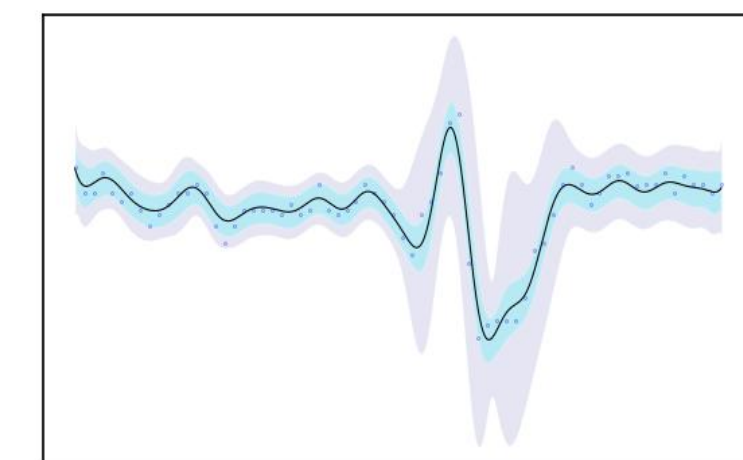
Model on $t = 2000$

N



Ventricular beat morphology

V



Model on $t = 5035$

N

Applications

Applications: myocardial conduction analysis

Applications: myocardial conduction analysis

MIT-BIH Arrhythmia Database



48 half-hour two-channel ambulatory
ECG records

~110000 beat-level annotations

Complex combinations of arrhythmias

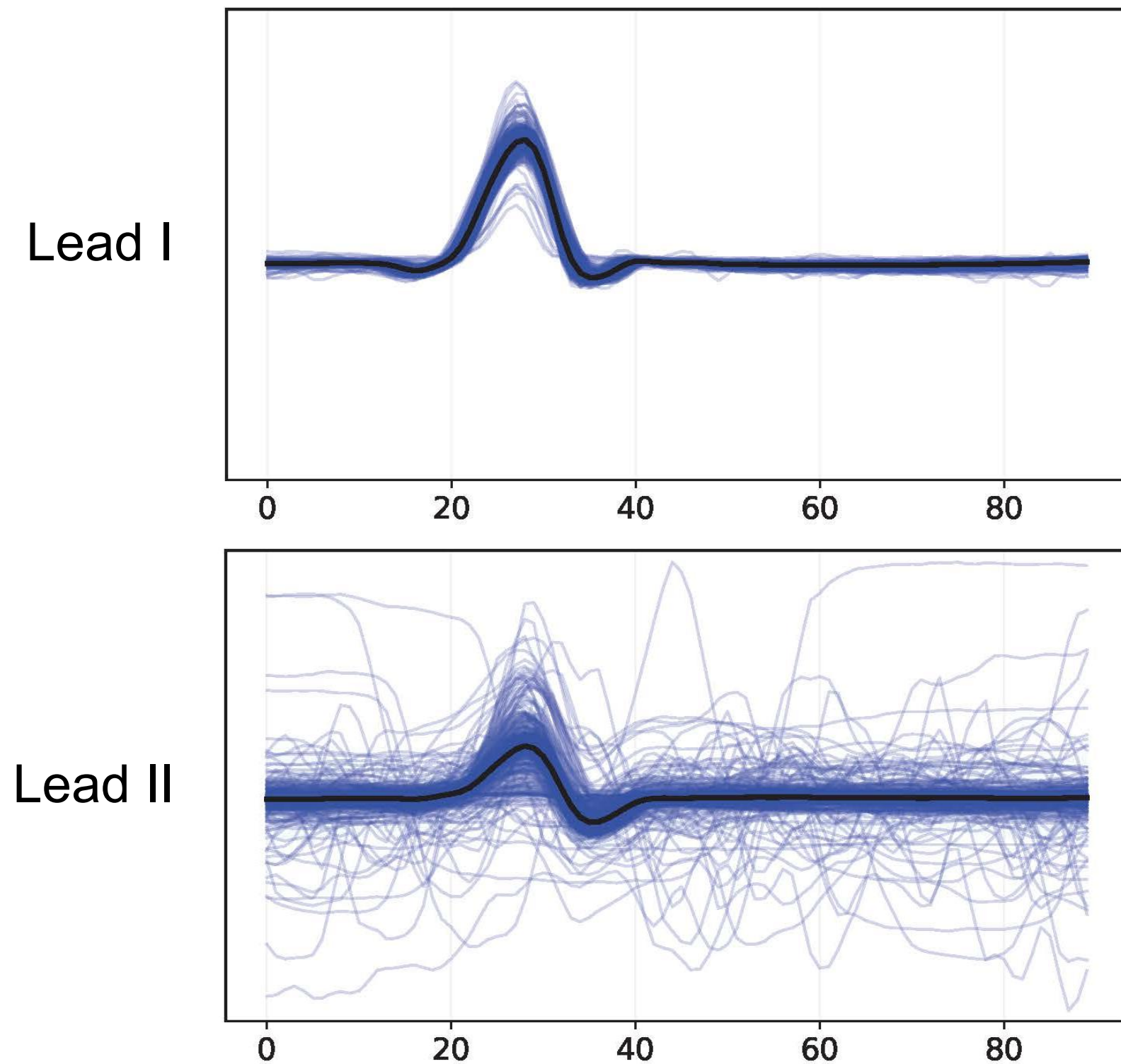
Morphological variations

Noise levels

Symbol	Description	Count	Proportion
N	Normal beat	75052	68.250%
L	Left bundle branch block beat	8075	7.343%
R	Right bundle branch block beat	7259	6.601%
A	Atrial premature beat	2546	2.315%
a	Aberrated atrial premature beat	150	0.136%
J	Nodal (junctional) premature beat	83	0.075%
S	Supraventricular premature or ectopic beat	2	0.002%
e	Atrial escape beat	16	0.015%
j	Nodal (junctional) escape beat	229	0.208%
n	Supraventricular escape beat (atrial or nodal)	0	0.000%
V	Premature ventricular contraction	7130	6.484%
E	Ventricular escape beat	106	0.096%
F	Fusion of ventricular and normal beat	803	0.730%
/	Paced beat	7028	6.391%
f	Fusion of paced and normal beat	982	0.893%
Q	Unclassifiable beat	33	0.030%
!	Ventricular flutter wave	472	0.429%

Applications: myocardial conduction analysis

Multi-output SNR



$w_n^{(1)}$ ↑

$w_n^{(2)}$ ↓

Signal-to-noise ratio (SNR) using the GP prototype

$$p(\mathbf{y}_n | s_n, \mathcal{X}, \mathcal{G}) \propto \prod_{l=1}^L \left[p(\mathbf{y}_n^{(l)} | s_n, \mathcal{X}^{(l)}, \mathcal{G}^{(l)}) \right]^{w_n^{(l)}}$$

SNR
softmax

$$w_n^{(l)} = \frac{\exp(\varphi_n^{(l)})}{\sum_{l'=1}^L \exp(\varphi_n^{(l')})}$$

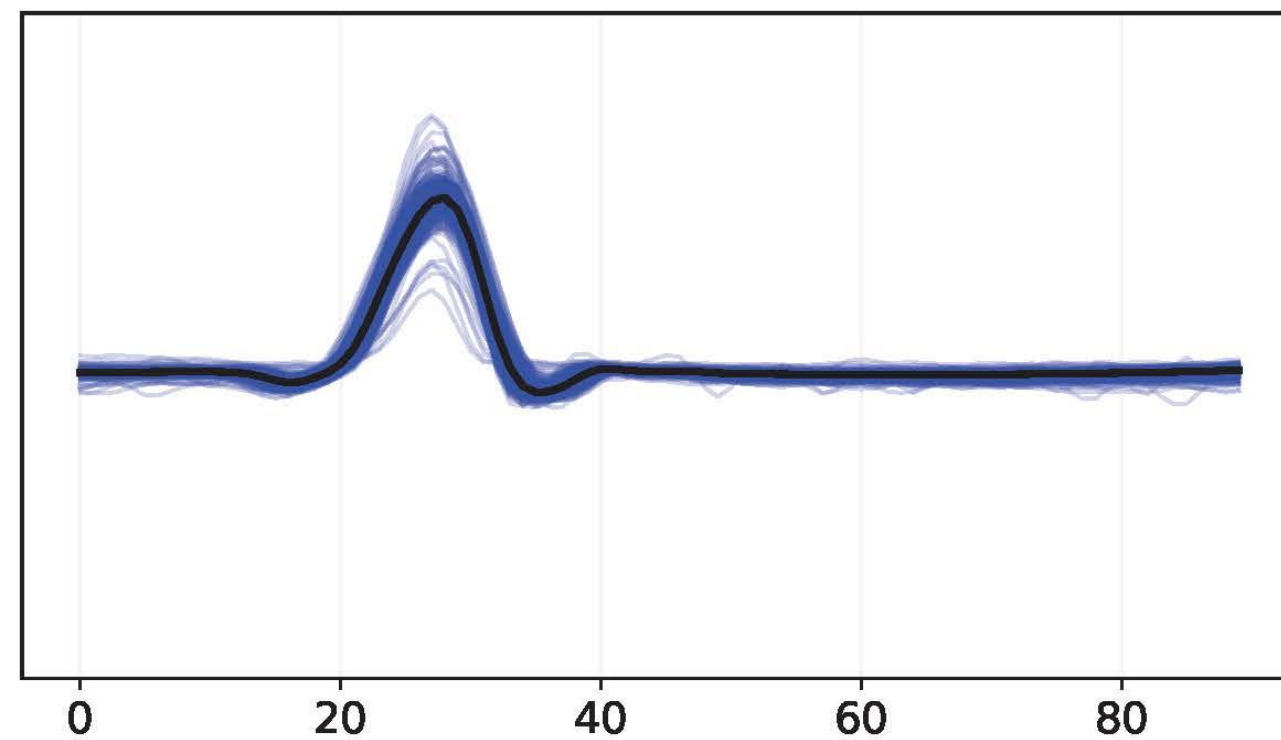
SNR

$$\varphi_n^{(l)} = 10 \log_{10} \left(\frac{\|\mathbf{y}_n^{(l)}\|_2^2}{\|\mathbf{y}_n^{(l)} - \mathbf{x}_n^{(l)}\|_2^2 + \varepsilon} \right)$$

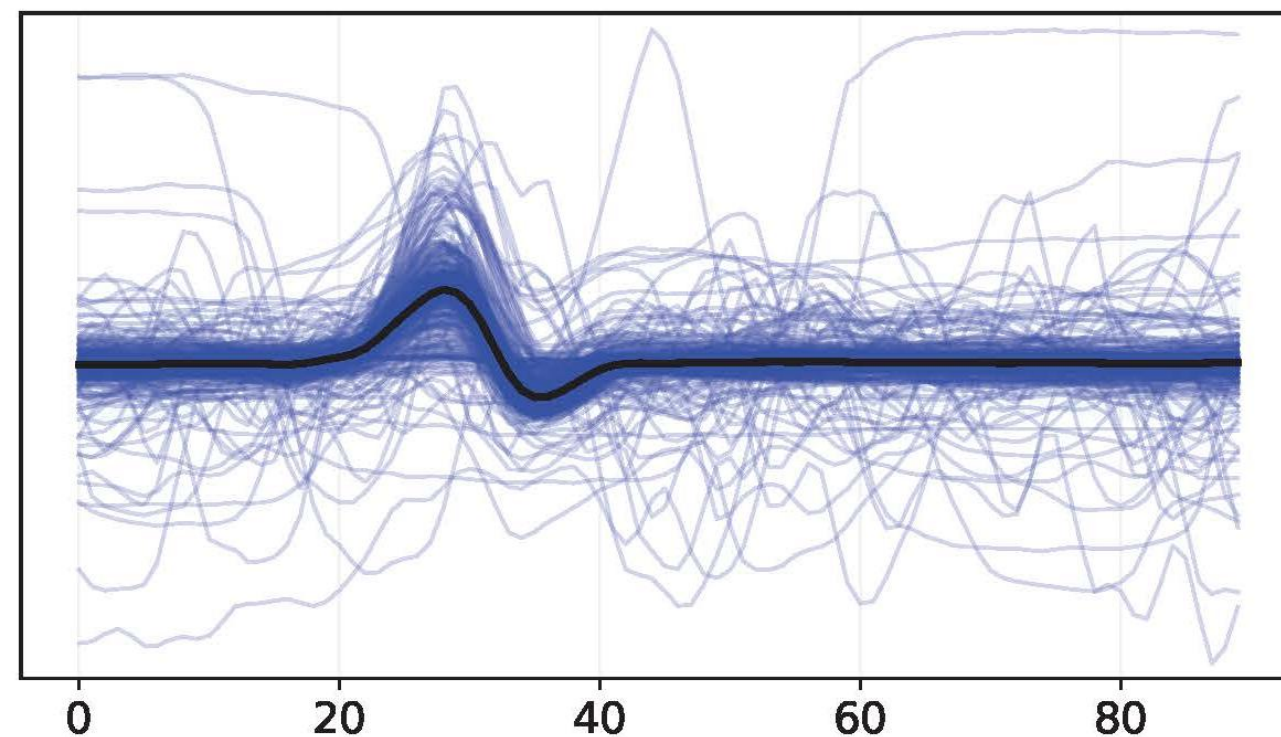
Applications: myocardial conduction analysis

Multi-output SNR

Lead I

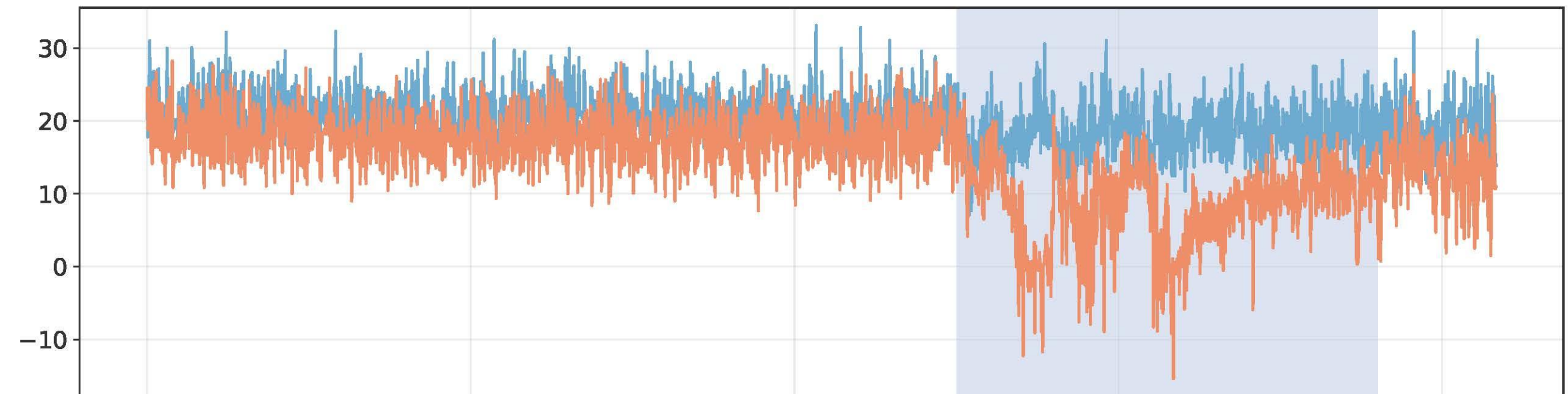


Lead II



SNR

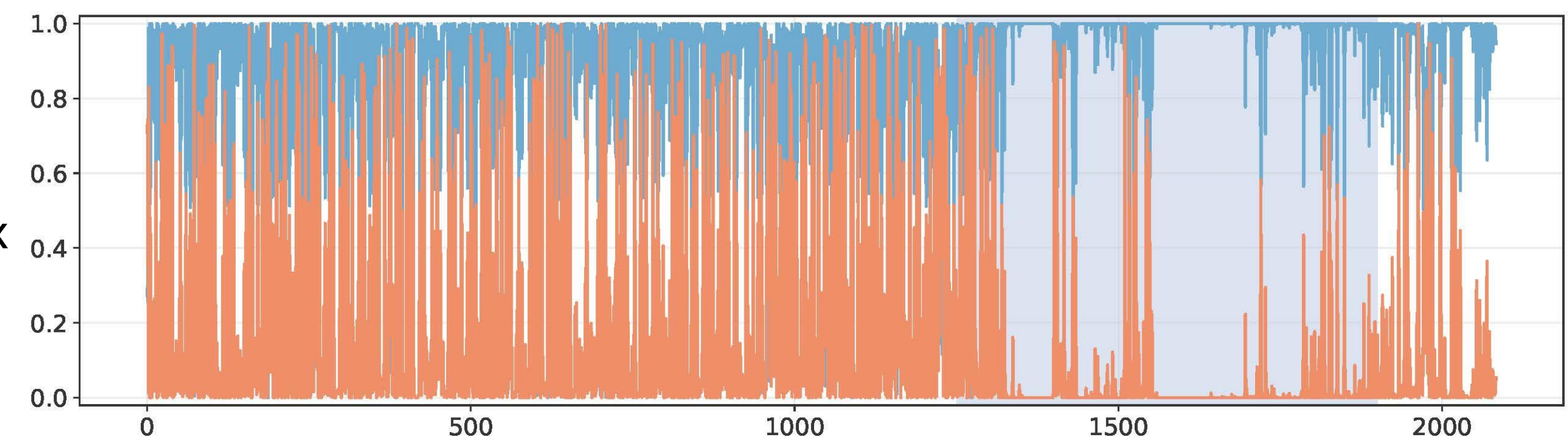
$\varphi_n^{(l)}$



SNR

softmax

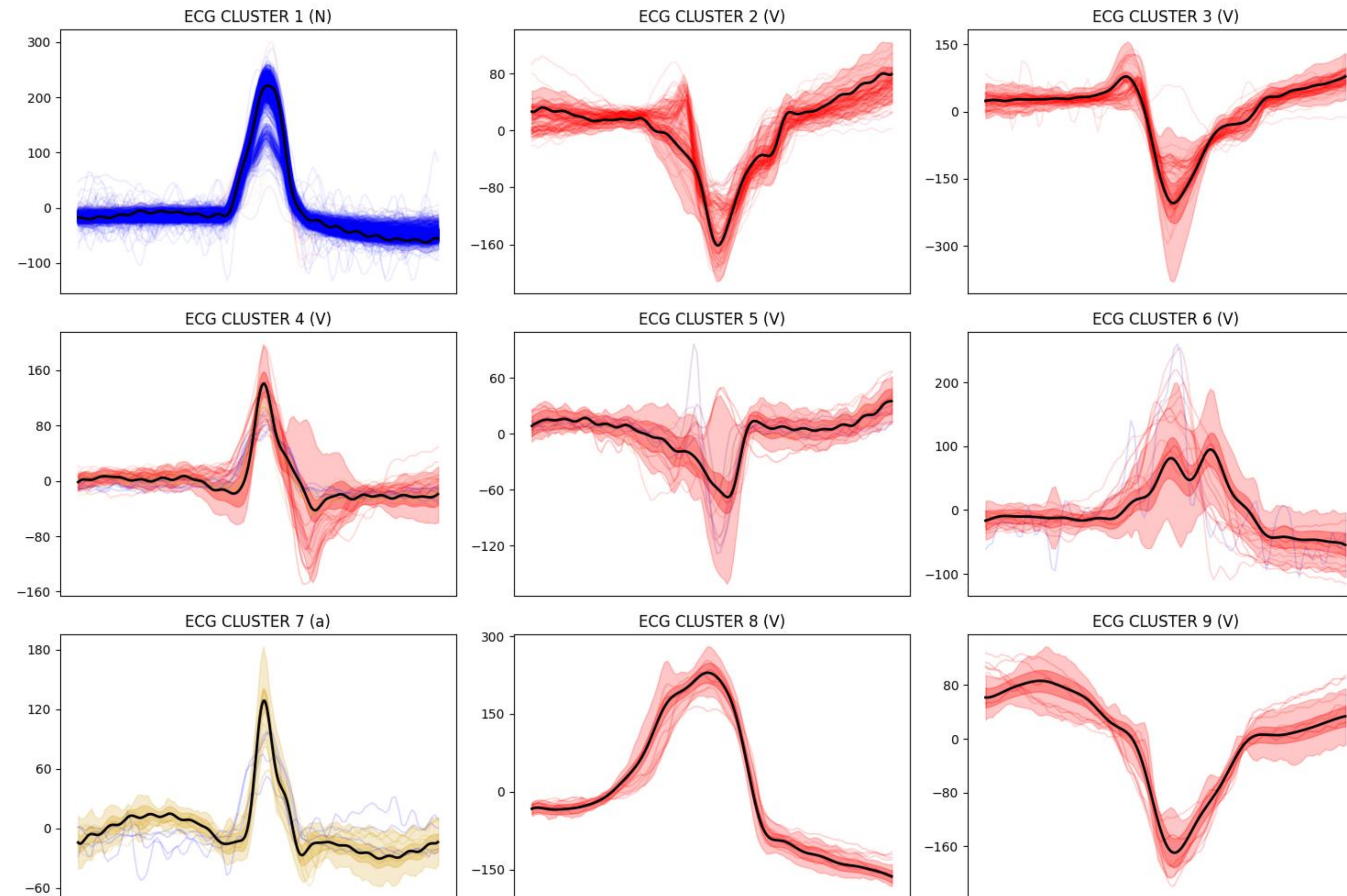
$w_n^{(l)}$



Lead I $w_n^{(1)}$
Lead II $w_n^{(2)}$

Applications: myocardial conduction analysis

Illustrative results



Applications: myocardial conduction analysis

Results and comparison

Method	Clusters	Acc./Purity	ARI	AMI	NMI	Homo.	Comp.
Lagerholm et al. [22]	1200	0.985	–	–	–	–	–
Castro et al. [23]	629	0.972	–	–	–	–	–
DTW+k-medoids [24], [25]	119	0.943 ± 0.063	0.224 ± 0.340	0.221 ± 0.306	0.222 ± 0.305	0.288 ± 0.322	0.232 ± 0.330
DTW+hierarchical-average [24], [26]	107	0.934 ± 0.087	0.367 ± 0.399	0.340 ± 0.372	0.341 ± 0.372	0.355 ± 0.373	0.506 ± 0.413
CAE+DBSCAN [27], [28]	168	0.946 ± 0.064	0.373 ± 0.308	0.323 ± 0.270	0.325 ± 0.269	0.490 ± 0.302	0.287 ± 0.243
CAE+DPGMM [27], [29]	906	0.979 ± 0.041	0.271 ± 0.326	0.298 ± 0.281	0.301 ± 0.281	0.754 ± 0.278	0.225 ± 0.237
DEC [30]	128	0.958 ± 0.059	0.375 ± 0.421	0.352 ± 0.383	0.353 ± 0.382	0.461 ± 0.371	0.364 ± 0.409
HDP-GPC MO offline	342	0.976 ± 0.045	0.434 ± 0.356	0.435 ± 0.312	0.437 ± 0.312	0.665 ± 0.350	0.414 ± 0.319

Applications: myocardial conduction analysis

Discussion

Method	Clusters	Acc./Purity	ARI	AMI	NMI	Homo.	Comp.
HDP-GPC MO offline	342	0.976 ± 0.045	0.434 ± 0.356	0.435 ± 0.312	0.437 ± 0.312	0.665 ± 0.350	0.414 ± 0.319

Interpretability emerges from the model structure

Cluster assignments

GP prototypes

Uncertainty bands

LDS dynamics

Warping functions

Applications: myocardial conduction analysis

Future work

Method	Clusters	Acc./Purity	ARI	AMI	NMI	Homo.	Comp.
HDP-GPC MO offline	342	0.976 ± 0.045	0.434 ± 0.356	0.435 ± 0.312	0.437 ± 0.312	0.665 ± 0.350	0.414 ± 0.319

Where does variability go?

Cluster assignments

GP prototypes

Uncertainty bands

LDS dynamics

Warping functions

Applications: myocardial conduction analysis

Future work

Fully variational warp computation.

Truly multi-output modelling (tested in supervised degree's thesis).

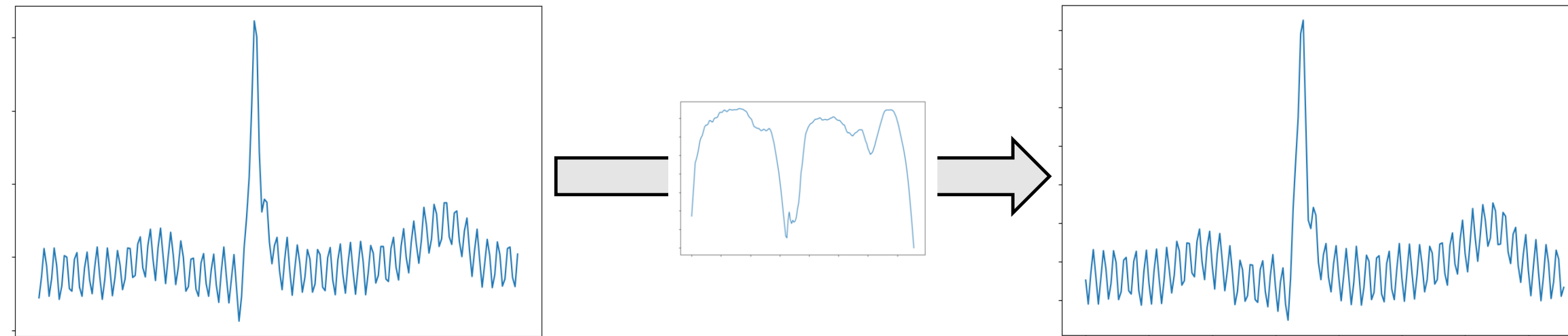
Explore different kernels to deal with artifacts as pacemakers.

Applications: breath estimation

Applications: breath estimation

Interpretability

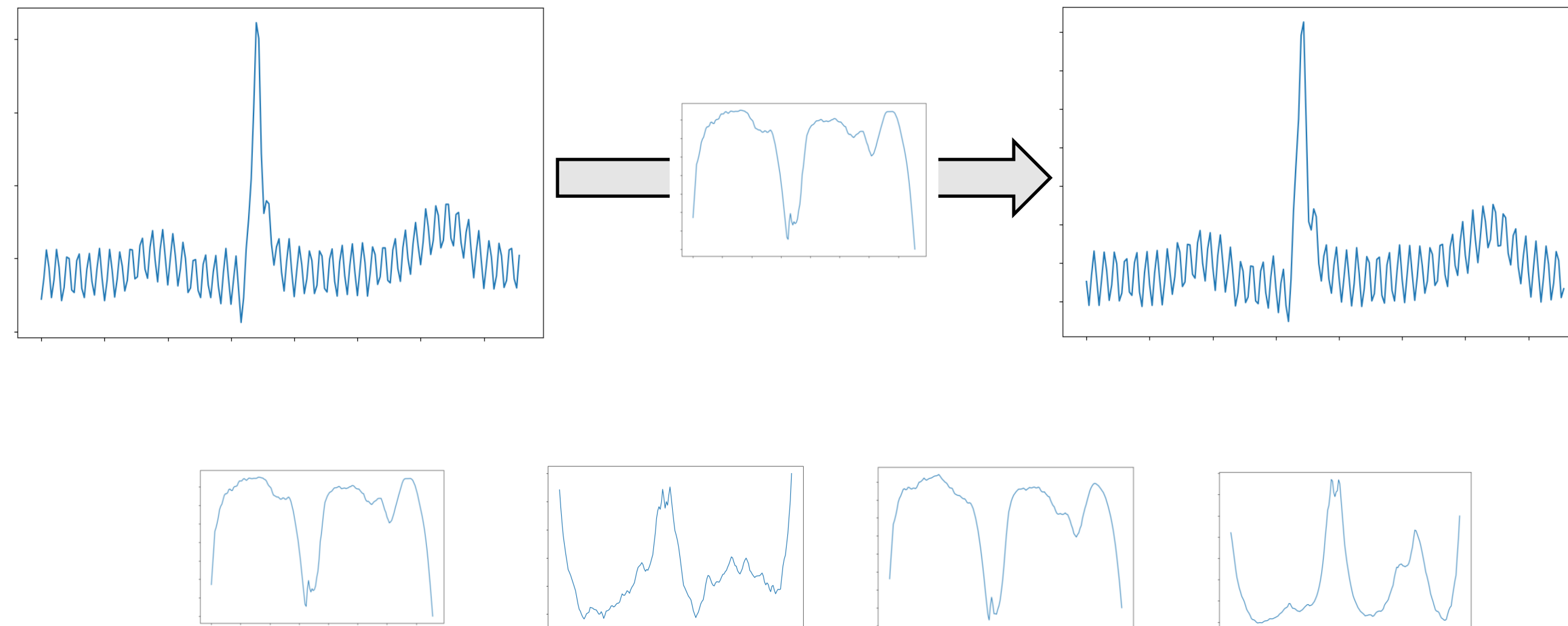
Generative modelling enables new questions to be asked about how the observations have been generated.



Applications: breath estimation

Interpretability

Generative modelling enables new questions to be asked about how the observations have been generated.

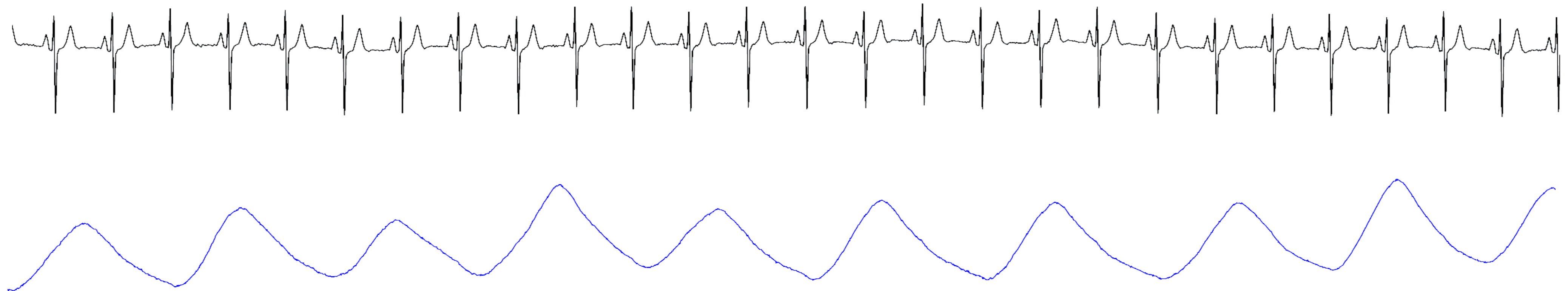


Applications: breath estimation

Interpretability

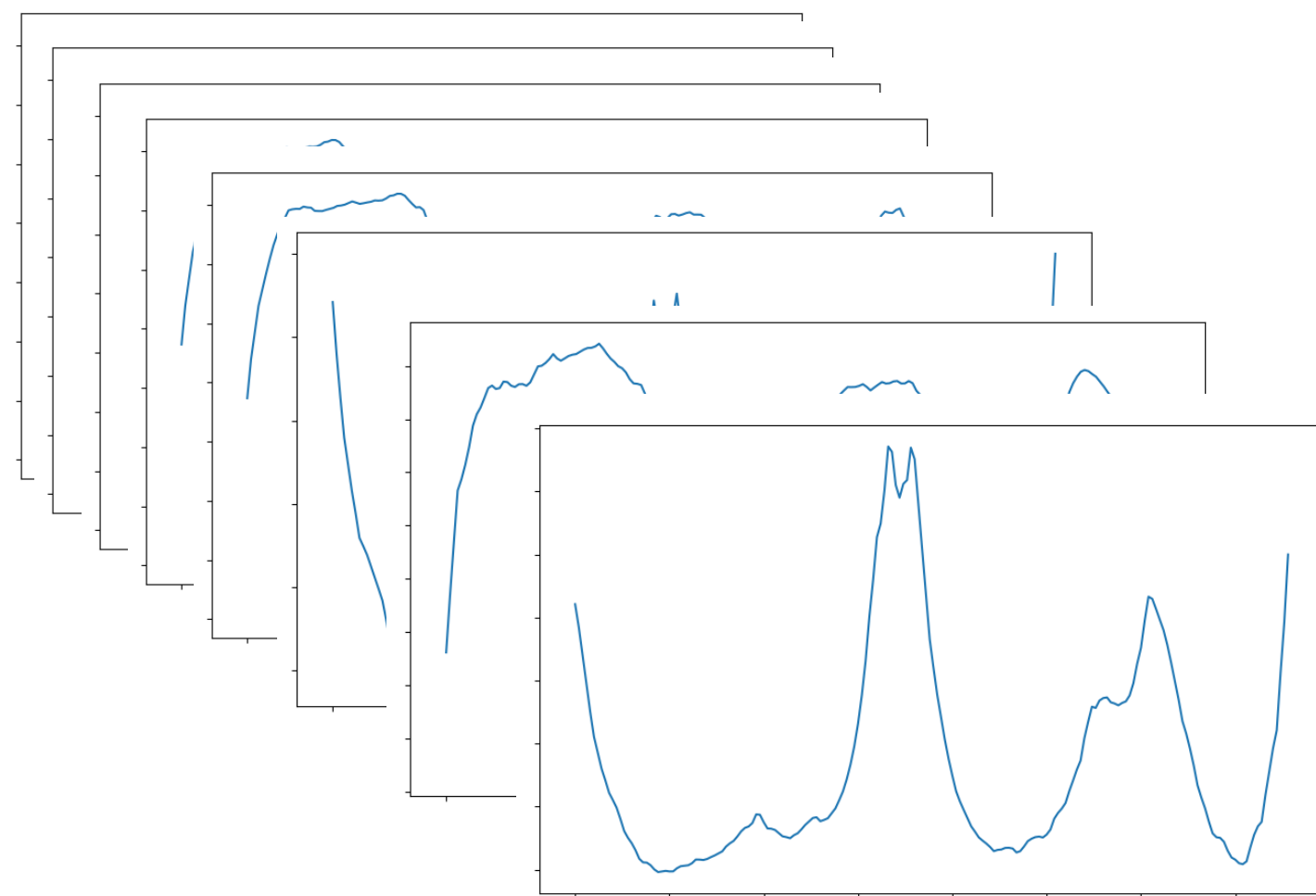
We hypothesized that inspiration-expiration cycle causes a delay in the ECG signal due to the dielectric medium (air), added to the propagation medium, in the lungs.

Can we estimate the breathing signal from the ECG signal?



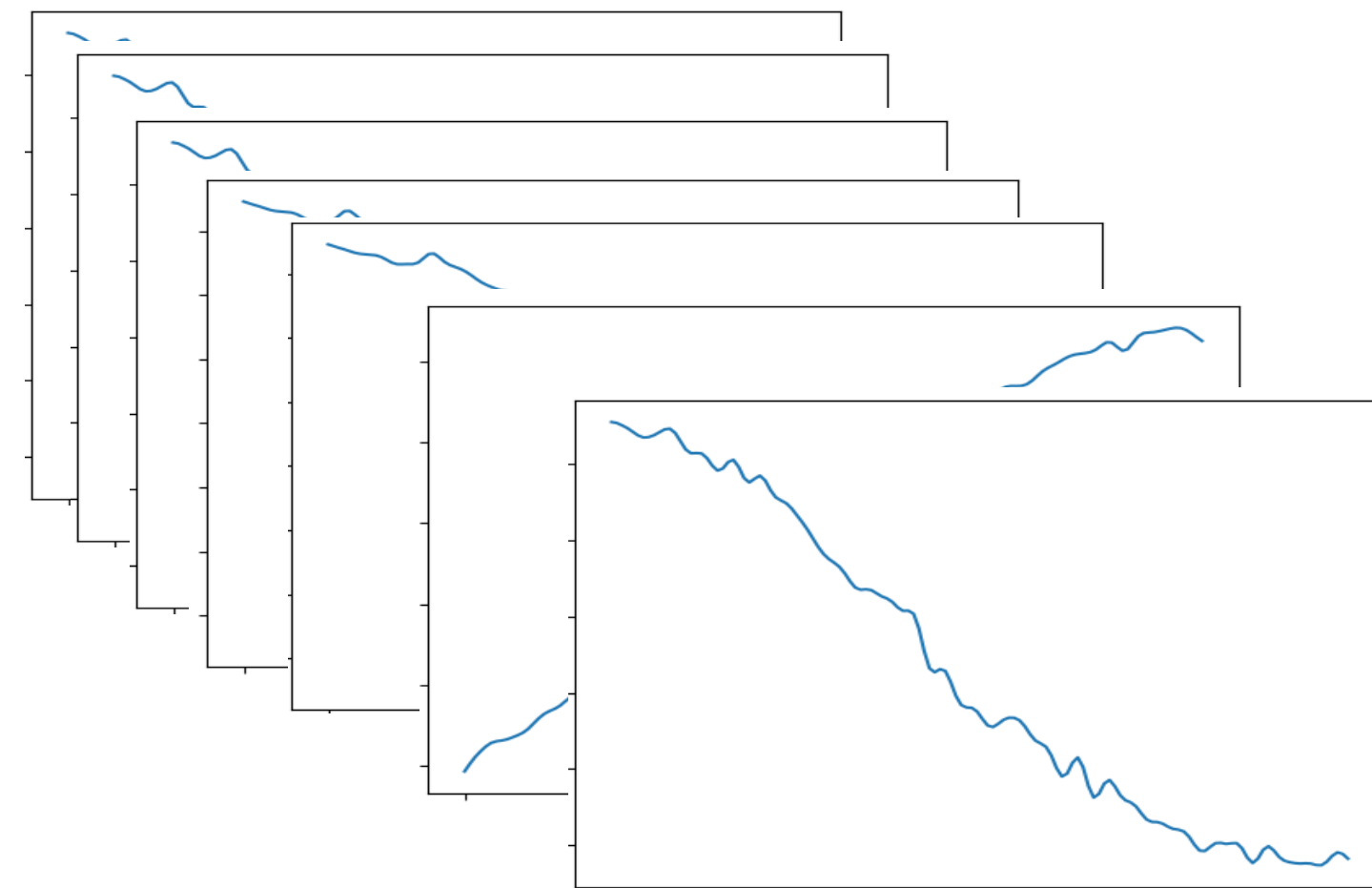
Appendix. Breath estimation

Uncover a relation between the warps and the corresponding breathing signal segment.



$$A$$
$$n_{train} \times q$$

$$\min_{X \in \mathbb{R}^q} \|AX - B\|$$

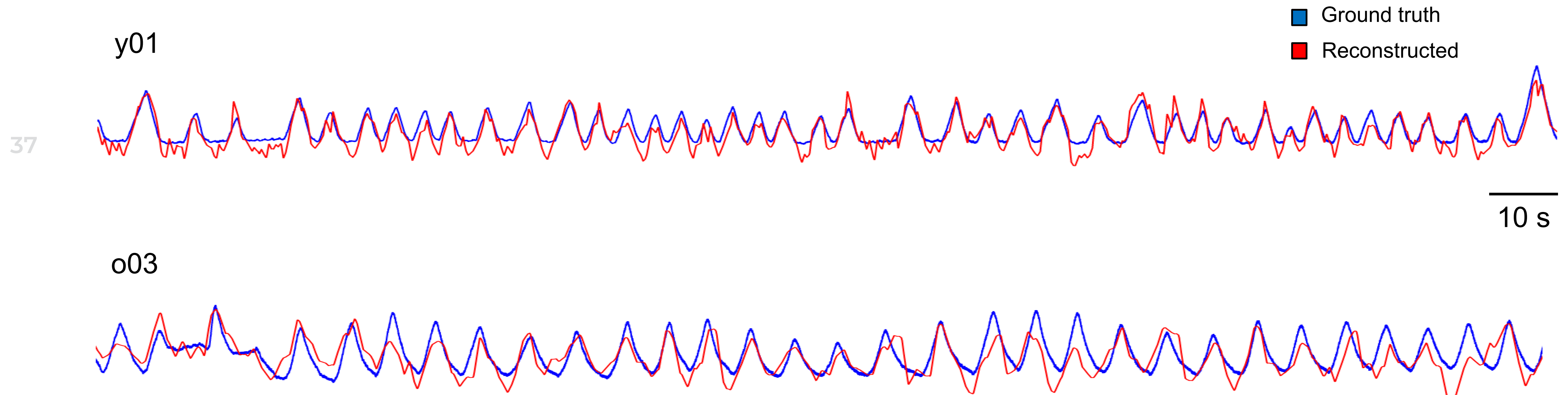


$$B$$
$$n_{train} \times q$$

Applications: breath estimation

Can we estimate the breathing signal from the ECG signal?

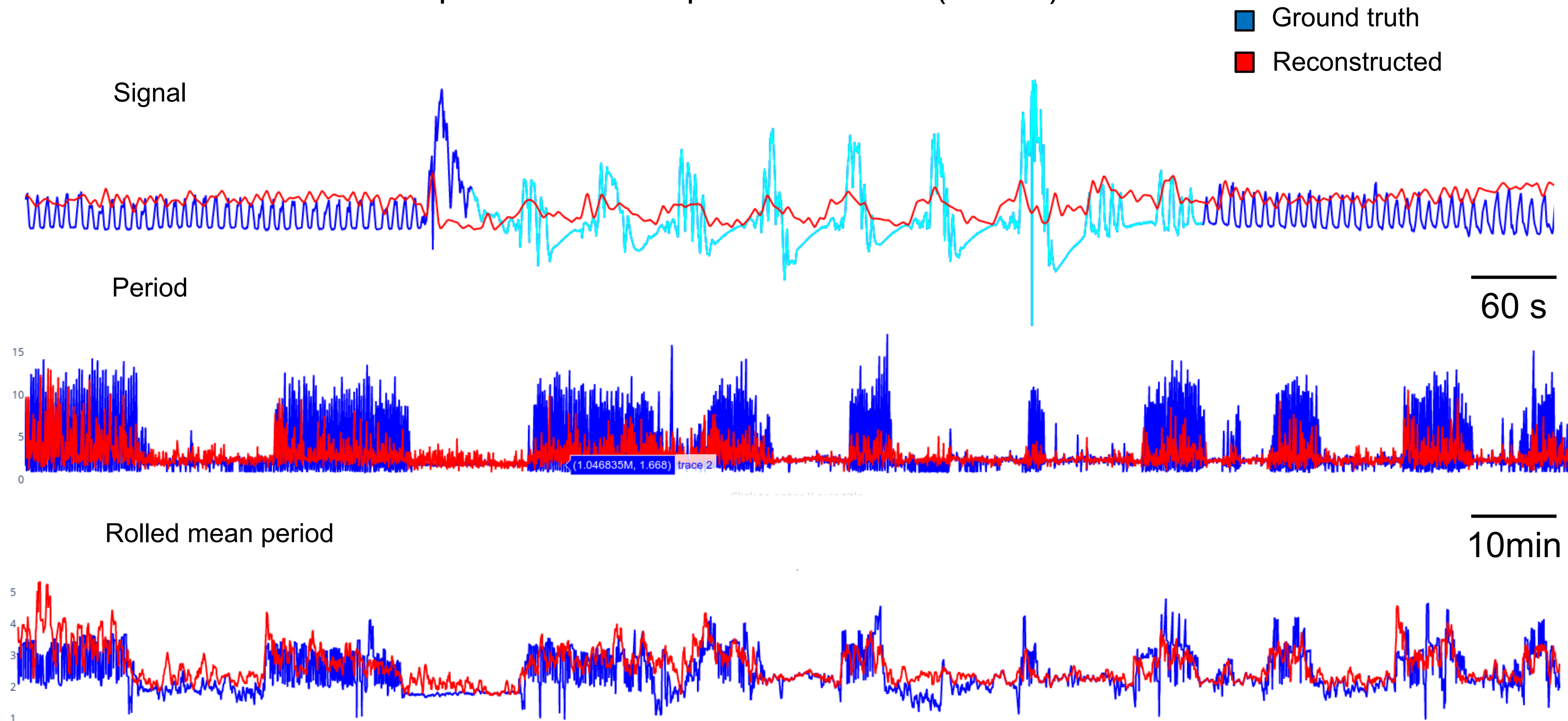
Reconstructing breath signal: Fantasia Database (250Hz)



Applications: breath estimation

Can we estimate the breathing signal from the ECG signal?

Period estimation in apnea records: Apnea Database (100Hz)



Applications: breath estimation

Discussion and future work

Components encode physiology.

Can the signal be reversed, work as an input for the warp computation?

With enough data, may we be able to estimate a breathing pattern without any respiration signal? (maybe not amplitude, but period)

Applications: takotsubo identification

Applications: takotsubo identification

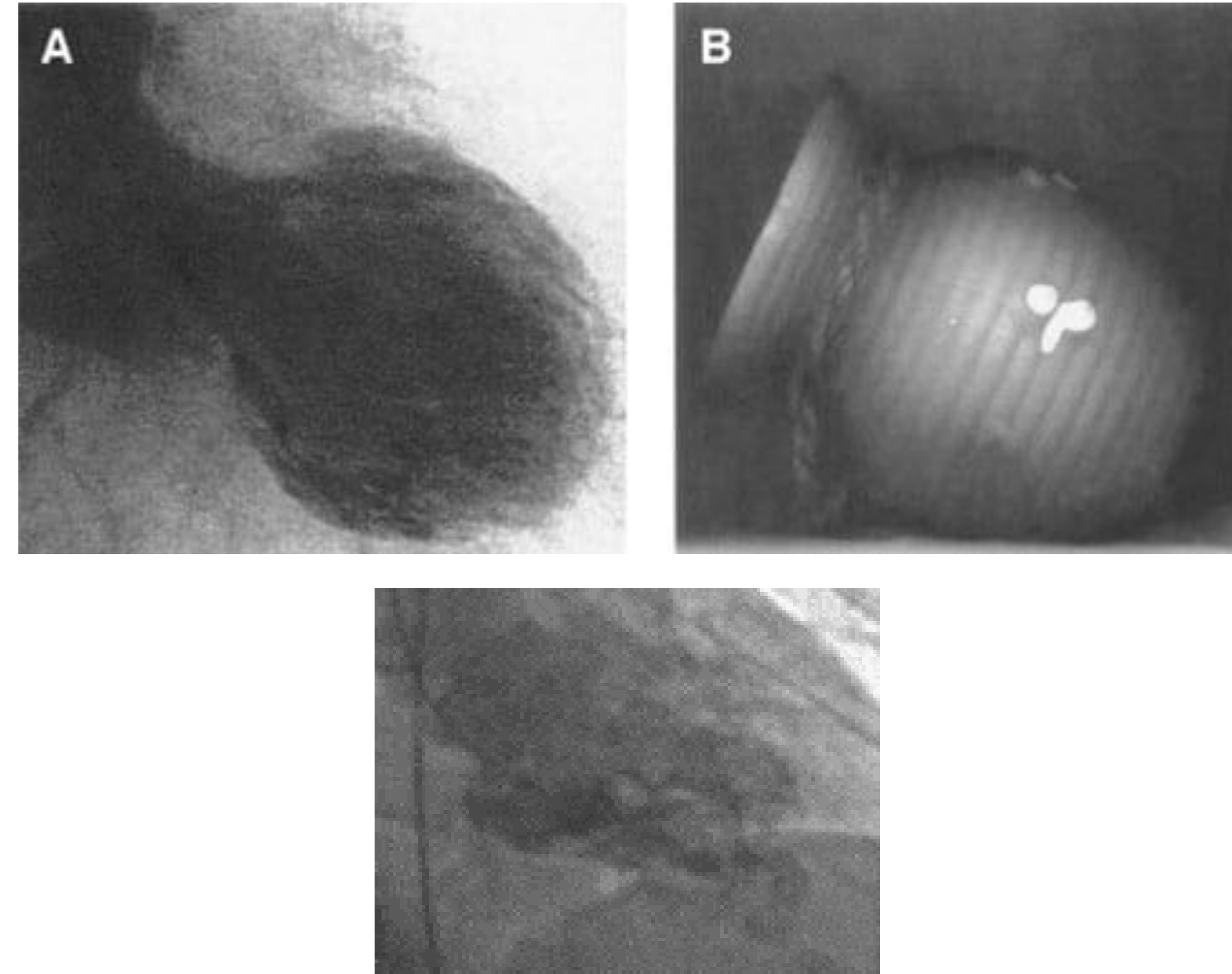
Acute cardiac condition which mimics myocardial infarction early clinical features

Uncommon but clinically relevant:
estimated to be 1-3% of acute coronary
syndrome presentations.

Affects mostly post-menopause woman.

Physical or emotional stress is frequent
but not always present.

There is no gold standard diagnostic test.

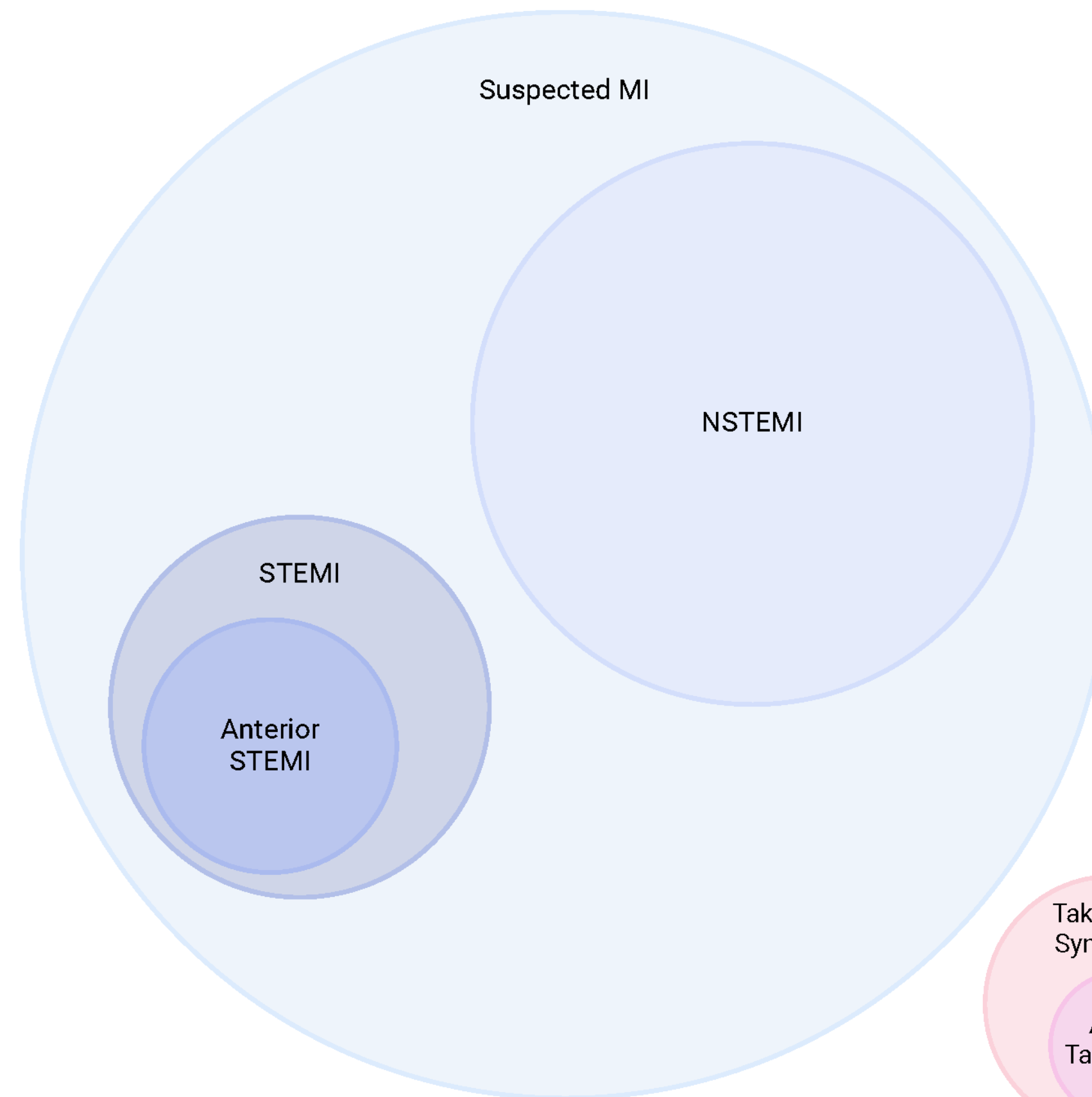


Kurisu S, Sato H, Kawagoe T, et al. Tako-tsubo-like left ventricular dysfunction with ST-segment elevation: a novel cardiac syndrome mimicking acute myocardial infarction

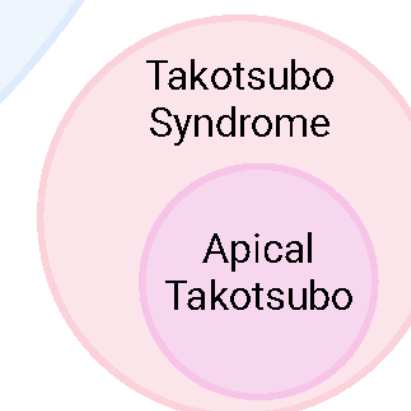
Applications: takotsubo identification

SCAAR database. Multi-centre retrospective study from Stockholm, Sweden. ECG database with patients admitted to hospital with a suspicion of MI and an indication for coronary angiography.

~15000 controls

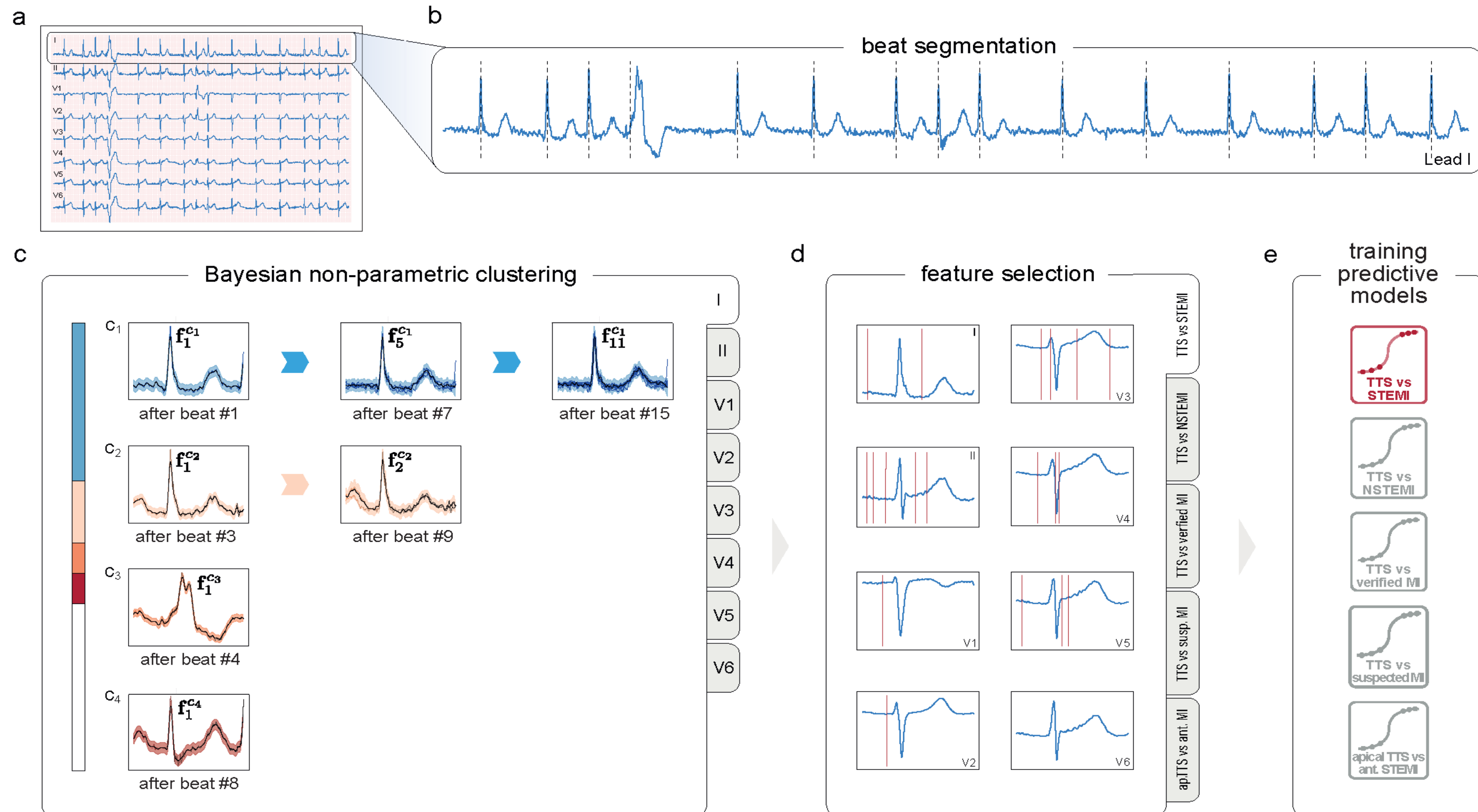


~500 Takotsubo cases



Applications: takotsubo identification

Graphical abstract



Applications: takotsubo identification

Results and comparison on the different scenarios.

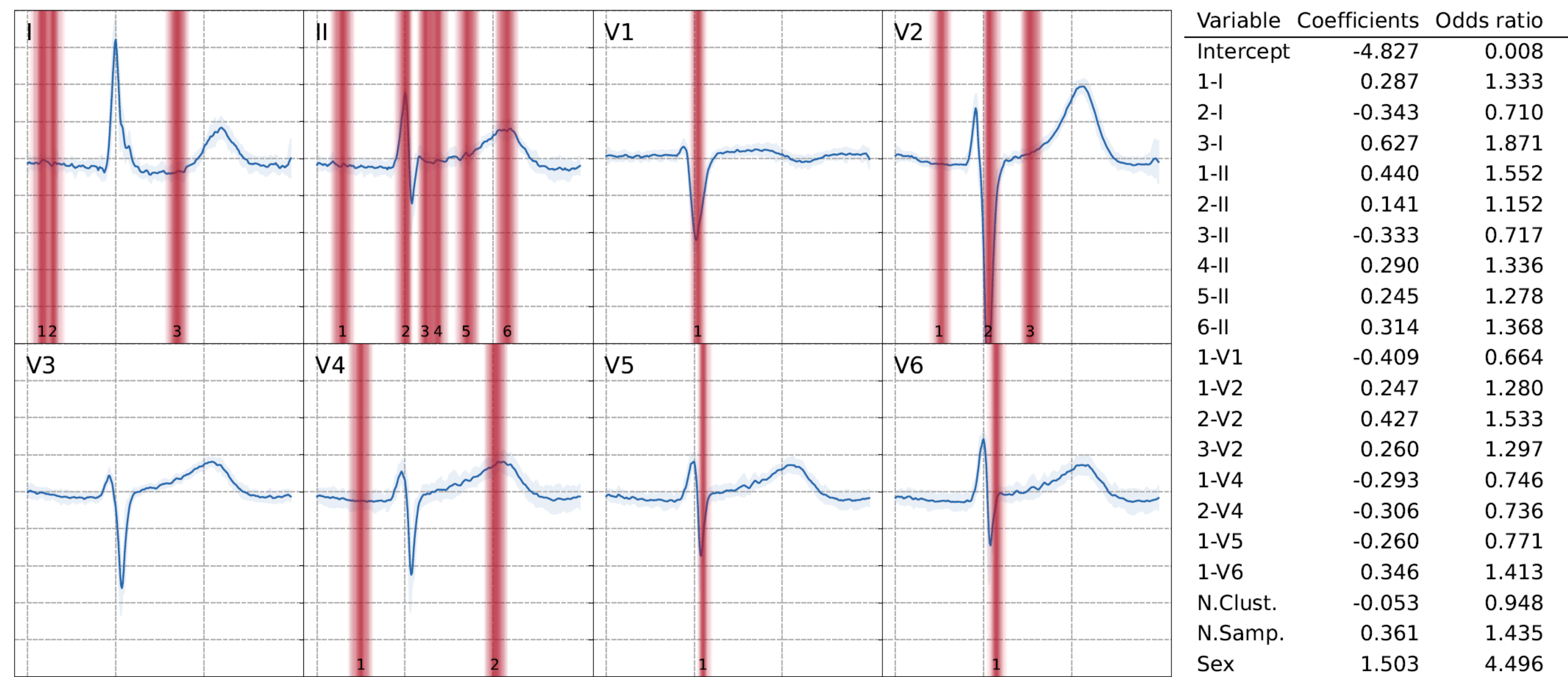
Metric	TTS vs suspected MI	TTS vs STEMI	TTS vs NSTEMI	TTS vs verified MI	apical TTS vs anterior STEMI
ROC AUC	0.89 (0.88-0.91)	0.94 (0.93-0.95)	0.91 (0.89-0.92)	0.92 (0.90-0.93)	0.90 (0.88-0.92)
PR AUC	0.29 (0.25-0.33)	0.84 (0.80-0.87)	0.59 (0.55-0.64)	0.53 (0.48-0.58)	0.71 (0.65-0.75)
Sensitivity	0.84 (0.80-0.87)	0.87 (0.84-0.90)	0.83 (0.79-0.86)	0.82 (0.78-0.85)	0.89 (0.85-0.92)
Specificity	0.80 (0.79-0.81)	0.86 (0.85-0.88)	0.84 (0.83-0.85)	0.87 (0.86-0.88)	0.79 (0.76-0.81)
F1 score	0.22 (0.20-0.24)	0.73 (0.70-0.75)	0.54 (0.51-0.57)	0.49 (0.46-0.52)	0.65 (0.61-0.68)
Accuracy	0.80 (0.80-0.81)	0.86 (0.85-0.88)	0.84 (0.83-0.85)	0.87 (0.86-0.87)	0.81 (0.79-0.83)
PPV	0.12 (0.11-0.14)	0.62 (0.58-0.66)	0.40 (0.37-0.43)	0.35 (0.32-0.38)	0.51 (0.46-0.55)
NPV	0.99 (0.99-0.99)	0.96 (0.95-0.97)	0.98 (0.97-0.98)	0.98 (0.98-0.99)	0.97 (0.96-0.98)

Scenario	Håkansson <i>et al.</i>		HDP-GPC approach	
	ROC AUC	PR AUC	ROC AUC	PR AUC
TTS vs suspected MI	0.85	0.26	0.89 (0.88–0.91)	0.29 (0.25–0.33)
TTS vs verified MI	0.87	0.48	0.92 (0.90–0.93)	0.53 (0.48–0.58)
TTS vs STEMI	0.88	0.72	0.94 (0.93–0.95)	0.84 (0.80–0.87)
TTS vs NSTEMI	0.86	0.54	0.91 (0.89–0.92)	0.59 (0.55–0.64)
Apical TTS vs anterior STEMI	0.84	0.59	0.90 (0.88–0.92)	0.71 (0.65–0.75)

Applications: takotsubo identification

Logistic regression provides interpretable features.

TTS vs suspected MI



Applications: takotsubo identification

Criteria identified from the interpretable features.

	Novelty	[1,2,3,4,5]	[3,5]	[1,2,4]	[1,2,4]	[1]	[1,2,4]
Findings	Prolonged P-wave or biphasic shape	ST-elevation	ST-elevation + reciprocal changes	ST-depression	Absence of ST-depression	T-wave inversion	QT-prolongation
TTS from suspected MI	I	II	–	–	I	V4	–
TTS from verified MI	I, II	II, V5, V6	–	–	I	V4	II
TTS from STEMI	I, II	II	I, II, V2, V3	–	I	–	V3
TTS from NSTEMI	I, II, V4	I, II	–	–	–	V4	–
Apical TTS from anterior STEMI	II, V1, V2	I, II	I, II	V1	–	I	I

[1] Christian Templin, Jelena R. Ghadri, Johanna Diekmann, et al. Clinical features and outcomes of Takotsubo (stress) cardiomyopathy. *New England Journal of Medicine*, 373:929–938, 2015.

[2] Antonio H. Frangieh, Slayman Obeid, Jelena-Rima Ghadri, et al. ECG criteria to differentiate between Takotsubo (stress) cardiomyopathy and myocardial infarction. *Journal of the American Heart Association*, 5, 2016.

[3] Gassan Moady, Gal Rubinstein, Loai Mobarki, and Shaul Atar. Takotsubo syndrome in the emergency room — diagnostic challenges and suggested algorithm. *Rev. Cardiovasc. Med.*, 23(4):131, 2022.

[4] Jelena-R. Ghadri, Victoria L. Cammann, Stjepan Jurisic, et al. A novel clinical score (InterTAK diagnostic score) to differentiate Takotsubo syndrome from acute coronary syndrome: results from the international Takotsubo registry. *European Journal of Heart Failure*, 19:1036–1042, 2017.

[5] Riyo Ogura, Yoshikazu Hiasa, Takefumi Takahashi, et al. Specific Findings of the Standard 12-Lead ECG in Patients With ‘Takotsubo’ Cardiomyopathy-Comparison With the Findings of Acute Anterior Myocardial Infarction. *Circulation Journal*, 67:687–690, 2003.

Applications: takotsubo identification

Discussion and future work

External validation

Promising results using the GP mean, but what happens with the rest of the model parameters?

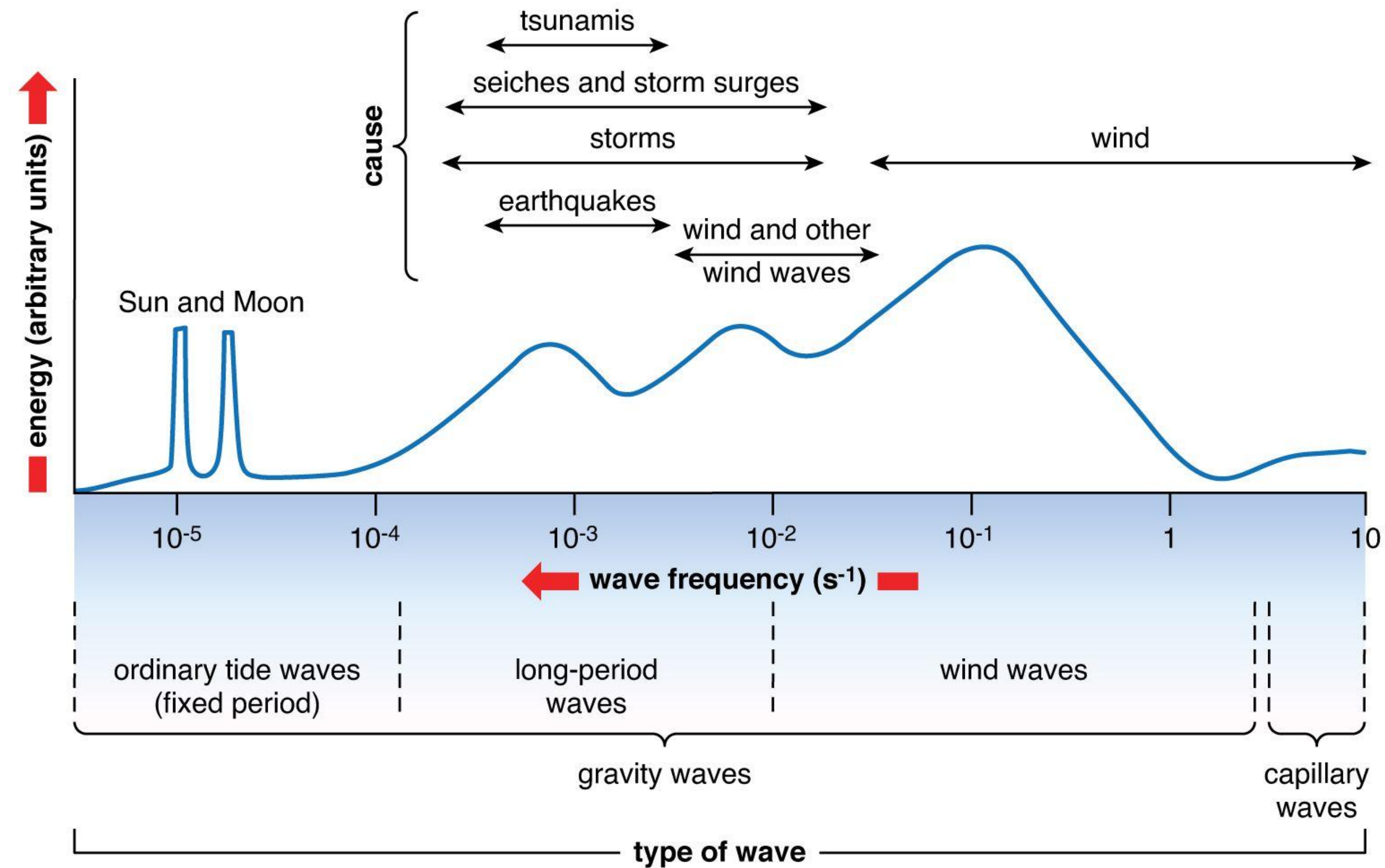
LDS parameters – encode dynamical information

Need of longer records

Applications: wave spectra analysis

Applications: wave spectra analysis

The surface wave field is affected by phenomena which ranges very different temporal and spatial scales.

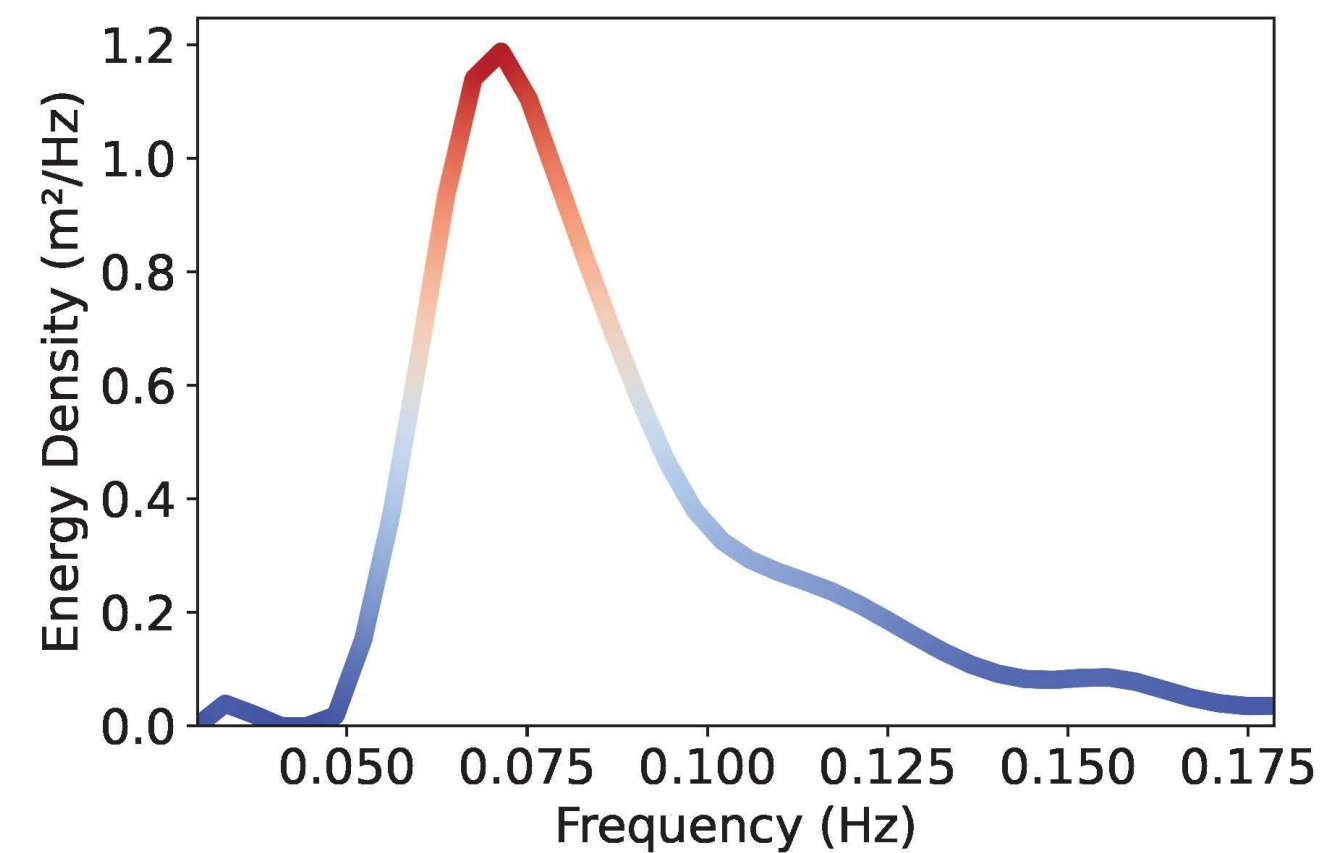


Applications: wave spectra analysis

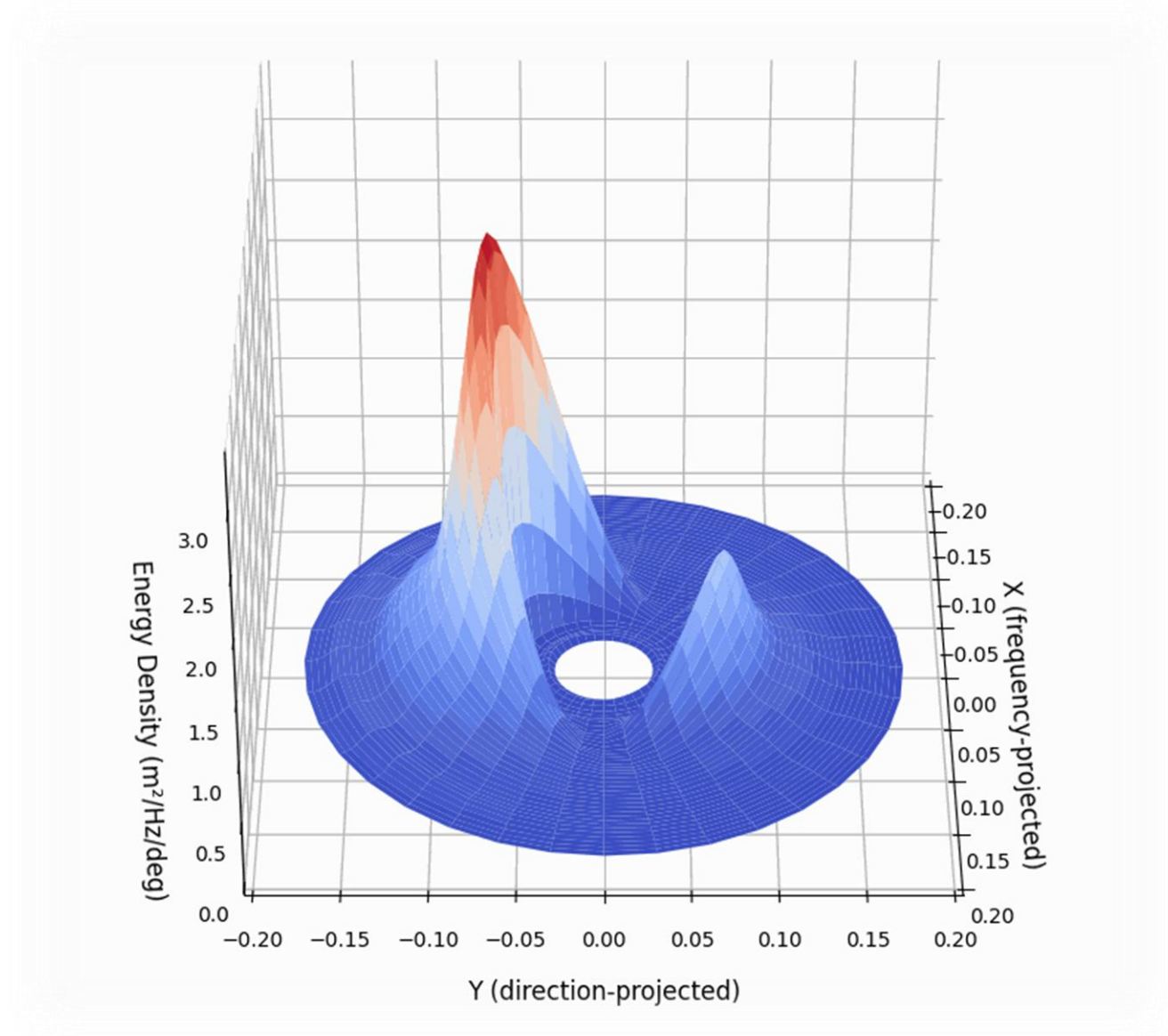
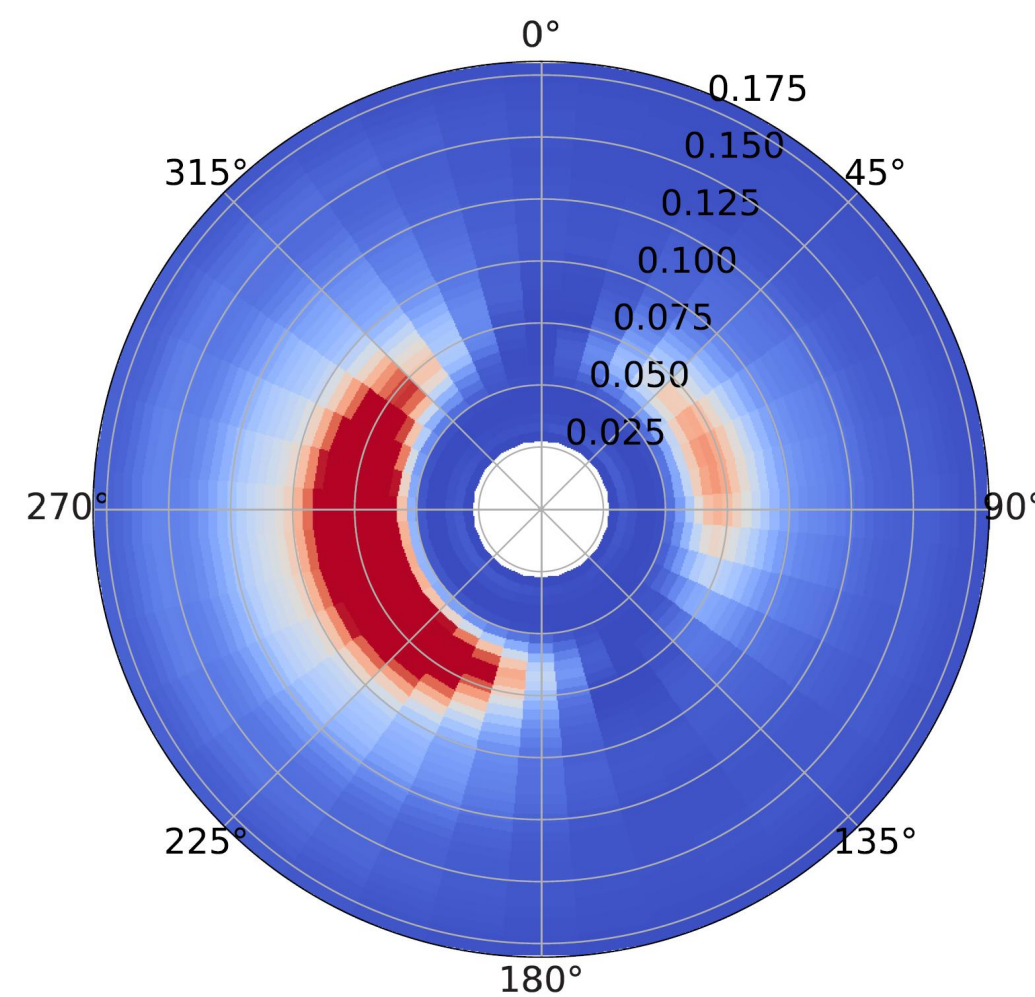
Directional spectra

The directional component of the spectra could be integrated but we lose some information.

$$S_n(\nu)$$



$$S_n(\nu, \theta)$$



$$S_n(\nu) = \int S_n(\nu, \theta) d\theta.$$

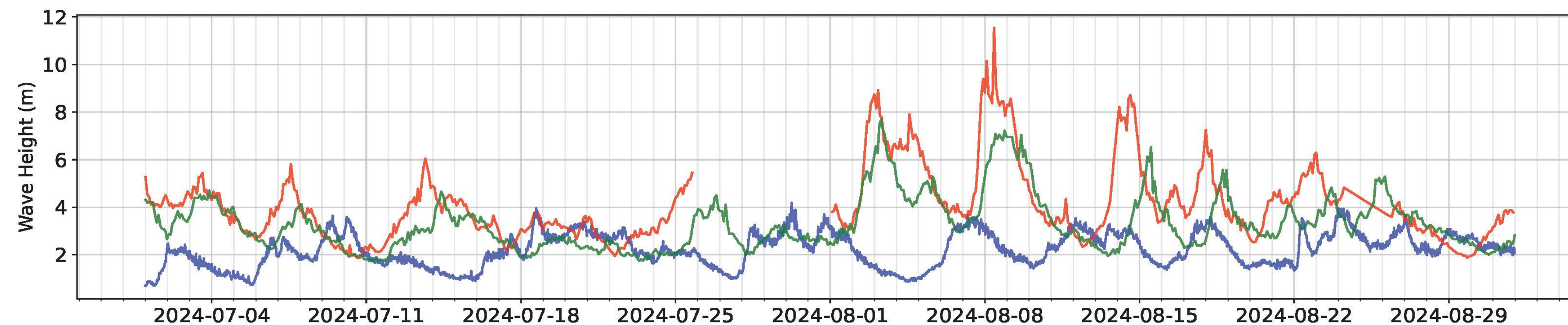
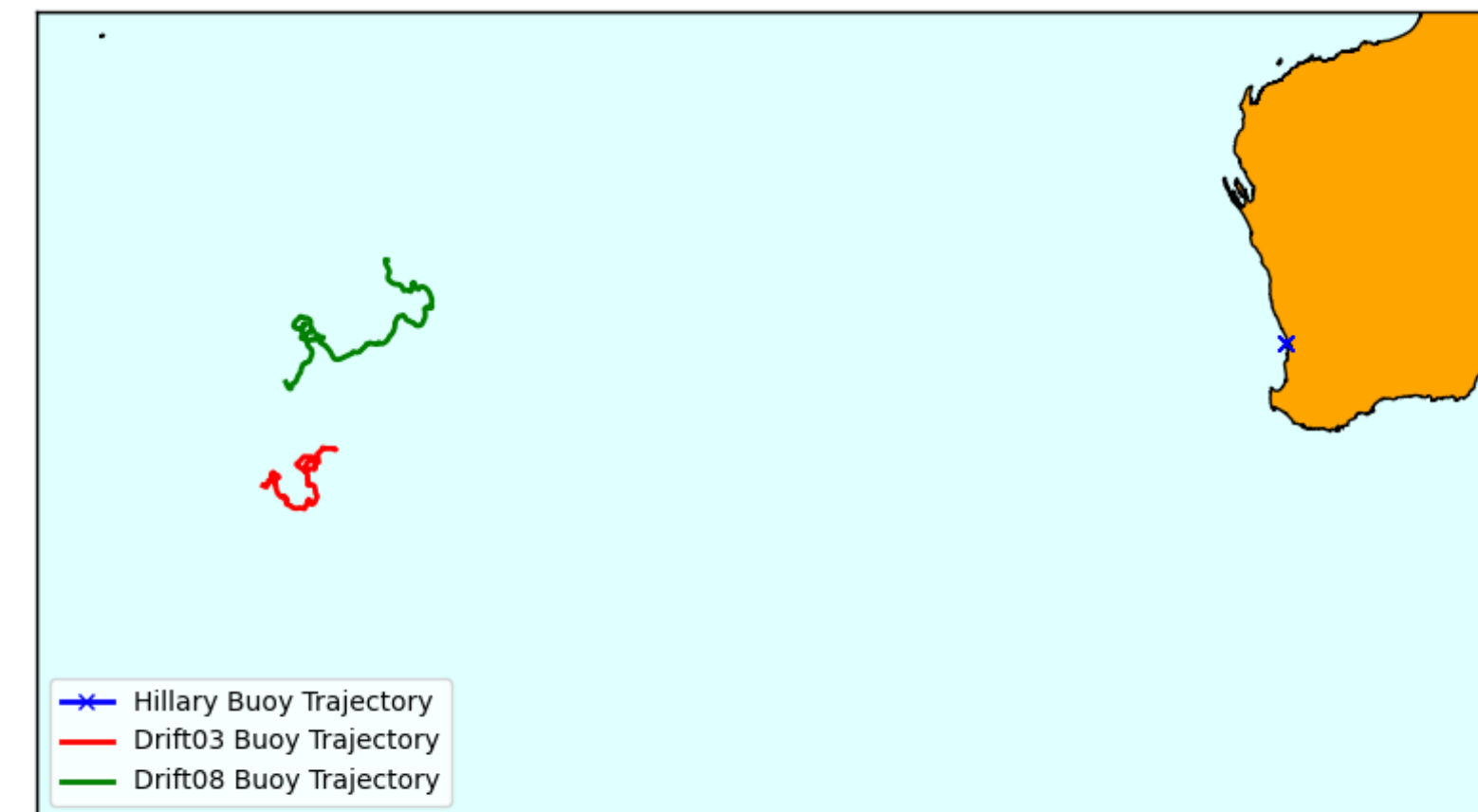
Applications: wave spectra analysis

Australian three-buoy dataset (Indian Ocean) during austral winter of 2024 July-August

One coastal moored buoy: **Hillarys**

Two drifting buoys Sofar Ocean Spotter devices: **Drift03** and **Drift08**

Point estimates fails due to asynchronous event registration due to drifting and asymmetric origin.



Applications: wave spectra analysis

Multi-output approach

We worked with directional spectra using the multi-output inference method.

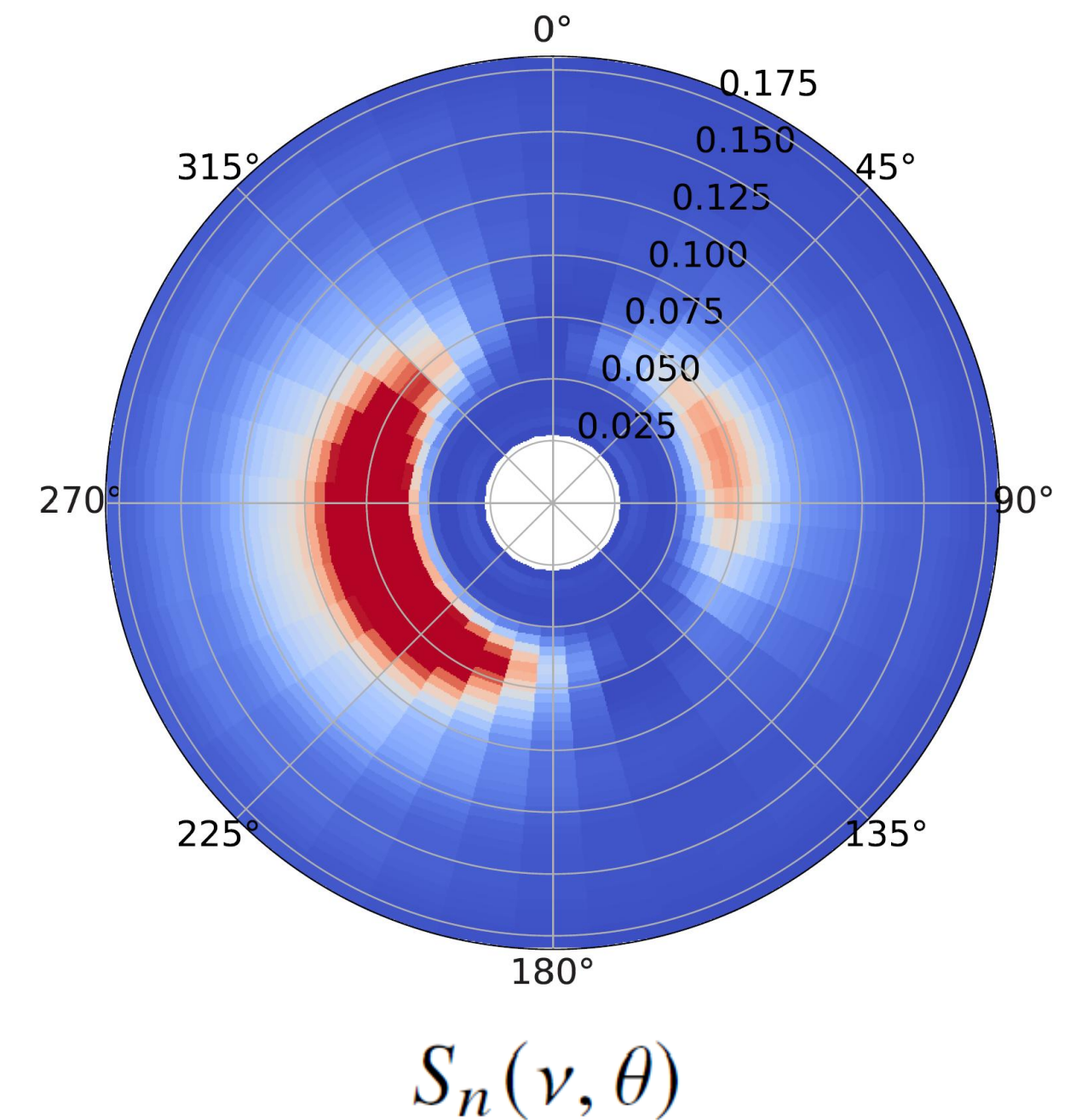
$$p(\mathbf{y}_n | s_n, \mathcal{X}, \mathcal{G}) \propto \prod_{l=1}^L \left[p(\mathbf{y}_n^{(l)} | s_n, \mathcal{X}^{(l)}, \mathcal{G}^{(l)}) \right]^{w_n^{(l)}}$$

We define the weights per direction using the energy proportion per direction

$$\log \zeta_{njk} = \mathbb{E}[\log \pi_{jk}] + \sum_{l=1}^L w_n^{(l)} \log \zeta_{njk}^{(l)}$$

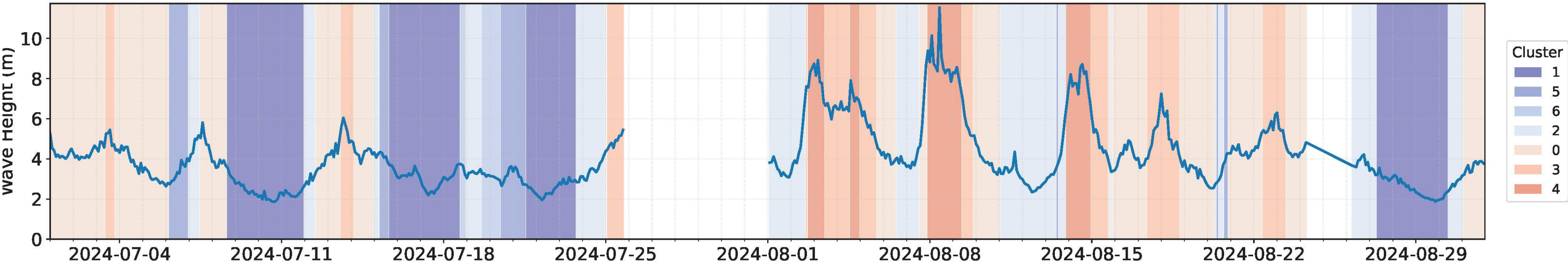
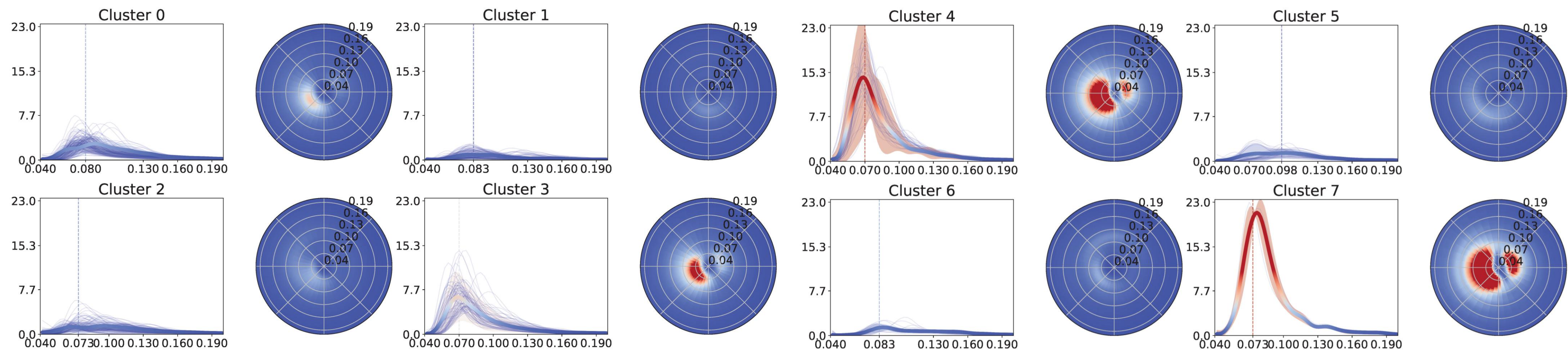
where

$$w_n(l) = \frac{\sum_{j=1}^{N_f} S_n(f_j, \theta_l)}{\sum_{j=1}^{N_f} \sum_{l'=1}^{N_l} S_n(f_j, \theta_{l'})} = \frac{\text{total energy in direction } \theta_l}{\text{total energy across all directions}},$$



Applications: wave spectra analysis

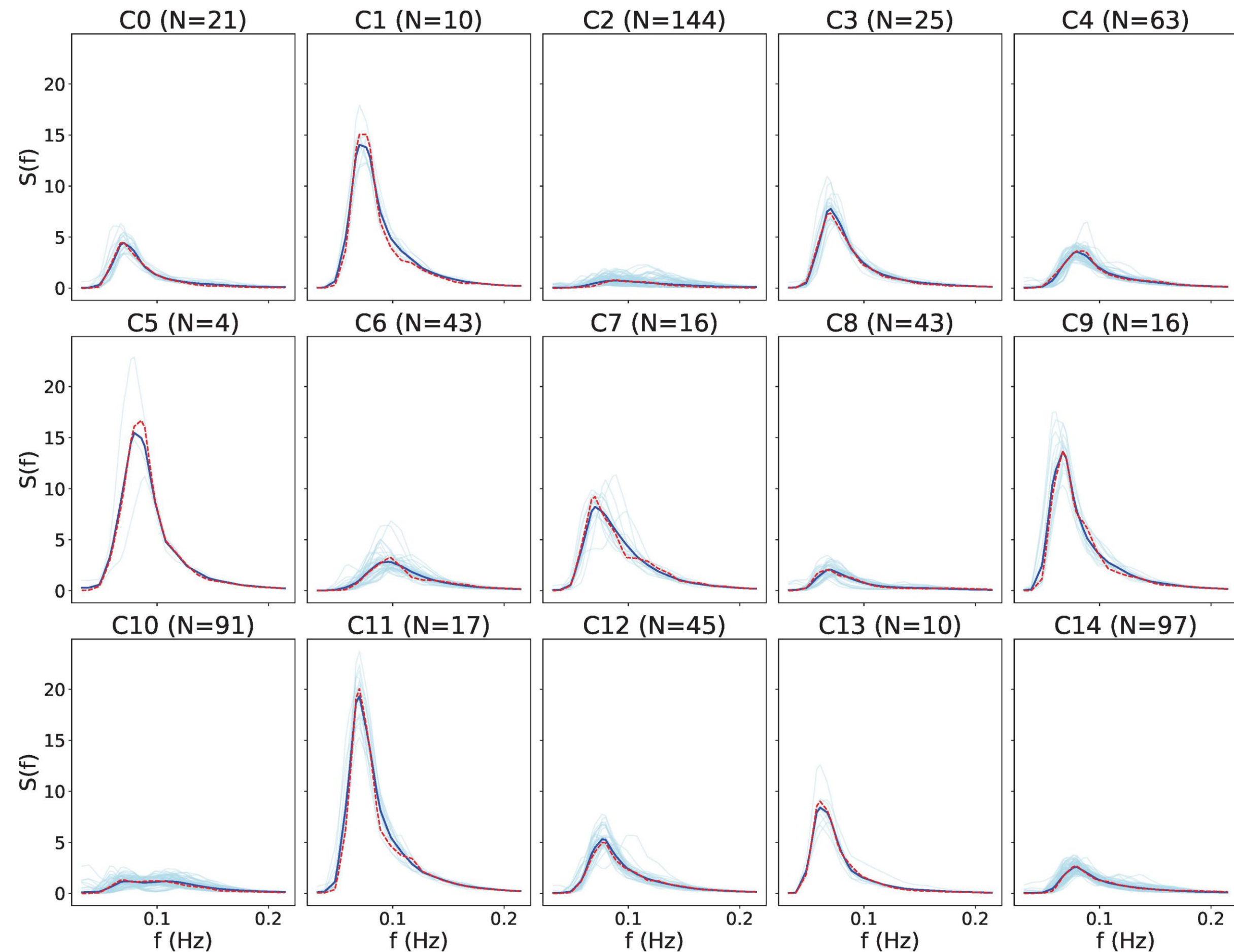
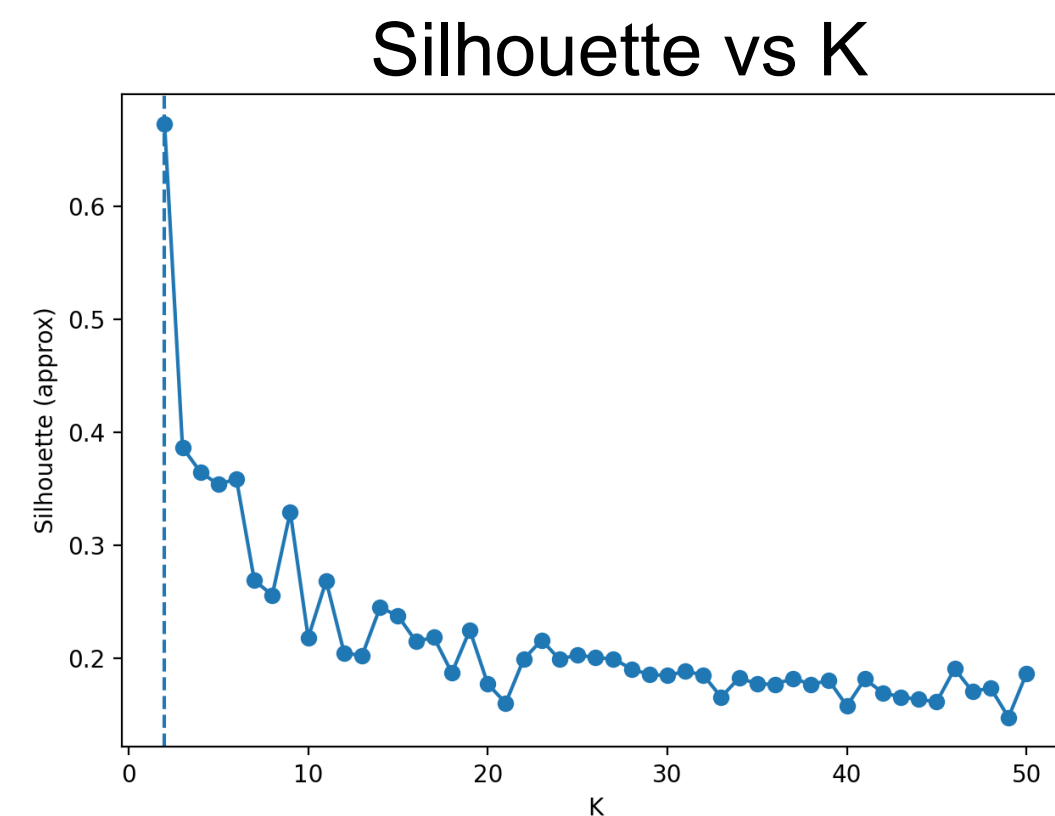
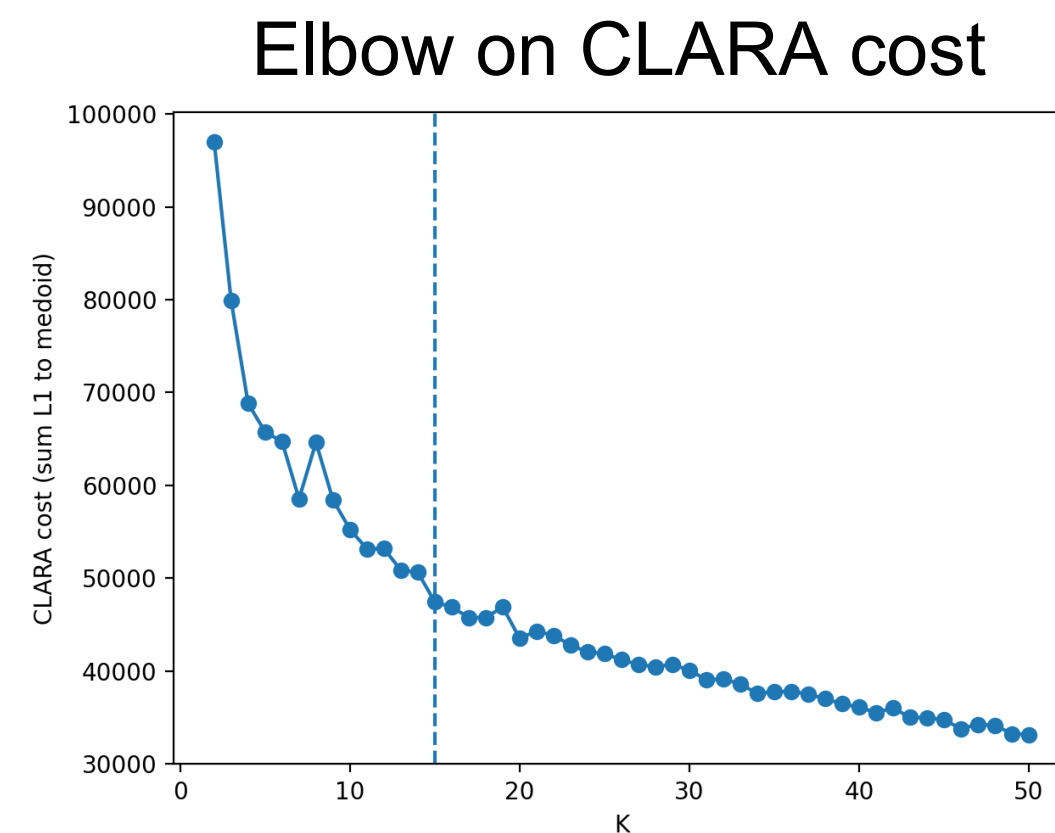
Directional clustering Drift03



Applications: wave spectra analysis

CLARA (Hamilton, 2010) comparison on Drift03

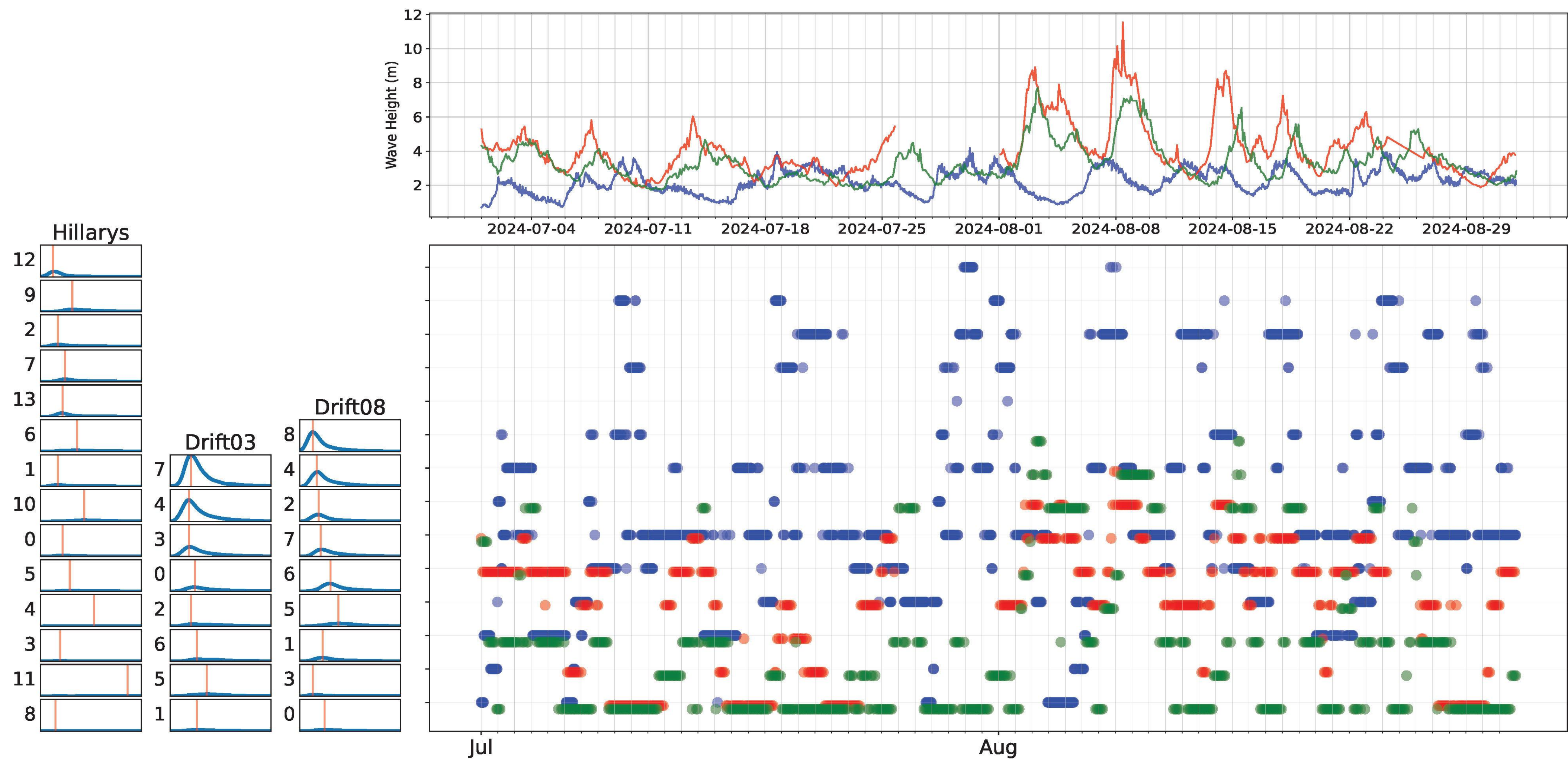
Inconsistent criteria to select number of clusters:



Applications: wave spectra analysis

Event identification

53



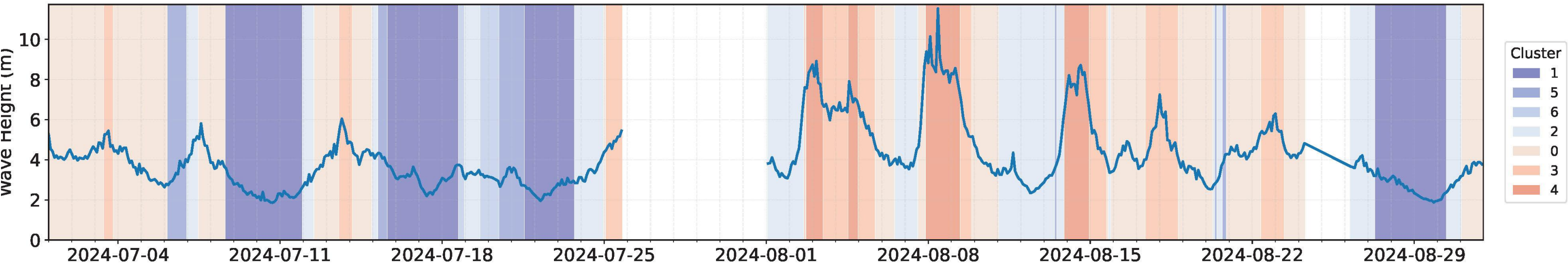
Applications: wave spectra analysis

Event pair identification

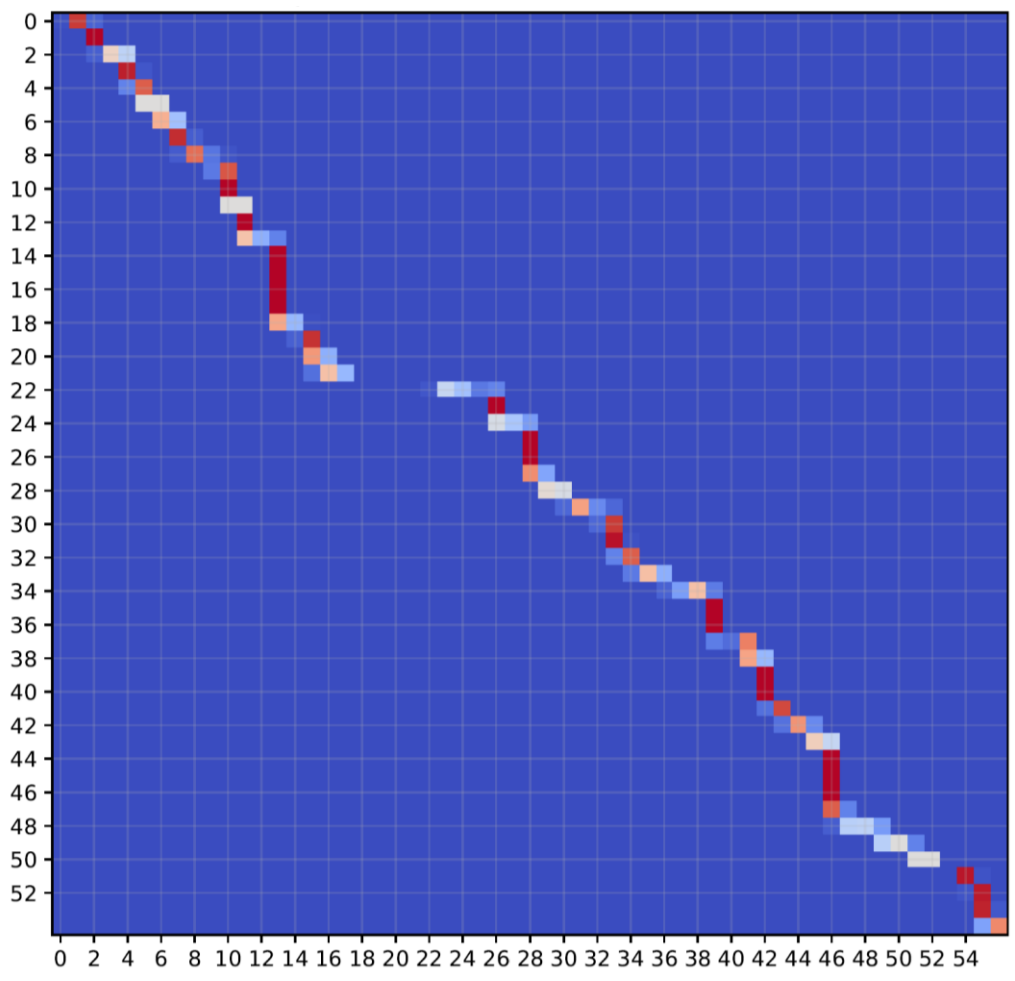
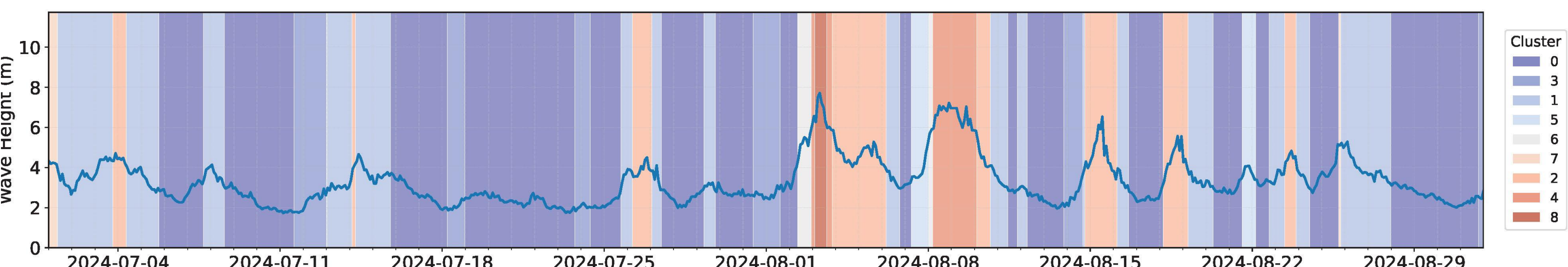
Using a physics-based approach we defined an event similarity function composed by three terms: time-lag, angular mismatch and Kullback-Leibler divergence between prototypes.

$$\Delta t_{m,m'} \approx \frac{d_{m,m'}^{p||q}}{c_g(f_{p,m})}, \quad \Delta \theta_{m,m'} = (\theta_{p,m} - \theta_{p,m'} - r360^\circ)^2, \quad D_{\text{KL}}(P||Q) = \frac{1}{2} \left[(\boldsymbol{\mu}_{08} - \boldsymbol{\mu}_{03})^\top \boldsymbol{\Sigma}_{08}^{-1} (\boldsymbol{\mu}_{08} - \boldsymbol{\mu}_{03}) + \text{tr}(\boldsymbol{\Sigma}_{08}^{-1} \boldsymbol{\Sigma}_{03}) - k + \ln \frac{|\boldsymbol{\Sigma}_{08}|}{|\boldsymbol{\Sigma}_{03}|} \right].$$

Drift03



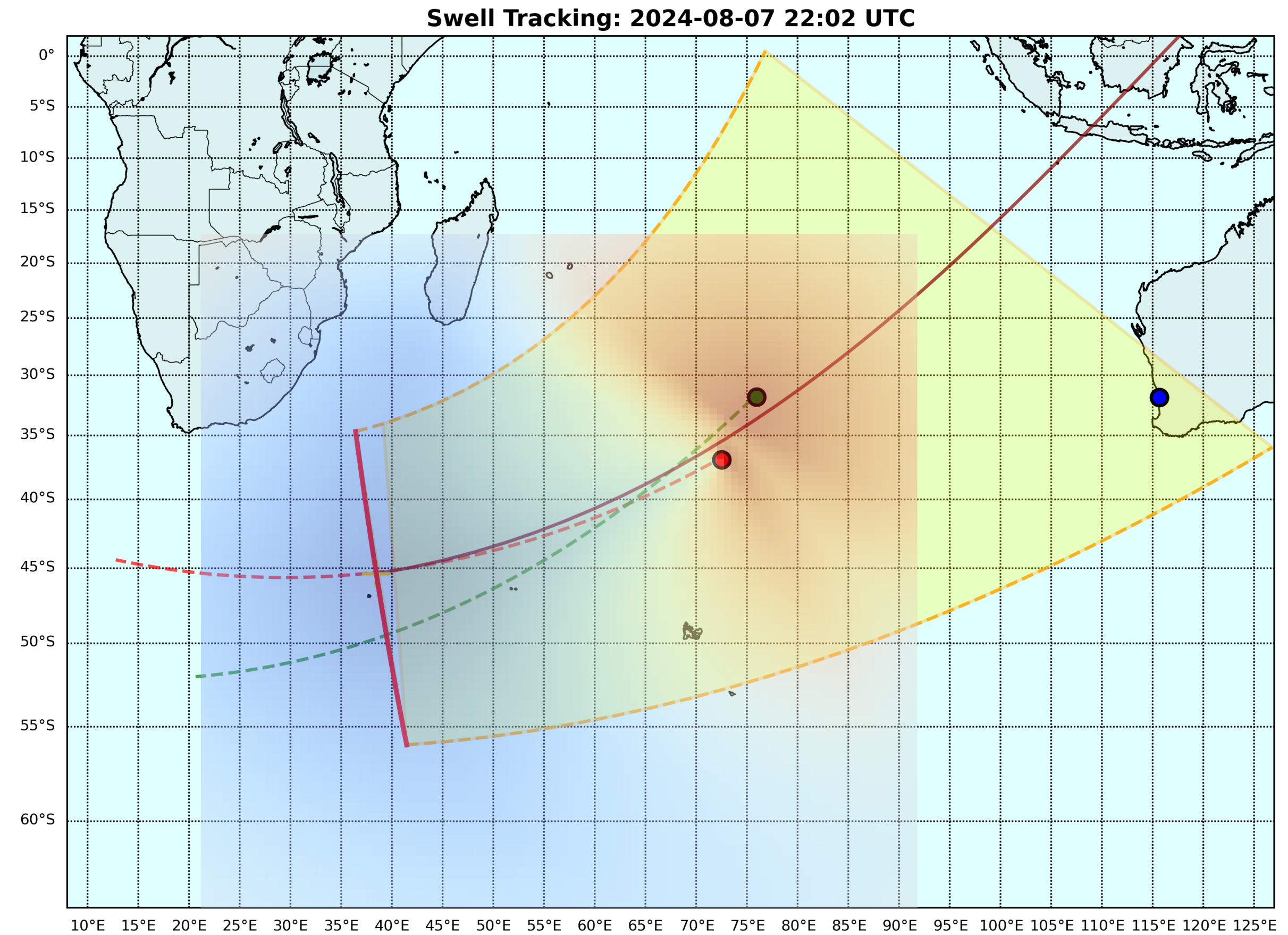
Drift08



Applications: wave spectra analysis

Source location and prediction using paired events

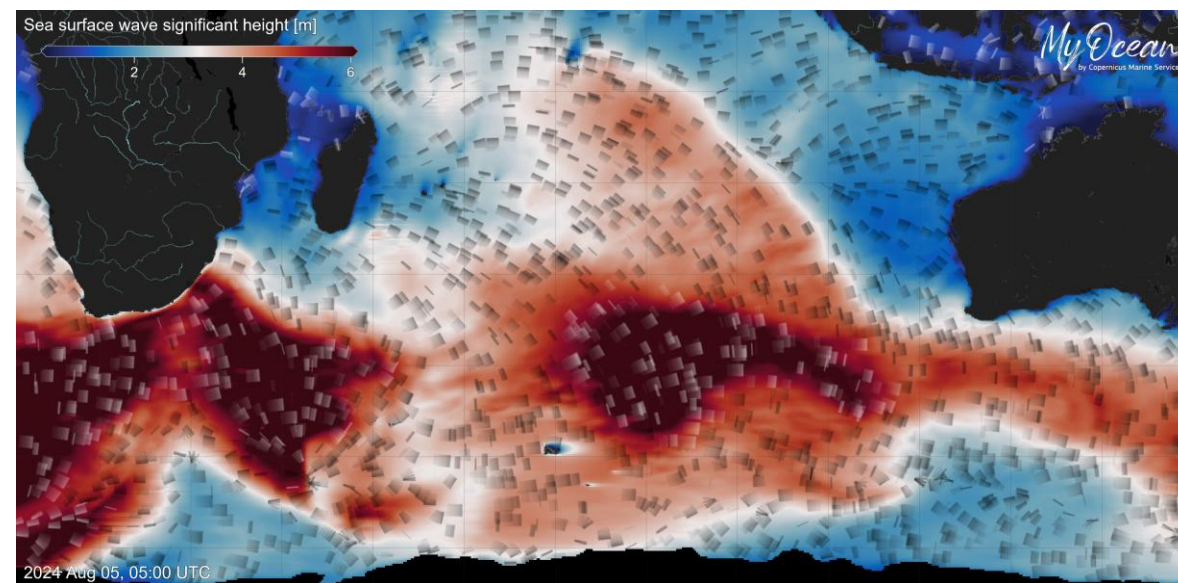
Defining a loss function depending on information of the paired events we were able to estimate the source and the reaching time.



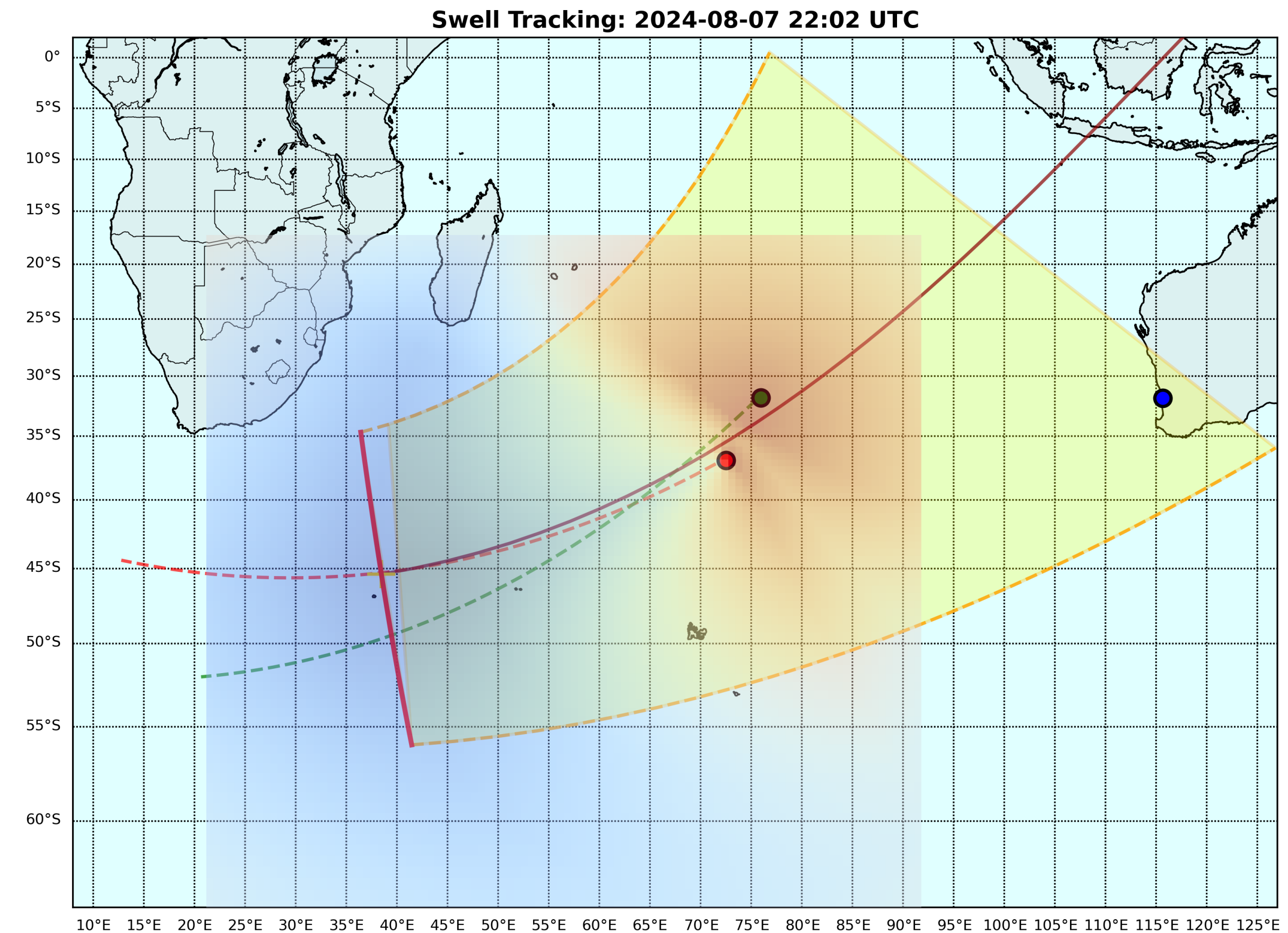
Applications: wave spectra analysis

Source location and prediction using paired events

Defining a loss function depending on information of the paired events we were able to estimate the source and the reaching time.

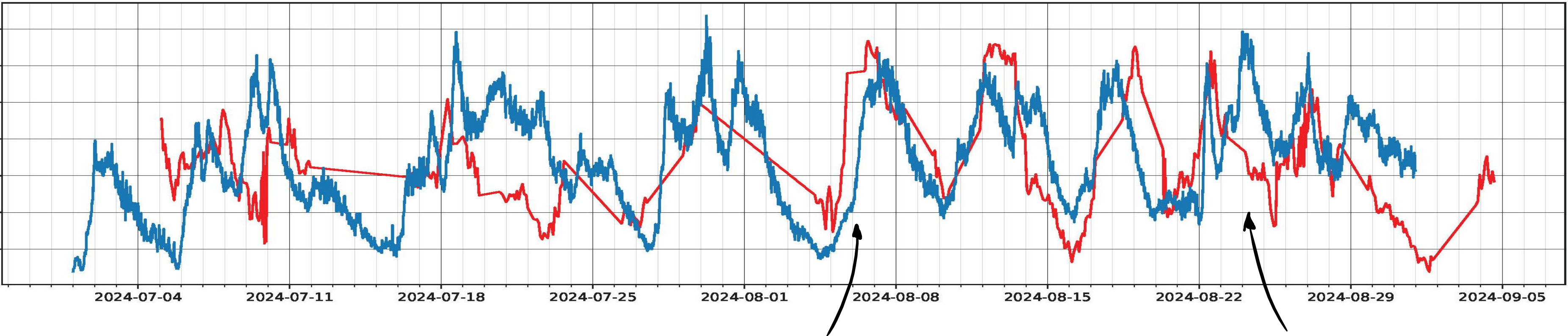
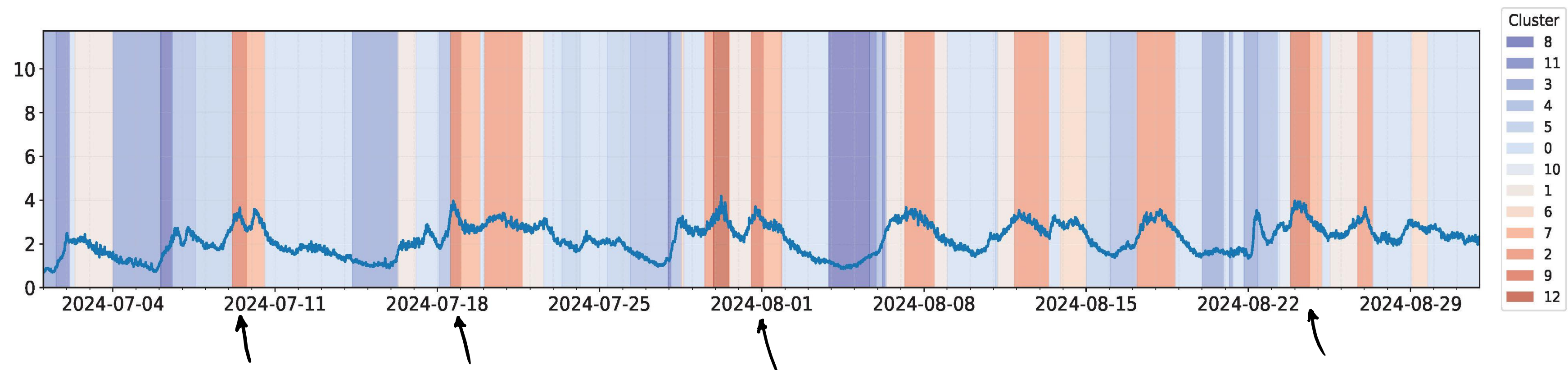


Copernicus Marine Service. Sea surface wave significant height [m] for 2024 August 05 05:00 UTC.



Applications: wave spectra analysis

Source location and prediction using paired events



Arrival time of wave front prediction

Detection of local events 9

Applications: wave spectra analysis

Discussion and future work

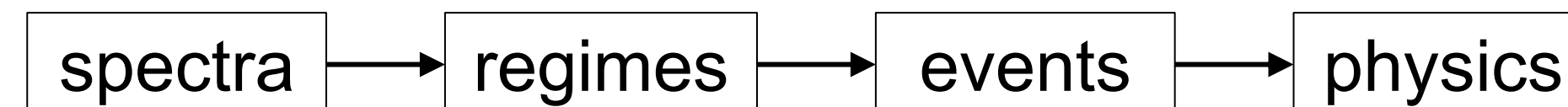
Clustering becomes event discovery.

Directional spectra preserve multi-system wave structure.

HDP avoids choosing K as in classic approaches as CLARA.

LDS avoids artificial fragmentation.

Avoid observation-by-observation inconsistencies.



Future work

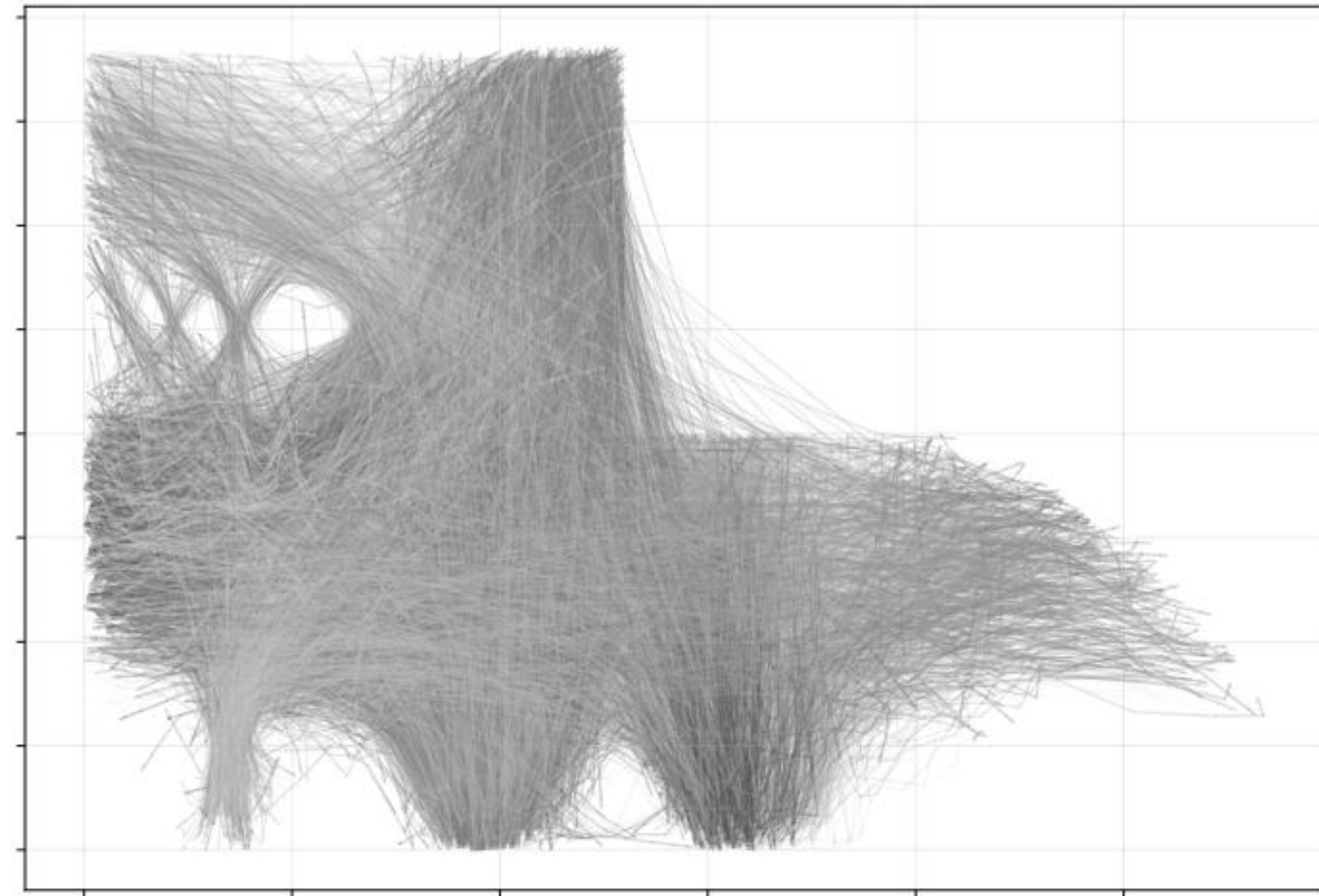
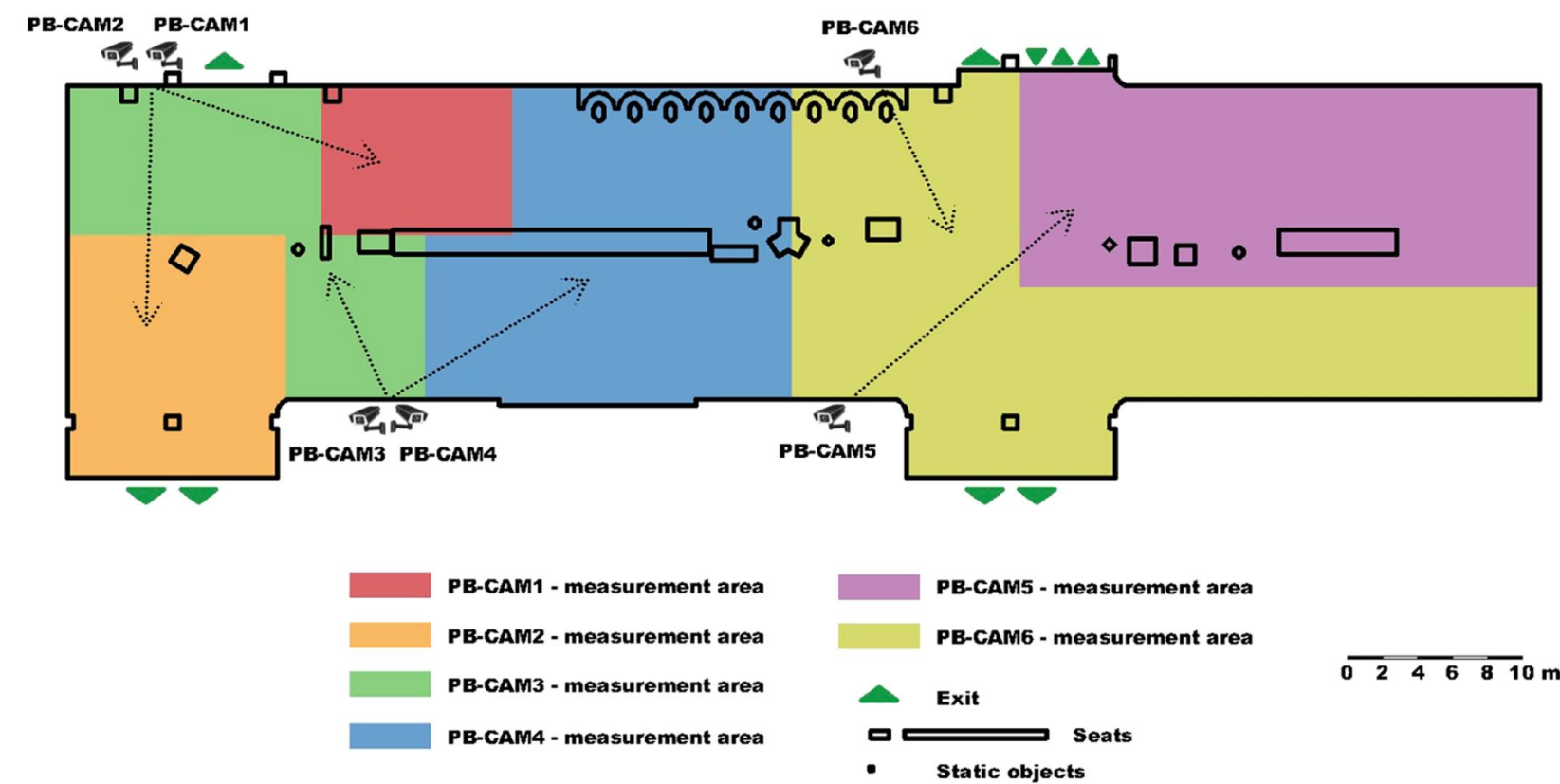
Joint buoy modelling / spatio-temporal GP representation.

Regime dynamics for arrival prediction.

**Applications: pedestrian
trajectory clustering**

Applications: pedestrian trajectory clustering

Spatiotemporal Trajectories of Pedestrian Mobility at the Train Station



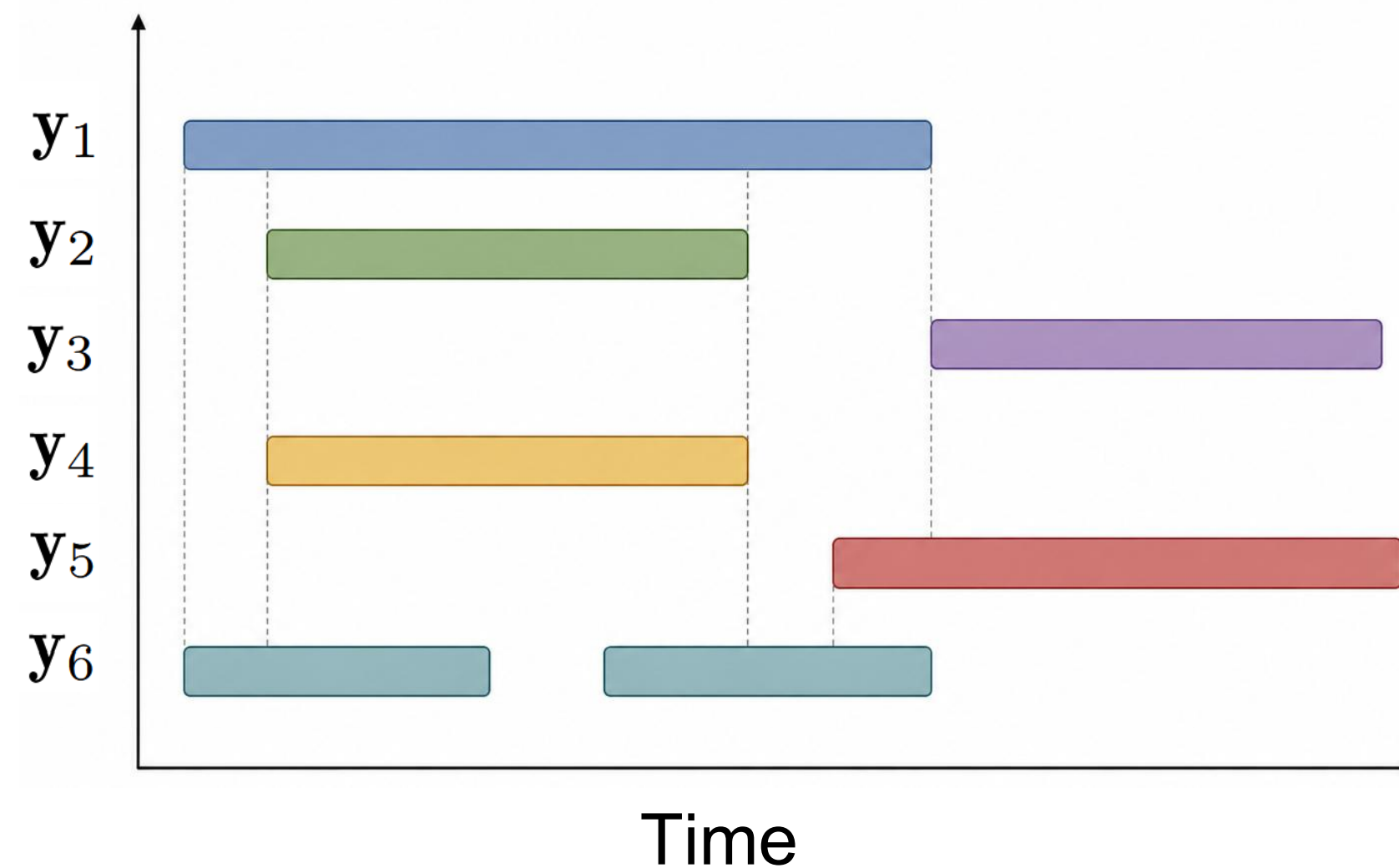
Tomáš Apeltauer, Ondřej Uhlík, Jiří Apeltauer, and Vojtěch Juřík. Spatiotemporal trajectories of pedestrian mobility at the train station: Evidence of 24 million trajectories. *Scientific Data*, 11(1):1254, 2024.

Applications: pedestrian trajectory clustering

Time problem

Ordering is local within each trajectory, but not globally well-defined in the database.

$$y_n(\tau) = (x_n(\tau), z_n(\tau)), \quad \tau \in [0, 1]$$



Nested intervals

Overlapping intervals

Interrupted trajectories

Interrupted IDs

No obvious information on the ordering

Applications: pedestrian trajectory clustering

Time problem

Ordering is local within each trajectory, but not globally well-defined in the database.

$$\mathbf{y}_n(\tau) = (x_n(\tau), z_n(\tau)), \quad \tau \in [0, 1]$$

We can collapse all transitions rows to the global HDP weights,

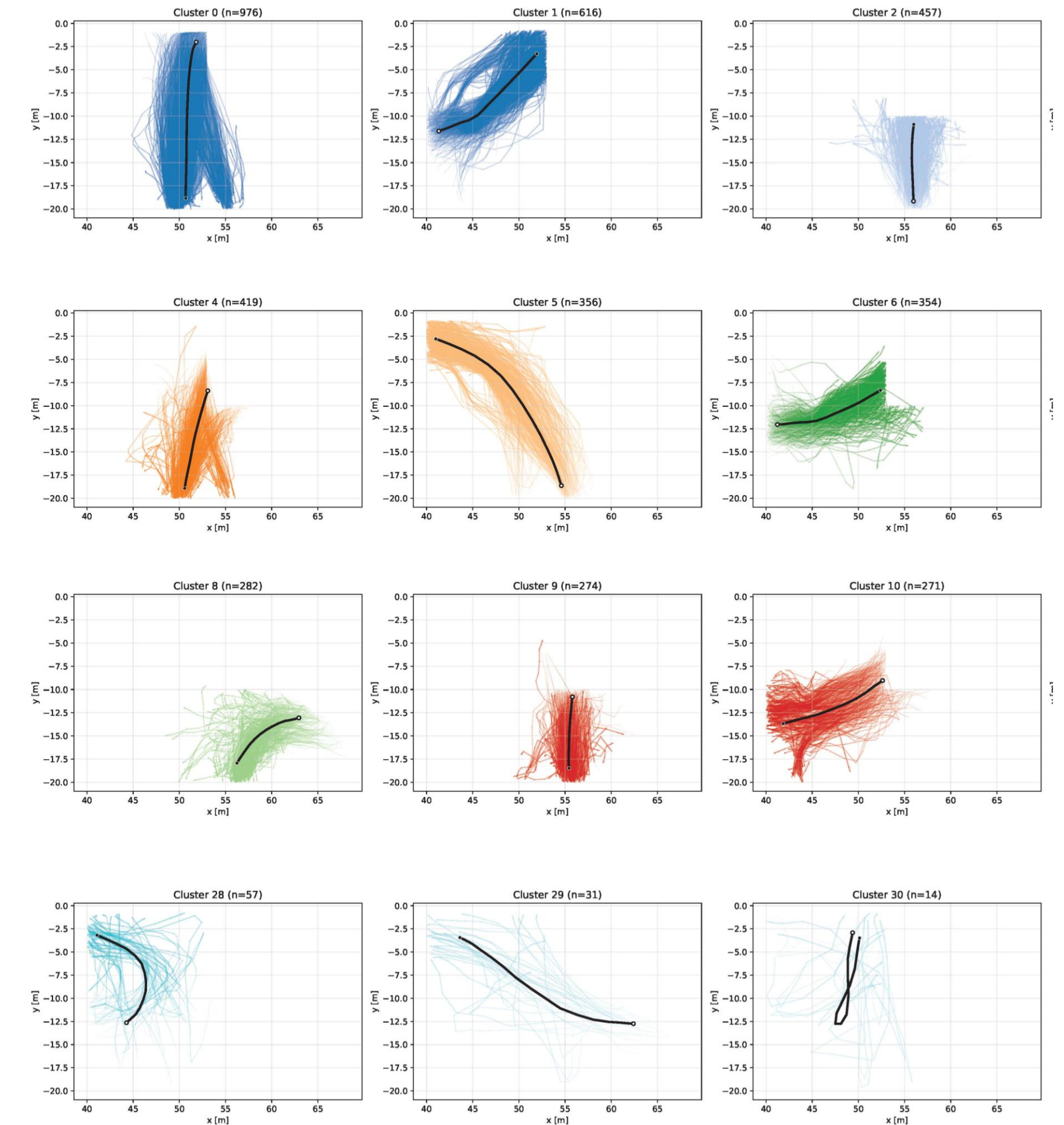
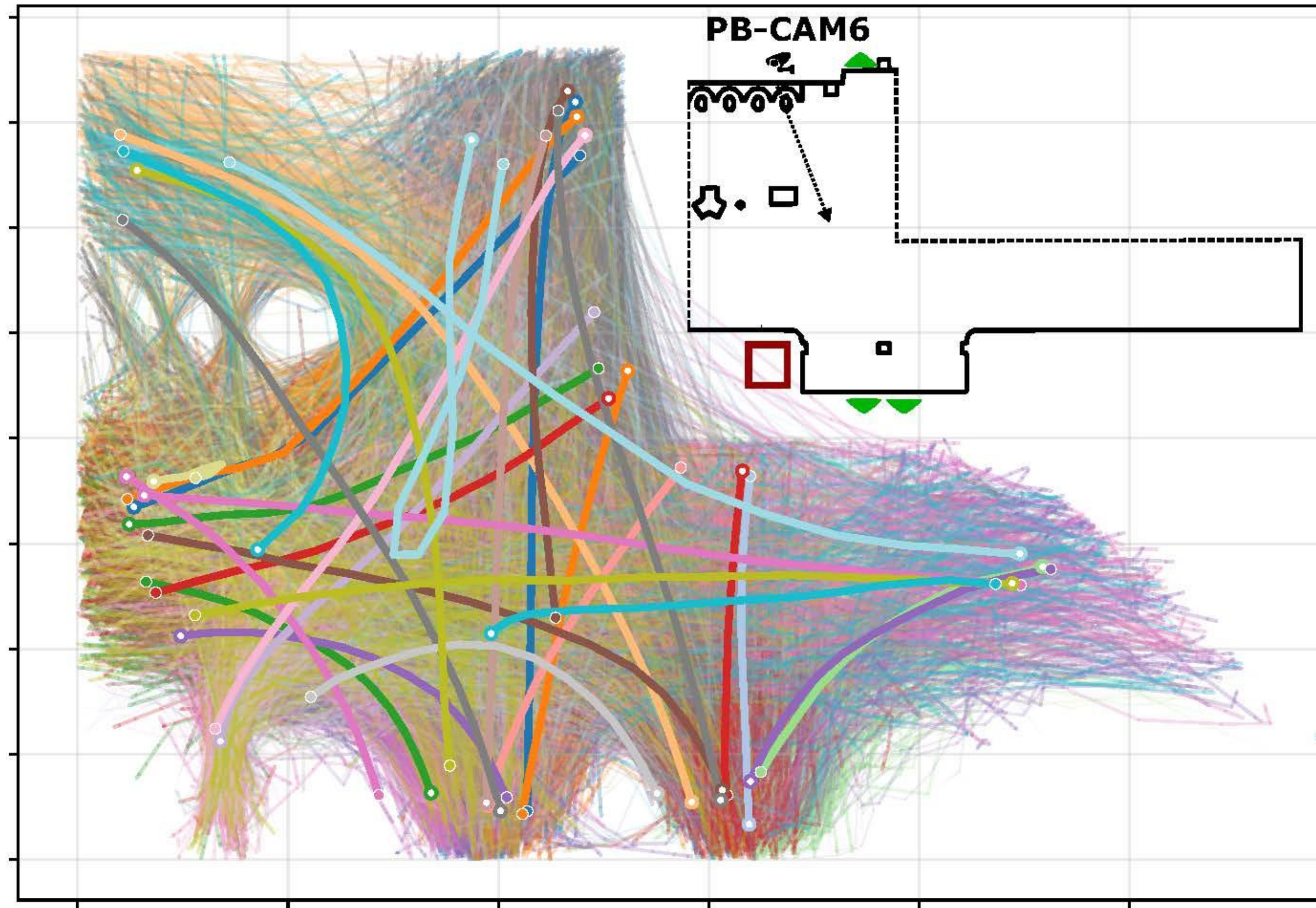
$$\pi^{(j)} = \beta, \quad s_n \sim \beta,$$

so, assignments are exchangeable rather than Markovian. In addition, we can fix the dynamical $\mathbf{A}^{(k)} = \mathbf{I}, \mathbf{C}^{(k)} = \mathbf{I}$. Observations are then emitted directly from the representative GP:

$$\mathbf{f}^{(k)}(\cdot) \sim \mathcal{GP}(m_{\theta^{(k)}}(\cdot), k_{\theta^{(k)}}(\cdot, \cdot))$$
$$\mathbf{y}_n(\tau) = \mathbf{f}^{(s_n)}(\tau) + \boldsymbol{\epsilon}_n(\tau).$$

Applications: pedestrian trajectory clustering

Clustering of PB-CAM6



Applications: pedestrian trajectory clustering

Discussion and future work

Assumptions – model decisions.

No route ground truth, exploratory analysis.

Longitudinal study – compare patterns between days.

Potential order and dynamical clustering.

Conclusions

Conclusions

- ▶ A domain-agnostic generative model was proposed covering each of the motivation challenges
- ▶ Temporal structure becomes scientific discoveries

Conclusions

- ▶ **A domain-agnostic generative model was proposed covering each of the motivation challenges**
- ▶ Temporal structure becomes scientific discoveries.

Conclusions

- ▼ **A domain-agnostic generative model was proposed covering each of the motivation challenges**

Unknown number of regimes

Evolving morphologies

Uncertainty-calibrated functional representations

Explicit temporal alignment

- **Temporal structure becomes scientific discoveries**

Conclusions

- ▶ A domain-agnostic generative model was proposed covering each of the motivation challenges
- ▶ **Temporal structure becomes scientific discoveries**

Conclusions

- ▶ A domain-agnostic generative model was proposed covering each of the motivation challenges

- ▼ **Temporal structure becomes scientific discoveries**

Breathing estimation: alignment as physiological signal

Takotsubo: GP prototypes as common ECG representation

Wave spectra: temporal regimes as event segmentation

Pedestrian paths: modularity of the model

Conclusions

Future work

Domain extensions and validation

Human-motion

Gene expression

Climatology

UCR: spectrometry, pedestrian distributions

Latent spaces exploration:

Does a deep temporal model learn an organized dynamical latent space?

Can we recover that organization?

Conclusions

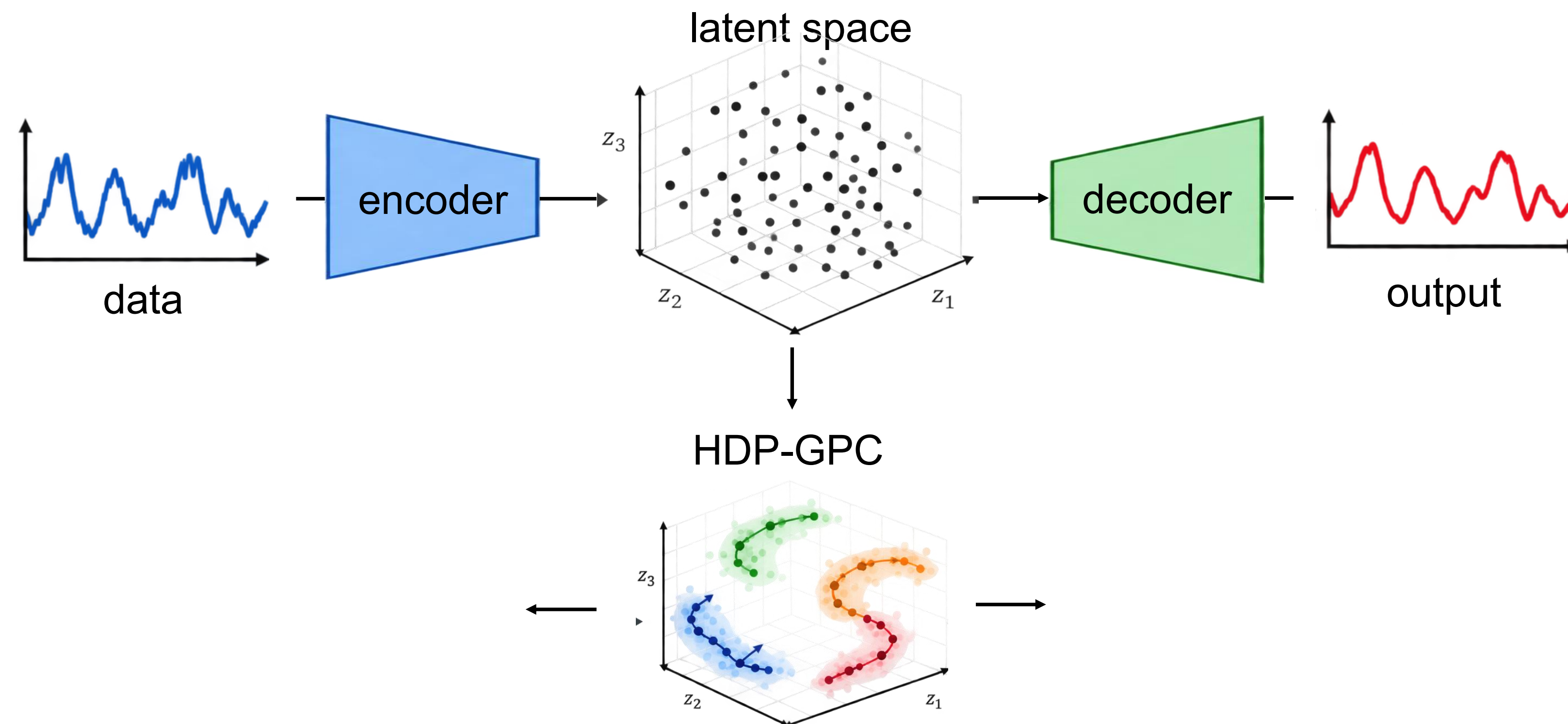
Future work

Latent spaces exploration:

Does a deep temporal model learn an organized dynamical latent space?

Can we recover that organization?

64



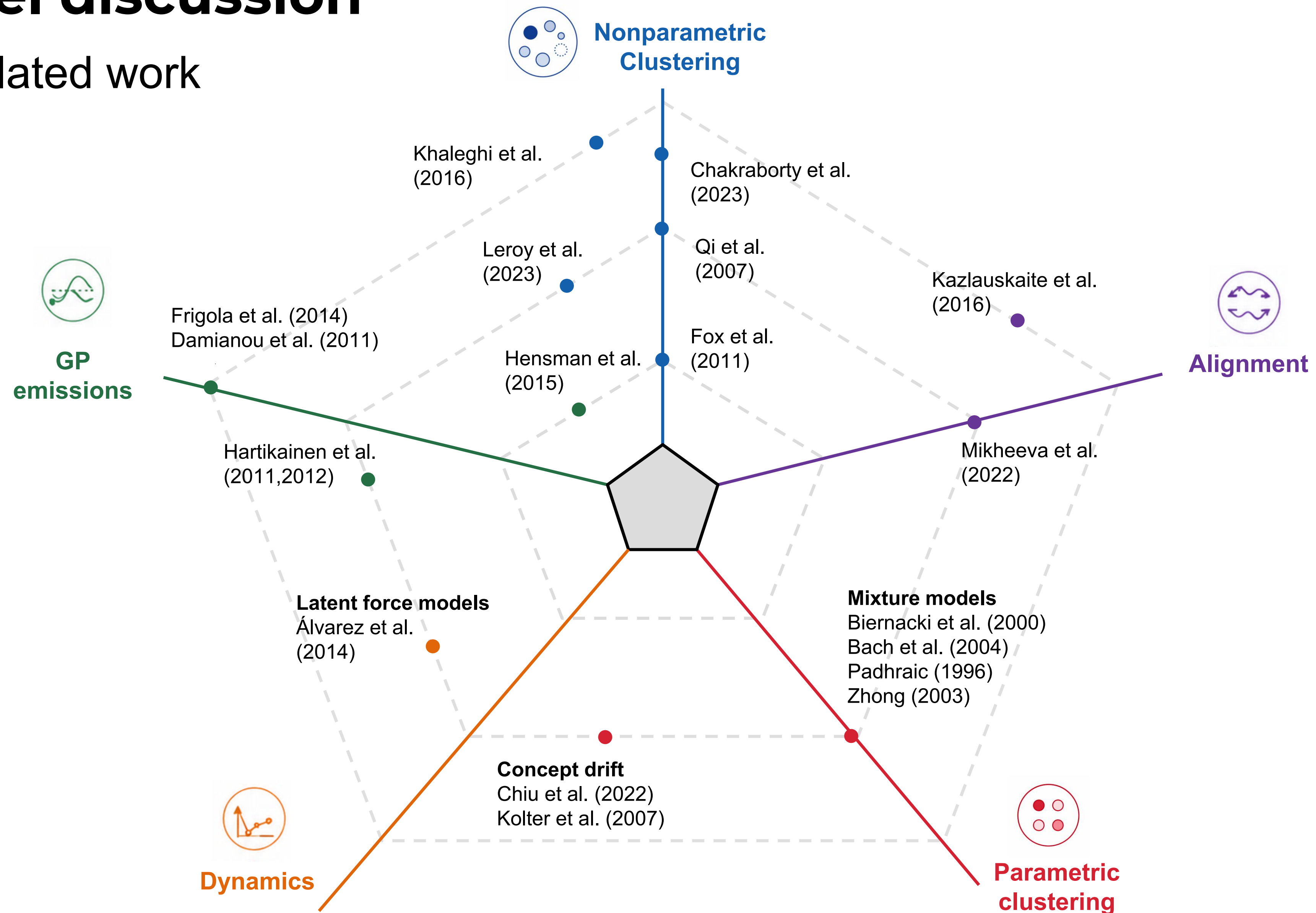
Conclusions

Code & contact

`https://github.com/AdrianPerezHerrero/HDP-GPC/`
`adrianperezherrero1997@outlook.es`

Model discussion

Related work



Conclusions

Clustering time-series helps understanding
complex systems, systems where
time itself matters





Gaussian Process Summer School

Manchester
14/09/2026

Bayesian non-parametric dynamical clustering of time series

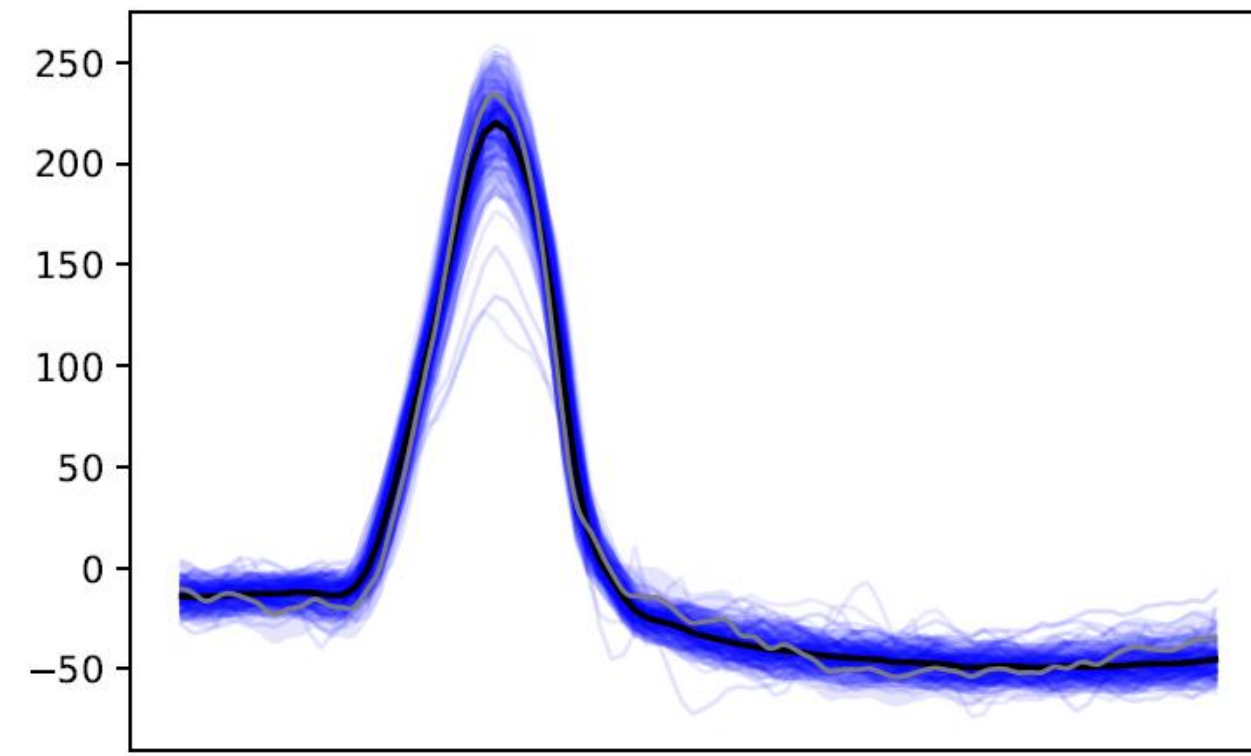
Adrián Pérez Herrero



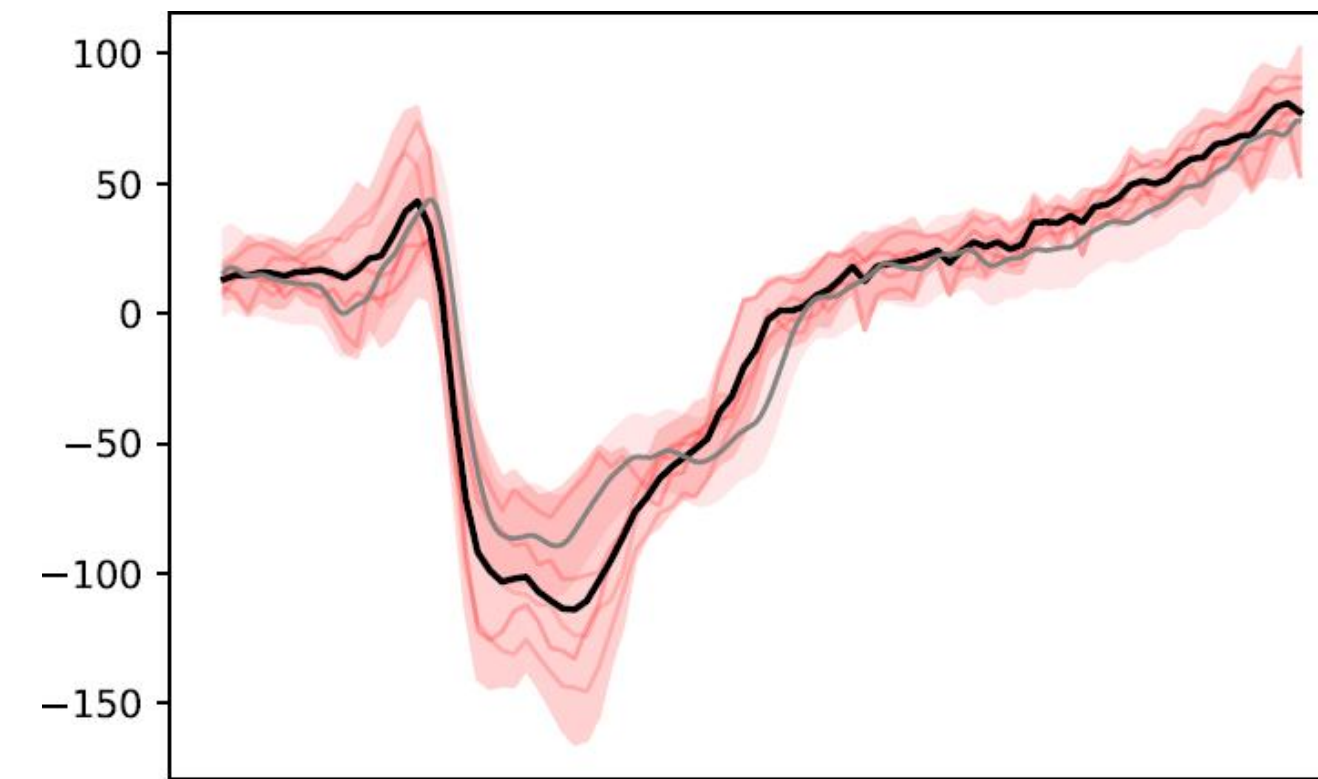
Appendix

Sampling

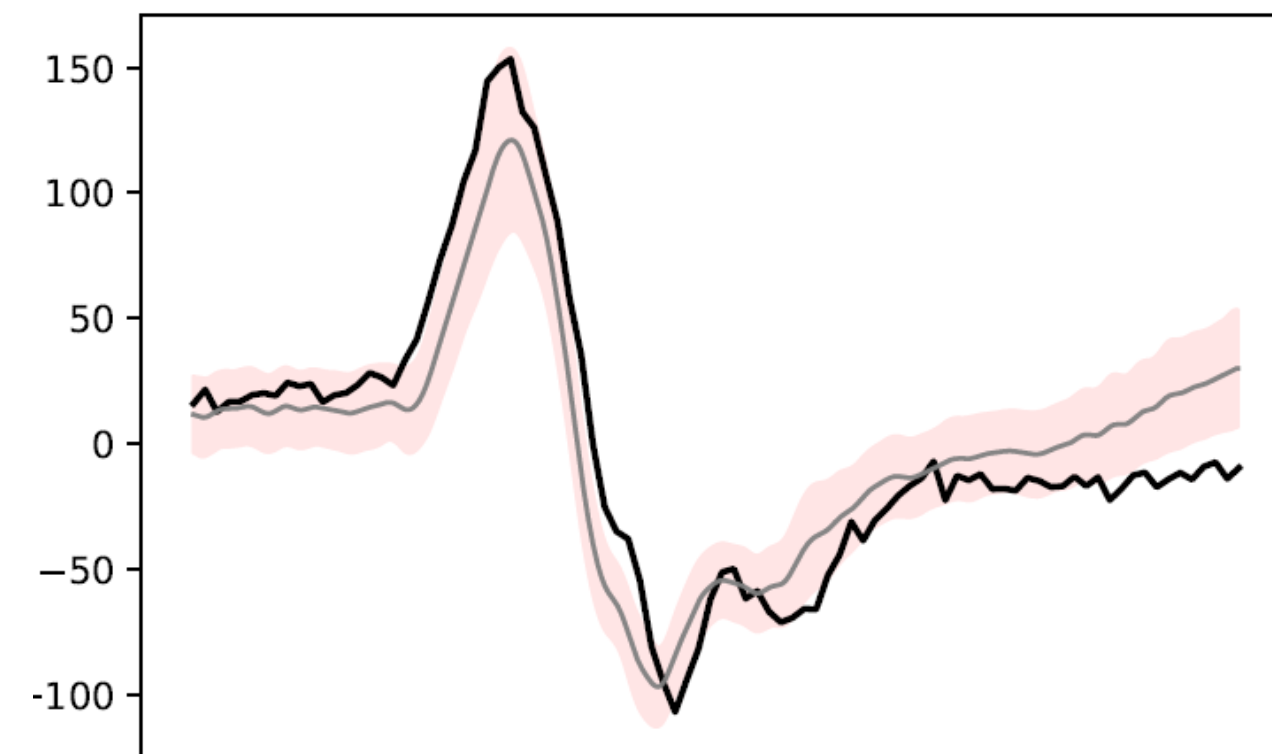
Cluster 1



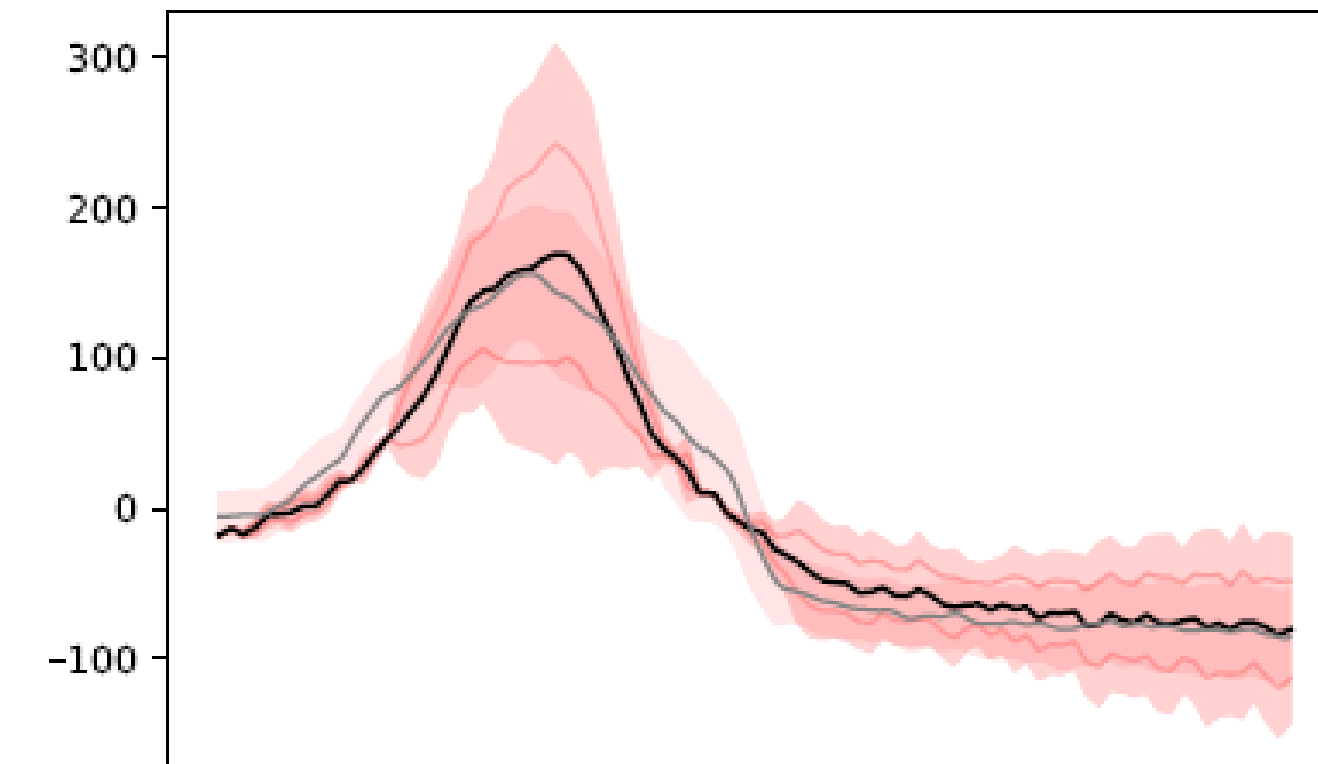
Cluster 3



Cluster 4



Cluster 8



Appendix. Group results MIT-BIH.

Record	[22]	N_{MO}	N_{OFF}	N_{ON}	Record	[22]	N_{MO}	N_{OFF}	N_{ON}
100	4	2	3	2	201	15	10	6	10
101	4	2	3	2	202	9	7	8	6
102	10	10	9	8	203	33	7	12	38
103	10	4	5	2	205	14	10	7	9
104	16	15	10	27	207	61	9	15	36
105	10	4	10	50	208	28	17	11	30
106	27	4	8	11	209	10	6	9	6
107	11	6	3	11	210	27	9	9	11
108	22	2	8	9	212	5	3	4	5
109	13	8	9	11	213	17	15	12	23
111	8	2	5	5	214	21	12	12	18
112	4	3	2	2	215	16	16	7	10
113	5	3	3	2	217	28	11	10	22
114	8	4	4	5	219	14	12	10	10
115	11	3	4	3	220	2	3	7	3
116	10	7	7	9	221	14	7	9	6
117	4	4	3	3	222	8	4	3	3
118	3	5	6	12	223	23	11	10	20
119	6	3	4	5	228	14	14	10	16
121	5	4	5	6	230	3	4	3	6
122	1	1	2	2	231	5	4	7	6
123	3	3	3	2	232	4	2	3	3
124	14	9	7	7	233	24	21	10	20
200	20	16	13	22	234	5	4	5	3
					TOTAL	629	342	335	539

MIT-BIH Arrhythmia Database

BASELINE:

[22] Castro D et al., A Method for Context-Based Adaptive QRS Clustering in Real Time. IEEE J Biomed Health Inform. 2015

Appendix. Performance results MIT-BIH.

Record	Beats	Clusters	Purity	NMI	ARI	Homo.	Comp.	M-F1
100	2272	2	0.985	0.092	0.056	0.048	1.000	0.664
101	1865	2	0.998	0.584	0.571	0.413	1.000	0.666
102	2187	10	0.991	0.452	0.328	0.872	0.305	0.932
103	2084	4	0.999	0.000	-0.002	0.002	0.000	0.500
104	2229	15	0.918	0.496	0.335	0.785	0.363	0.497
105	2572	4	0.986	0.496	0.661	0.619	0.413	0.566
106	2027	4	0.992	0.768	0.891	0.919	0.660	0.989
107	2137	6	0.997	0.236	0.095	0.850	0.137	0.972
108	1763	2	0.993	0.616	0.759	0.581	0.656	0.359
109	2532	8	0.998	0.224	0.155	0.903	0.128	0.653
111	2124	2	1.000	1.000	1.000	1.000	1.000	1.000
112	2539	3	0.999	0.000	-0.001	0.009	0.000	0.500
113	1794	3	0.997	0.016	0.004	0.254	0.008	0.721
114	1879	4	0.991	0.663	0.725	0.694	0.635	0.388
115	1952	3	1.000	0.000	0.000	1.000	0.000	1.000
116	2411	7	0.999	0.356	0.181	0.965	0.218	0.665
117	1535	4	0.999	0.001	-0.000	0.109	0.001	0.500
118	2278	5	0.958	0.311	0.247	0.196	0.747	0.659
119	1987	3	1.000	0.997	1.000	1.000	0.993	1.000
121	1863	4	0.999	0.091	0.057	0.501	0.050	0.667
122	2476	1	1.000	1.000	1.000	1.000	1.000	1.000
123	1518	3	1.000	0.089	0.041	1.000	0.047	1.000
124	1619	9	0.975	0.306	0.187	0.542	0.213	0.409
200	2601	16	0.978	0.674	0.814	0.836	0.565	0.492

201	1963	10	0.974	0.551	0.351	0.776	0.428	0.533
202	2136	7	0.982	0.246	0.134	0.510	0.162	0.781
203	2980	7	0.970	0.413	0.366	0.717	0.290	0.378
205	2656	10	0.995	0.716	0.824	0.908	0.592	0.664
207	2332	9	0.897	0.642	0.740	0.760	0.556	0.686
208	2955	17	0.969	0.711	0.799	0.908	0.585	0.703
209	3005	6	0.891	0.167	0.206	0.108	0.377	0.734
210	2649	9	0.981	0.561	0.670	0.769	0.441	0.384
212	2747	3	1.000	0.900	0.908	1.000	0.818	1.000
213	3250	15	0.961	0.498	0.472	0.745	0.375	0.680
214	2262	12	0.996	0.481	0.355	0.932	0.325	0.496
215	3363	16	0.999	0.786	0.973	0.968	0.661	0.665
217	2208	11	0.984	0.709	0.603	0.933	0.571	0.960
219	2154	12	0.996	0.403	0.334	0.854	0.263	0.623
220	2048	3	0.954	0.007	0.045	0.012	0.005	0.488
221	2427	7	1.000	0.835	0.962	1.000	0.716	1.000
222	2483	4	0.830	0.005	-0.017	0.004	0.007	0.227
223	2605	11	0.952	0.510	0.592	0.676	0.409	0.323
228	2053	14	0.999	0.425	0.196	0.978	0.271	0.666
230	2256	4	1.000	0.007	0.001	1.000	0.004	1.000
231	1571	4	0.999	0.904	0.926	0.982	0.838	0.749
232	1780	2	0.776	0.002	-0.002	0.001	0.034	0.291
233	3078	21	0.994	0.618	0.786	0.961	0.455	0.534
234	2753	4	0.987	0.404	0.519	0.319	0.551	0.829
Total/mean	109958	342	0.976	0.437	0.434	0.665	0.414	0.671
Global	109958	342	0.975	0.392	0.028	0.928	0.249	0.636

Appendix. Confusion matrix MIT-BIH.

	N	L	R	a	V	F	J	A	S	E	j	/	e	f	Q	!
N	74852	0	1	39	128	62	36	895	2	1	223	0	16	144	11	0
L	0	8075	0	0	18	3	0	1	0	1	0	0	0	0	2	18
R	0	0	6858	0	3	2	29	99	0	3	4	0	0	0	0	24
a	1	0	0	107	0	0	0	1	0	0	0	0	0	0	0	0
V	86	0	0	4	6824	74	0	2	0	0	1	0	0	1	4	0
F	68	0	0	0	46	662	0	1	0	0	0	0	0	0	0	0
J	5	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0
A	2	0	397	0	28	0	0	1546	0	0	1	0	0	0	0	38
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	16	0	0	0	0	94	0	0	0	0	0	41
j	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
/	0	0	0	0	6	0	0	0	0	0	0	7021	0	16	0	0
e	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
f	32	0	0	0	11	0	0	0	0	0	0	7	0	821	13	0
Q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0
!	0	0	3	0	49	0	0	0	0	7	0	0	0	0	0	351

Appendix. Variational Inference. Offline.

Algorithm 1 Offline variational inference

Input: A dataset $\mathcal{Y} = \{(\mathbf{t}_n, \mathbf{y}_n)\}_{n=1}^N$ of time sequences

Input: A model $p(\mathcal{M}, \mathcal{Y})$

Output: The variational density $q(\mathcal{M}|\mathcal{Y})$

```
1: procedure OFFLINEVARIATIONALINFERENCE( $\mathcal{Y}, \mathcal{M}$ )
2:   while  $\mathcal{L}$  has not converged do
3:     while  $\mathcal{L}_{OBS}$  has not converged do
4:       for all  $k = 1, 2, \dots$  do //local update
5:         for all  $n = 1, \dots, N$  do
6:           compute  $q(s_{nk})$  using eq. (3.14)
7:         end for
8:         for all  $m = 1, \dots, n_k$  do
9:           compute  $q(\mathbf{g}_m^k)$  numerically using eq. (3.22)
10:          compute  $q(\mathbf{f}_m^k)$  and  $q(\mathbf{x}_m^k)$  using eq. (3.31)
11:        end for
12:        compute  $q(\phi^k)$  using eq. (3.26)
13:      end for
14:    end while
15:    for all  $k = 1, 2, \dots$  do //global update
16:      compute  $q(\pi_k)$  using eq. (3.19)
17:      compute  $q(v_k)$  numerically using eq. (3.38)
18:    end for
19:    compute  $\mathcal{L} = \mathcal{L}_{OBS} + \mathcal{L}_{HDP} + \mathcal{L}_W + \mathbb{H}[q(S)]$ 
20:  end while
21:  return  $q(\mathcal{M}|\mathcal{Y})$ 
22: end procedure
```

Appendix. Variational Inference. Multi-output.

This observation contribution allows us to run the inference independently for each output channel but keeping a global assignment s_n and just iterating over leads

Algorithm 3 Off-line multi-output variational inference

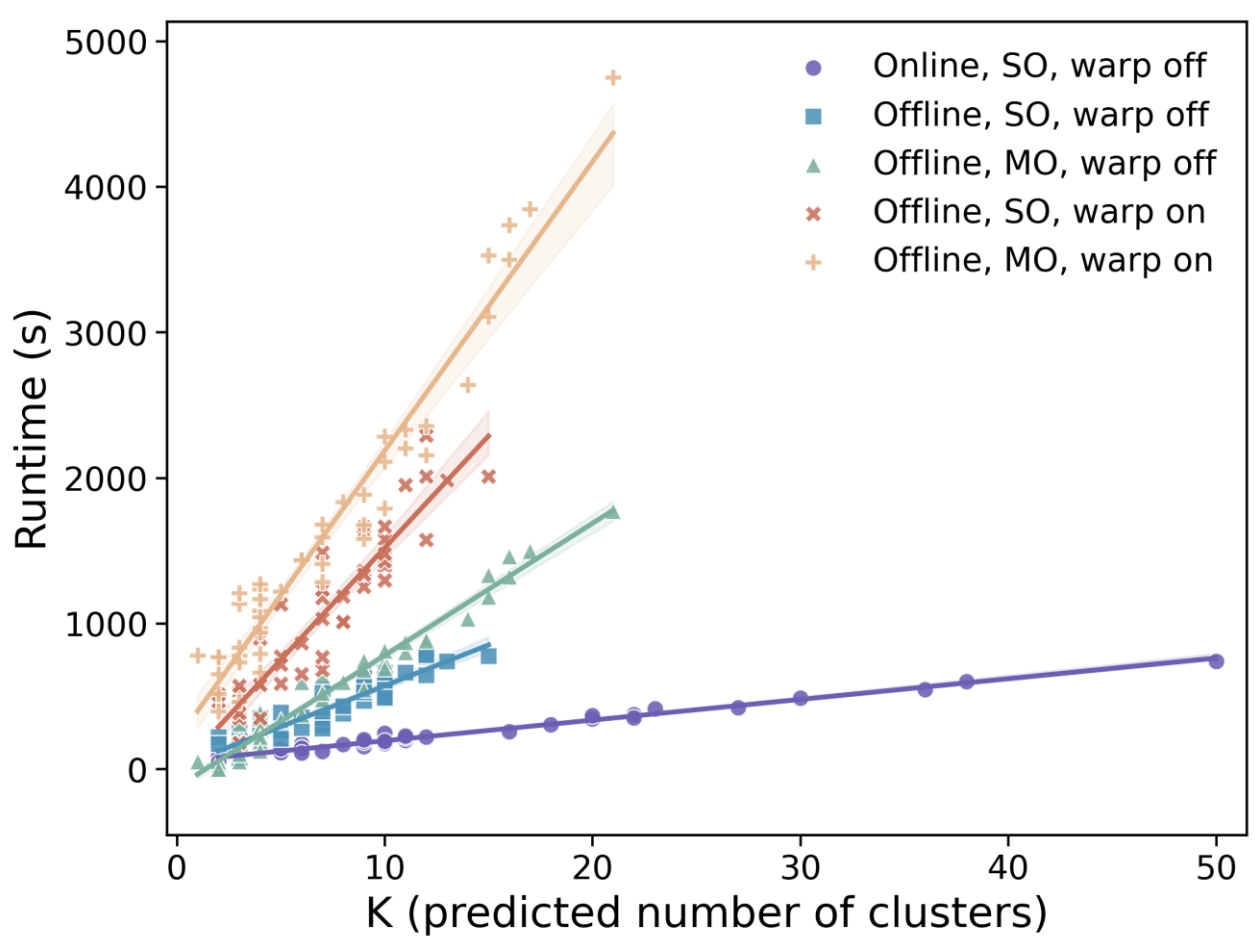
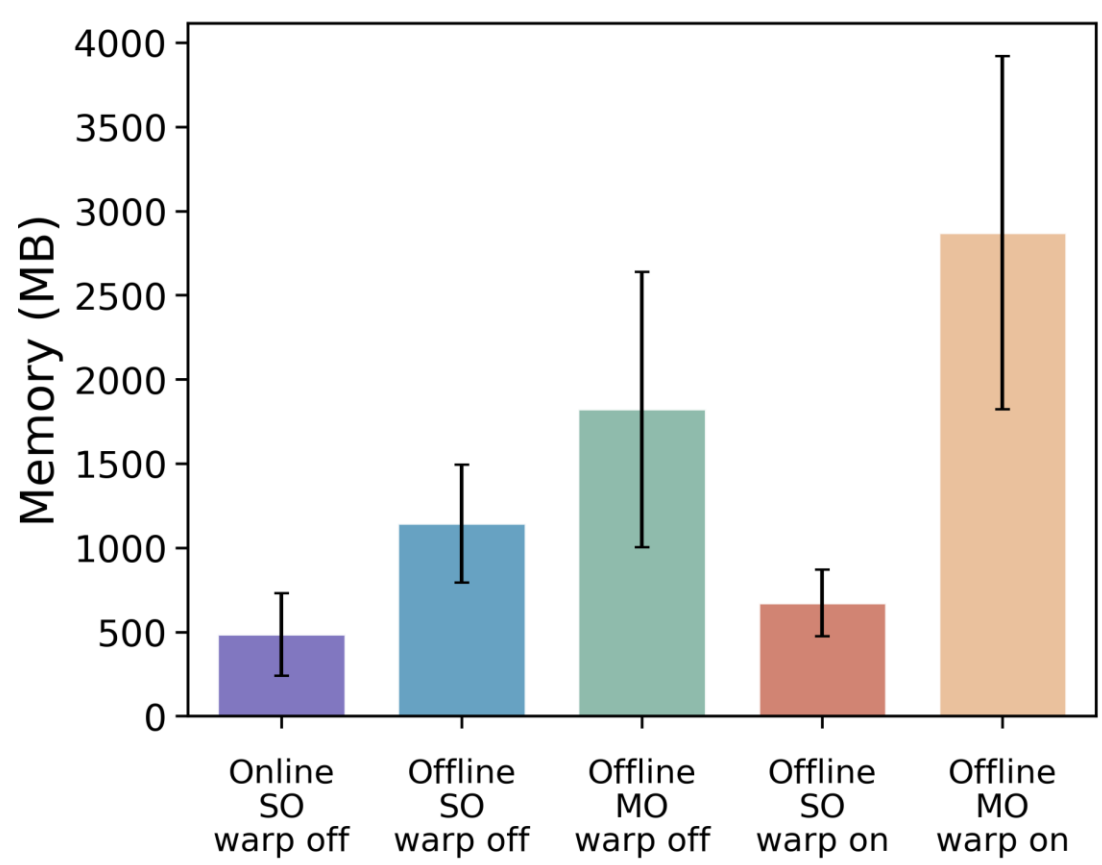
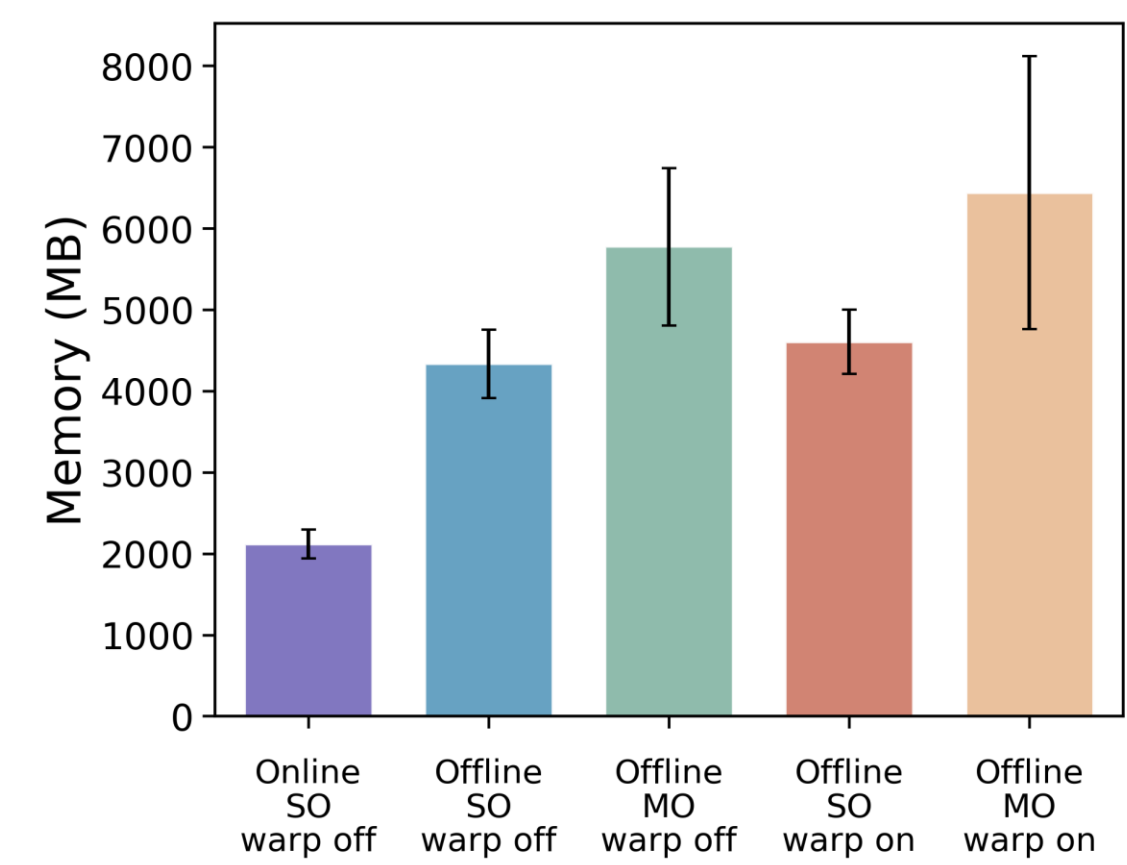
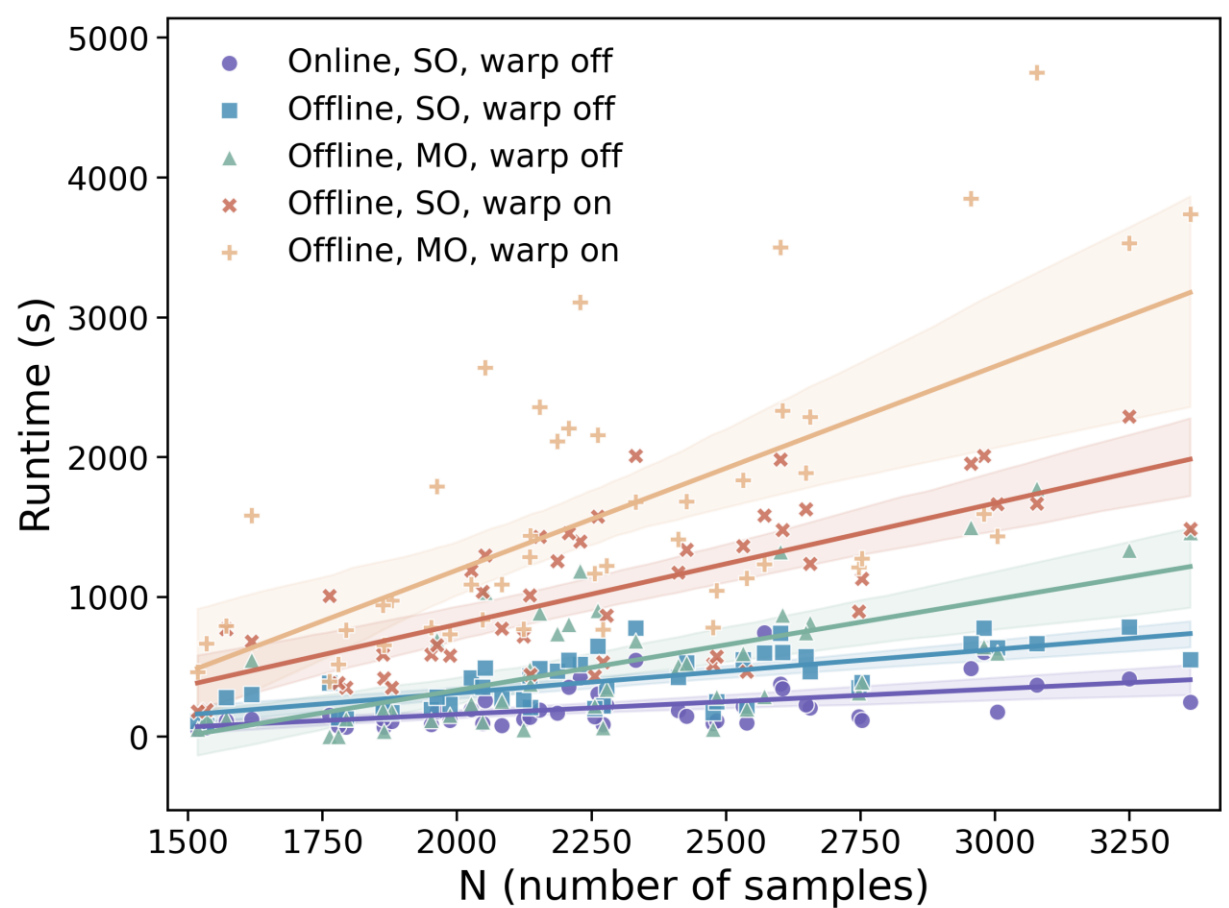
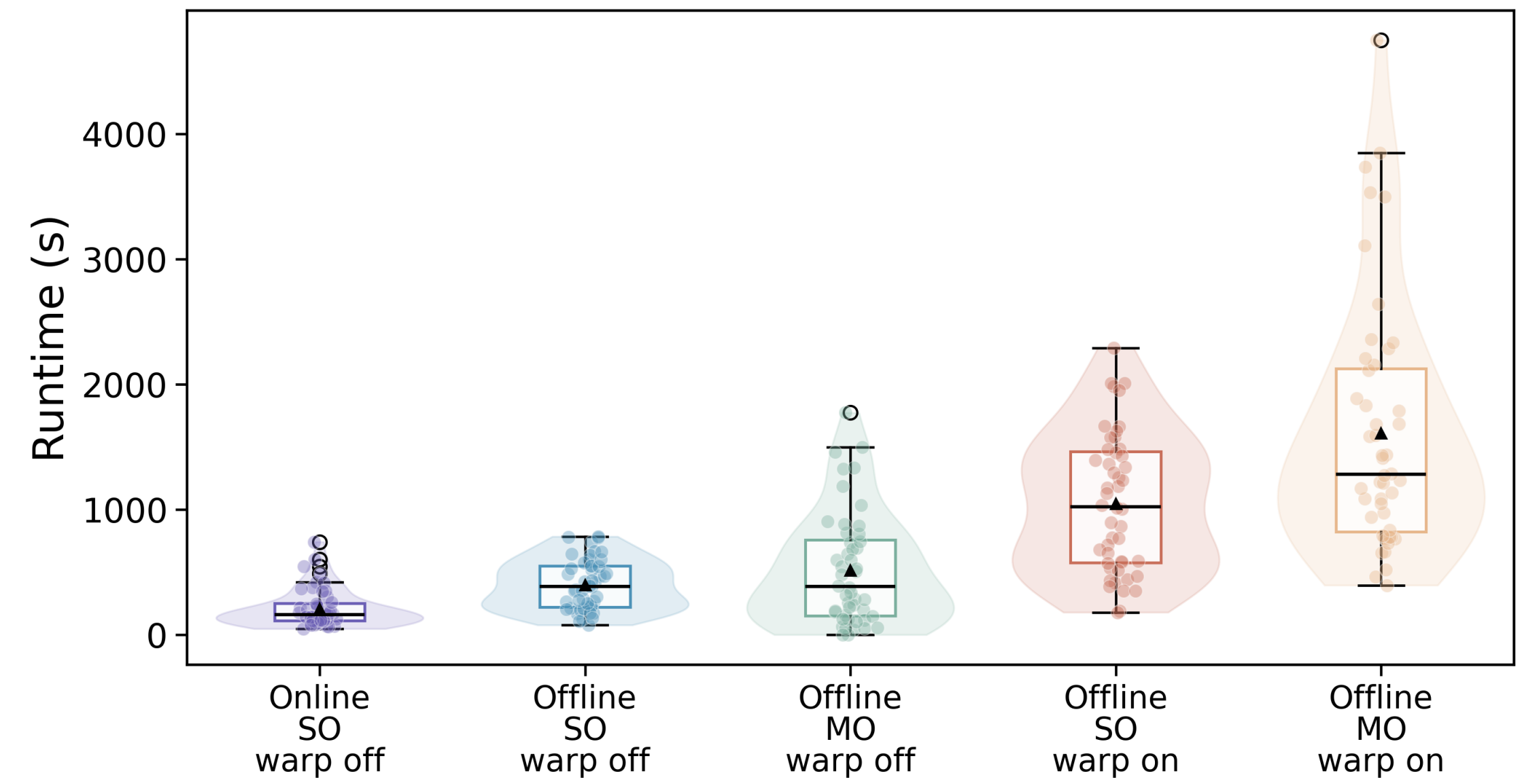
Input: A dataset $\mathcal{Y} = \{(t_n, y_n)\}_{n=1}^N$ of time sequences

Input: A model $p(\mathcal{M}, \mathcal{Y})$

Output: The variational density $q(\mathcal{M}|\mathcal{Y})$

```
1: procedure OFFLINEMOVARIATIONALINFERENCE( $\mathcal{Y}, \mathcal{M}$ )
2:   while  $\mathcal{L}$  has not converged do
3:     while  $\mathcal{L}_{OBS}$  has not converged do
4:       for all  $l = 1, 2, \dots, L$  do
5:         for all  $k = 1, 2, \dots$  do //local update
6:           for all  $n = 1, \dots, N$  do
7:             compute  $q(s_{nk})$  using eq. (3.14)
8:           end for
9:           for all  $m = 1, \dots, n_k$  do
10:            compute  $q(\mathbf{g}_m^k)$  numerically using eq. (3.22)
11:            compute  $q(\mathbf{f}_m^k)$  and  $q(\mathbf{x}_m^k)$  using eq. (3.31)
12:          end for
13:          compute  $q(\phi^k)$  using eq. (3.26)
14:        end for
15:      end for
16:    end while
17:    for all  $l = 1, 2, \dots, L$  do
18:      for all  $k = 1, 2, \dots$  do //global update
19:        compute  $q(\pi_k)$  using eq. (3.19)
20:        compute  $q(v_k)$  numerically using eq. (3.38)
21:      end for
22:    end for
23:    compute  $\mathcal{L} = \mathcal{L}_{OBS} + \mathcal{L}_{HDP} + \mathcal{L}_W + \mathbb{H}[q(S)]$ 
24:  end while
25:  return  $q(\mathcal{M}|\mathcal{Y})$ 
26: end procedure
```

Appendix. Empirical complexity.



Appendix. Takotsubo ablation

Results takotsubo identification, sex ablation.

Metric	TTS vs suspected MI	TTS vs STEMI	TTS vs NSTEMI	TTS vs verified MI	apical TTS vs anterior STEMI
ROC AUC (Sex)	0.80	0.83	0.81	0.82	0.82
ROC AUC (ECG)	0.77	0.85	0.83	0.83	0.76
ROC AUC (All)	0.89	0.94	0.91	0.92	0.90
PR AUC (Sex)	0.11	0.49	0.32	0.25	0.48
PR AUC (ECG)	0.14	0.68	0.42	0.34	0.48
PR AUC (All)	0.29	0.84	0.59	0.53	0.71

Appendix. Takotsubo ablation

Results takotsubo identification, mean ablation.

Metric	TTS vs suspected MI	TTS vs STEMI	TTS vs NSTEMI	TTS vs verified MI	Apical TTS vs ant. STEMI
ROC AUC (Std)	0.84	0.86	0.83	0.83	0.84
ROC AUC (Bayes)	0.86 (+1.9%)	0.88 (+3.3%)	0.83 (+0.1%)	0.83 (+0.0%)	0.88 (+4.4%)
PR AUC (Std)	0.17	0.68	0.40	0.34	0.51
PR AUC (Bayes)	0.19 (+12.9%)	0.78 (+14.5%)	0.45 (+10.1%)	0.36 (+6.2%)	0.61 (+19.9%)
Sens (Std)	0.86	0.86	0.88	0.82	0.86
Sens (Bayes)	0.86	0.86	0.88	0.82	0.86
Spec (Std)	0.72	0.72	0.59	0.75	0.78
Spec (Bayes)	0.75 (+4.3%)	0.82 (+14.9%)	0.69 (+17.1%)	0.81 (+7.2%)	0.82 (+5.8%)
Acc (Std)	0.72	0.76	0.63	0.76	0.79
Acc (Bayes)	0.75 (+4.0%)	0.83 (+10.3%)	0.71 (+14.2%)	0.81 (+6.3%)	0.83 (+4.3%)
F1 (Std)	0.18	0.66	0.36	0.39	0.59
F1 (Bayes)	0.19 (+9.7%)	0.74 (+12.4%)	0.43 (+18.0%)	0.44 (+14.0%)	0.64 (+8.6%)
PPV (Std)	0.10	0.53	0.23	0.25	0.45
PPV (Bayes)	0.11 (+11.2%)	0.65 (+21.5%)	0.28 (+23.7%)	0.30 (+19.1%)	0.51 (+13.1%)
NPV (Std)	0.99	0.93	0.97	0.98	0.96
NPV (Bayes)	0.99 (+0.1%)	0.94 (+0.9%)	0.98 (+0.4%)	0.98 (+0.1%)	0.97 (+0.6%)