Eliciting Preferences in Music

Jan Larsen
Cognitive Systems Section
Dept. of Informatics and Mathematical Modelling
Technical University of Denmark
jl@imm.dtu.dk, www.imm.dtu.dk/~jl

Preference elicitation refers to the problem of developing a decision support system capable of generating recommendations to a user, thus assisting him in decision making. It is important for such a system to model user's preferences accurately, find hidden preferences and avoid redundancy. This problem is sometimes studied as a **computational learning theory** problem

Ref: Wikipedia

Collaborative filtering is a method of making automatic predictions (filtering) about the interests of a user by collecting preferences or taste information from **many users** (collaborating).

Ref: Wikipedia

Main assumption

User preference recorded from behavior and interactions is a proxy for aspects of human cognition

Cognitive Systems - a vision for the future

An artificial cognitive system is the ***ultimate learning*** and thinking machine with ability to operate in ***open-ended environments*** with ***natural interaction*** with humans and other artificial cognitive systems and plays key role in the transformational society in order to achieve augmented *capabilities beyond* human and existing machines

J. Larsen: ”Cognitive Systems,” tutorial presented at IEEE Machine Learning for Signal Processing Workhsop, Cancun, Mexico, 2008. http://www2.immm.dtu.dk/pubdb/p.php?5705, http://www2.immm.dtu.dk/pubdb/p.php?5766

Cognitive systems

- Why: goals
- How: data, processing
- What: capabilities

How much is needed to qualify the system as being cognitive?

A tiered approach: from low to high-level capabilities

What - capabilities

Natural interaction

- Mediation and ontology alignment
- Handling of ambiguity, conflicts, uncertainties
- Communication
- Multi-goal achievement
- Locomotion and other physical actions

High-level emergent properties (strong AI)

- Consciousness
- Self-awareness
- Sentience (feeling)
- Empathy
- Emotion
- Intuition

COSOUND

A COGNITIVE SYSTEMS APPROACH TO ENRICHED AND ACTIONABLE INFORMATION FROM AUDIO STREAMS

DTU

DR

Syntonetic

Musikzonen

Geckon

UCL

Royal School of Library and Information Science

Hindenburg Systems

B&O

Queen Mary University of London

Aalborg University

Copenhagen University

University of Glasgow

State and University Library

Vision

The overall vision is to foster truly participatory, collaborative, and cross-cultural tools for enrichment of audio streams which can improve interactivity, findability, experienced quality, ability to co-create, and boost productivity in a broad sense.

users in the loop framework – required to study and evaluate interactive and participative (crowd) designs

CoSound Hypothesis

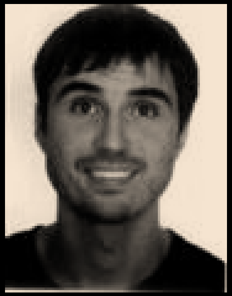
The main hypothesis is that the integration of bottom-up data derived from audio streams and top-down data streams from users can enable actionable cognitive representations, which will positively impact and enrich user interaction with massive audio archives, as well as facilitating new commercial success.

We will test the hypothesis at three different functionality levels: 1) personalized audio streams; 2) task driven navigation and organization; 3) sharing of enriched audio streams through editing and co-creation.

Acknowledgments and references

Bjørn Sand Jensen

Jens Brehm Nielsen



Jens Madsen

- B.S. Jensen, J.B. Nielsen, J. Larsen: "Efficient Preference Learning with Pairwise Continuous Obervations and Guassian Processes," IEEE Workshop on Machine Learning for Signal Processing (MLSP2011), Beijing, China, September 2011.
- B.S. Jensen, J.S. Gallego, J. Larsen: " A Predictive Model of Music Preference Using Pairwise Comparisons," IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP2012), Kyoto, Japan, March 2012.
- J. Madsen, J.B. Nielsen, B.S. Jensen, J. Larsen: "Modeling Expressed Emotions in Music using Pairwise Comparisons," accepted for 9'th International Conference on Computer Music Modellng and Retrieval (CMMR2012), Queen Mary University, UK, June 2012.
- J.B. Nielsen, B.S. Jensen, J. Larsen: "On Sparse Multi-Task Gaussian Process Priors for Music Preference Learning," In NIPS CMPL workshop, pages 1-8, December 2011.
- B. S. Jensen and J. B. Nielsen: "Pairwise Judgments and Absolute Ratings with Gaussian Process Priors," Technical report, DTU November 2011.
- J.B. Nielsen, B.S. Jensen, J. Larsen: "Learning from Pairwise Observations with Gaussian Processes: Review and Extensions,'" in preparatiom for submission Spring 2012.

Motivation

- Detailed view of a subjects preference, e.g. music or audio clips with the purpose of recommendation or other MIR tasks.
- Expressed emotions in e.g. music
- Fitting/optimization of personalized hearing aids (and other common aids/gadgets TVs, HiFi etc)

Parwise

- Elicitation by pairwise comparisons eliminates the need for absolute references, and explanation of multiple dimensions – no why questions!
- Difficult to articulate experience/opinion
- Maybe issues related to learning from limited number of song/clips

Direct

- Elicitates a specific aspect
- Learning from few songs might
- Complex due to perceptual and cognitive processes
- Difficult to understand/explain scale
- Difficult to consistently rate music/settings/emotions on direct scales (dimensional or categorical)
    - communication biases due to uncertainties in scales, anchors or labels
    - lack of references causes drift and inconsistencies

The background: Weber's law

'Just noticable difference' is relative to stimuli strength

$$dp = k\ dS/S$$

Perception

Stimuli, e.g. weight

prop. constant

$$p = k\ \ln\left(\frac{S}{S_0}\right)$$

"Weber's Law", *Encyclopedia Americana*, 1920.

Pairwise comparison versus direct scaling

- Thurnstones ”Priciple of comparative judments”
  - ”The discrimal process” – the total process of discrimating stimuli
  - Assumptions
    1. preference (utility function, or in Thurstone's terminology, *discriminal process*) for each stimulus
    2. The stimulus whose value is larger at the moment of the comparison will be preferred by the subject
    3. These unobserved preferences are normally distributed in the population
- The “phsycological scale is at best an artificial construct” (Thurnstone)
- Lockhead claims that everything is relative……

G. R. Lockhead, “Absolute Judgments Are Relative: A Reinterpretation of Some Psychophysical Ideas.,” Review of General Psychology, vol. 8, no. 4, pp. 265–272, 2004.

L. L. Thurstone, “A law of comparative judgement.,” Psychological Review, vol. 34, 1927.

A. Maydeu-Olivares: ”On Thutstone’s Model For Paired Comparisons and Ranking Data”, Barcelona Univ.

The background

- Pairwise preference learning
  - Bayesian formulation with GP
  - Audio preference
  - Food preference
  - Semi-supervised version with active learning

- W. Chu and Z. Ghahramani: “Preference learning with Gaussian Processes,” ICML 2005 - Proceedings of the 22nd International Conference on Machine Learning, pp. 137–144, 2005.
- P. Groot, T. Heskes, T. Dijkstra, and J. Kates: “Prdicting preference judgments of individual normal and hearing-impaired listeners with Gaussian Processes,” IEEE Transactions on Audio, Sound, and Language Processing, 2010.
- E. Bonilla, S. Guo, and S. Sanner: “Gaussian Process preference elicitation,” in Advances in Neural Information Processing Systems 23, J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R.S. Zemel, and A. Culott, Eds., pp. 262–270. 2010.
- W. Chu and Z. Ghahramani: “Extensions of Gaussian Processes for ranking: semi-supervised and active learning,” in Workshop Learning to Rank at Advances in Neural Information Processing Systems 18, 2005.
- E. Bonilla, K. Ming and C Williams: “Multi-task Gaussian Process Prediction.” Advances in Neural Information Processing Systems 20, pp. 153–260, 2008.
- C. E. Rasmussen and C. K. I. Williams: Gaussian Processes for Machine Learning, MIT Press, 2006.

Outline

- Introduction
- Methods and models
  - Likelihood models
    - Probit/Logistic Choice
    - Confidence rating
    - Degree of Difference
  - Models of preference value
    - Generalized Linear Models (GLM)
    - Gaussian Process (GP) framework
      - The Basics
      - Sparse extensions
      - Multi-task / multi-subject
  - Experimental Design
    - Sequential design / active learning
- Applications
    - Music Preference
    - Emotion in Music
    - Optimization of Hearning Aids

Our Framework

Methodology

- Standard and new likelihoods for discrete / continuous ratings with pairwise for fast learning
- A Bayesian and non-parametric approach using Gaussian Process as the underlying regression model
- Support for various experimental paradigms: Two-alternative forced choice (2AFC), continuous paired ratings (and absolute scaling)
- Sequential design for (semi-) optimal experimental design (aka D-optimal designs, active learning)

Tools and documentation

- A Matlab Toolbox (available soon)
  - Many specialized likelihoods
  - Multi-task kernels
  - Generative kernels (e.g. Probability Product Kernel for GMMs, topic models, etc.)
  - Sequential Design / Active Learning
  - Inference
    - Laplace approximation for all likelihood functions
    - EP for some likelihood functions
    - MCMC for some likelihood functions

Methods and models

Notation

- A input instance (scalar, vector, matrix, string, distribution: x
- Full set of all input instansens: $\mathcal{X}$, $(\mathrm{i.e.}\, x \in \mathcal{X})$
- A response (discrete): $d_k\,(-1 \text{ or } +1)$
- A response (continuous, bounded): $\mathrm{y}_k \in [0;1]$
- A response (continuous, bounded): $\mathrm{z}_k \in [-1;1]$,
- A pairwise experiment: input one $\mathrm{u}_k \in \mathcal{X}$, input two $\mathrm{v}_k \in \mathcal{X}$ and response d_k and/or y_k
- Likelihood is a function of fbs difference of preference values for two instances:

$$\Delta \mathrm{f} = f(u_k) - f(v_k)$$

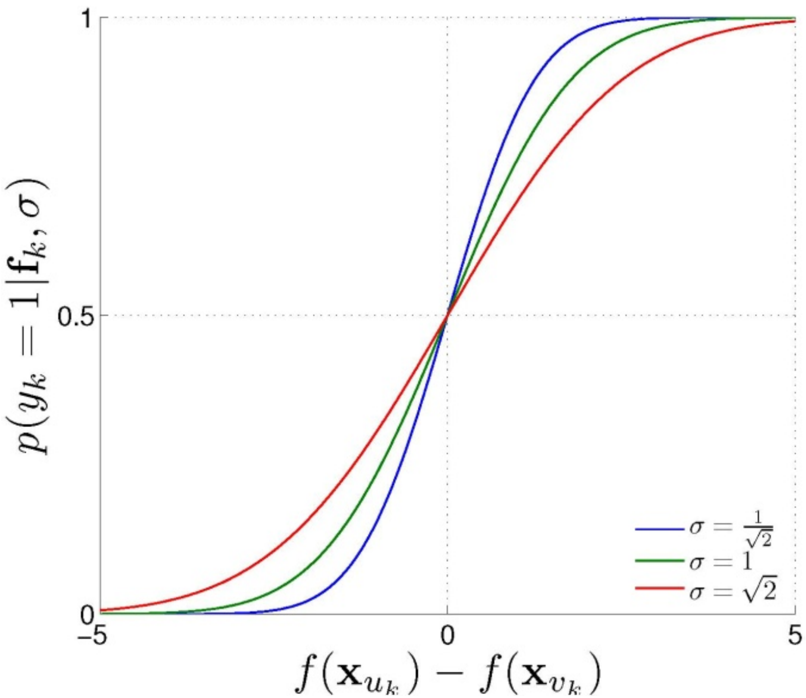
Likelihoods: Discrete

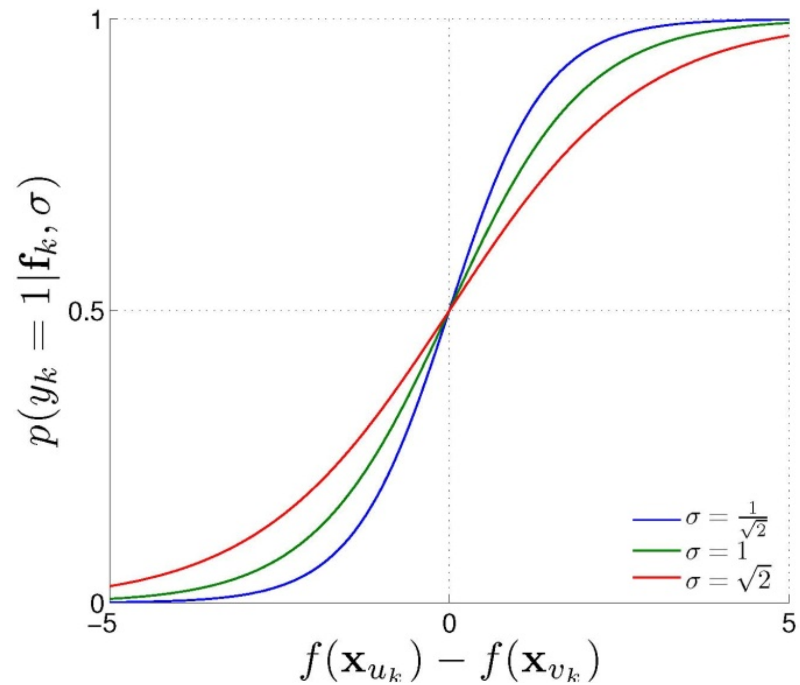
Discrete choice

Latent preference

Probit Choice $\mathcal{L}(\boldsymbol{f}_k, \sigma | u_k, v_k) = \mathrm{p}(d_\mathrm{k} | \boldsymbol{f}_k, \sigma) = \Phi\left(d_k \frac{f(u_k) - f(v_k)}{\sqrt{2}\sigma}\right)$

Logistic Choice $\mathcal{L}(\boldsymbol{f}_k, \sigma | u_k, vk) = \mathrm{p}(d_\mathrm{k} | \boldsymbol{f}_k, \sigma) = \dfrac{1}{1 + exp\left(d_k \frac{f(u_k) - f(v_k)}{\sigma}\right)}$





Three ways of interacting with users

- Discrete choice
- Discrete choice with confidence rating
- Degree of difference

Evaluating all $n^2/2$-n pairwise situations is infeasible (curse of dim.); hence, model and prior should caputure regularity

Discrete Choice with Confidence Rating

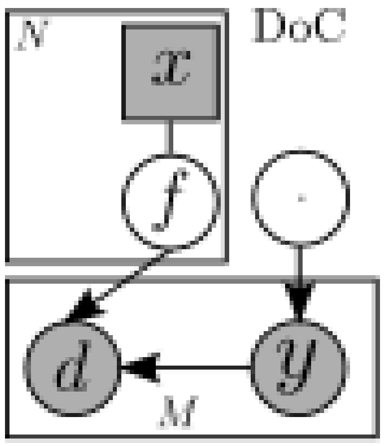
Objective: Obtain faster and robust learning by trusting observations where the subject is confidenet about the answer.

Case I: two ordered questions

1) A discrete choice, d_k
2) Followed by a confidence, $y_k \in [0; 1]$

Generic joint likelihood assuming that y does not depend on f:

$$\mathcal{L}(\mathbf{f}_k | d_k, y_k, u_k, v_k, \sigma) = p(d_k | \boldsymbol{f}_k, y_k) p(y_k | \cdot)$$



Discrete Choice with Confidence Rating

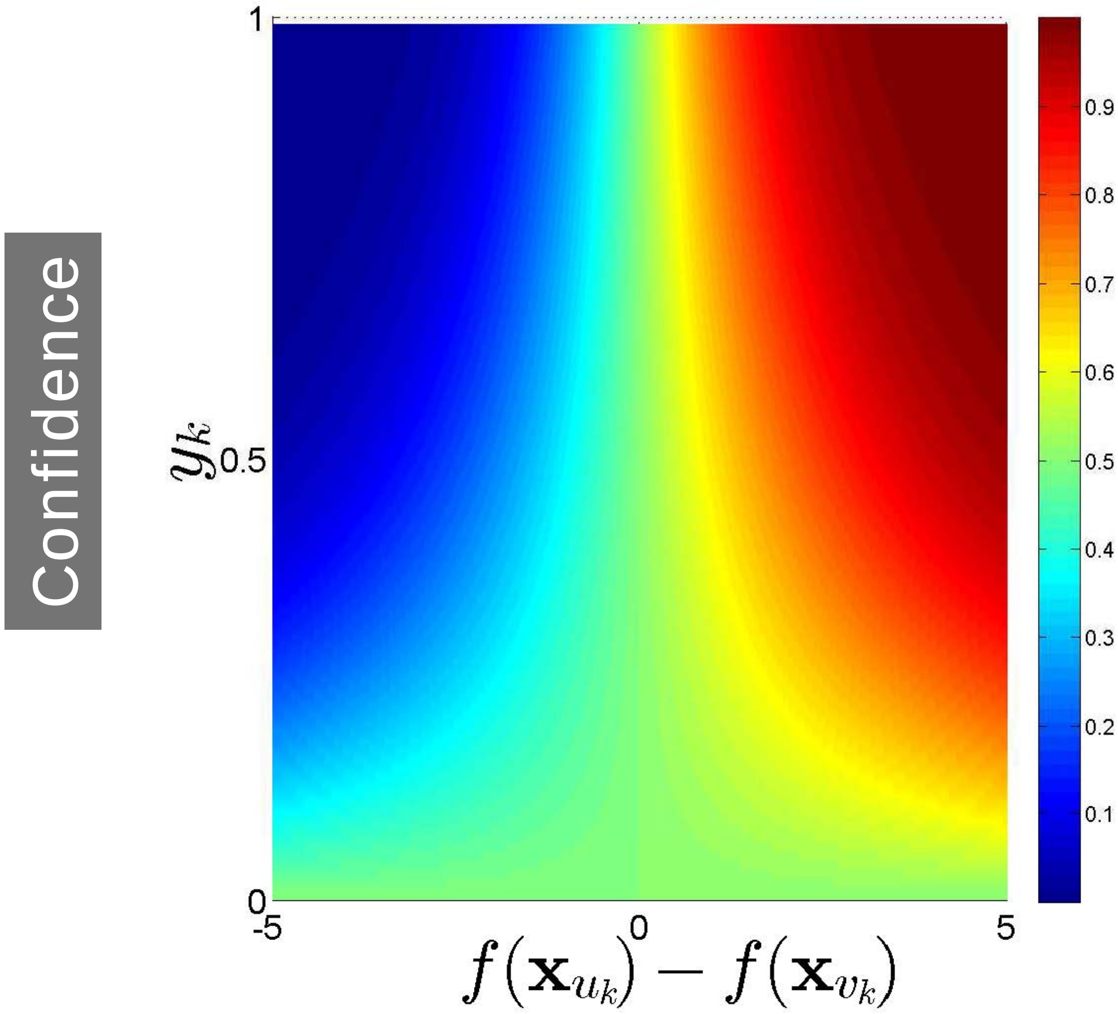
Case Ia - simple $p(y_k|\cdot)$:

- Uniform $p(y_k|\cdot) = 1$, joint equals the conditional
- Probit with $\sigma = \frac{1}{y_k}$

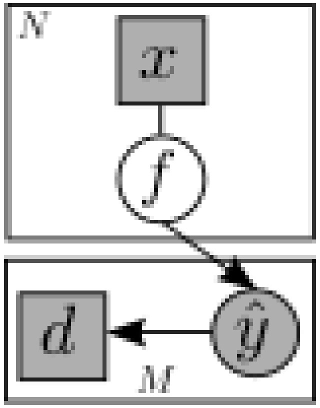
$$\mathcal{L}(\mathbf{f}_k|d_k, y_k, u_k, v_k) = p(d_k, y_k|\boldsymbol{f}_k) = \Phi\left(d_k y_k \frac{f(u_k) - f(v_k)}{\sqrt{2}}\right)$$

- Properties:
  - The choice of confidence y can not switch the binary decision decision.
  - With zero confidence, y=0 , the observation has no influence on $\boldsymbol{f}$.

Discrete Choice with Confidence Rating Case Ia - Likelihood



Discrete Choice with Confidence Rating



Case II – single observation

We observe that d_k and y_k cannot separate – so an equivalent (technical) likelihood is:

Continuous choice $\mathrm{z} = \begin{cases} [-1;0] & \text{if } \ d=-1 \\ [0;+1] & \text{if } \ d=+1 \end{cases}$

$$\mathcal{L}(\mathbf{f}_k | d_k, z_k, u_k, v_k) = p(d_k, z_k | \boldsymbol{f}_k) = \Phi\left(z_k \frac{f(u_k) - f(v_k)}{\sqrt{2}} \right)$$

The likelihood has the same properties in terms of inference, but in terms of the subjective and psychological meaning it constitutes a different paired comparison paradigm. We can e.g. not consider it as a forced choice.

Note that $p(z_k | \cdot)$ has support on [-1;1].

Degree of difference

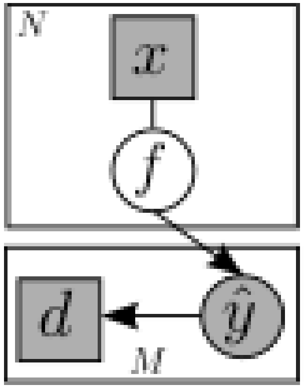
- Modeling the perceived degree of preference to which one option preferred over another
- Modeling consistency of the scale

Degree of difference

Objective: Possible to rate a particular difference between instances, e.g. similarity, degree of preference etc.

Case I: The paradigm involves observing the continuous degree of difference, y_k, between the two paired objects.

$$y_k = \begin{cases}]-1;0[& \text{if} \quad \text{difference} < 0 \\ 0 & \text{if} \quad \text{difference} = 0 \\ ]0;+1[& \text{if} \quad \text{difference} > 0 \end{cases}$$



Thus

$$\mathcal{L}(\mathbf{f}_k | y_k, u_k, v_k, \sigma) = p(y_k | \boldsymbol{f}_k)$$

We may ask question such as: What is the probability that p(y>0) which in effect can be exploited to derive a discrete outcome . But in this case we do not actually observe it and it is given deterministically by y, so there is no explicit noise model for the discrete choice, p(y>0|f).

Likelihood

$$\pi_k \in \,]0,1[$$

$$p\left(\pi_k \middle| \mathbf{f}_k\right) = \mathrm{Beta}\left(\pi_k \middle| \alpha(\mathbf{f}_k), \beta(\mathbf{f}_k)\right)$$

Inconsistency param.=precision

For fixed μ, large v is small variance

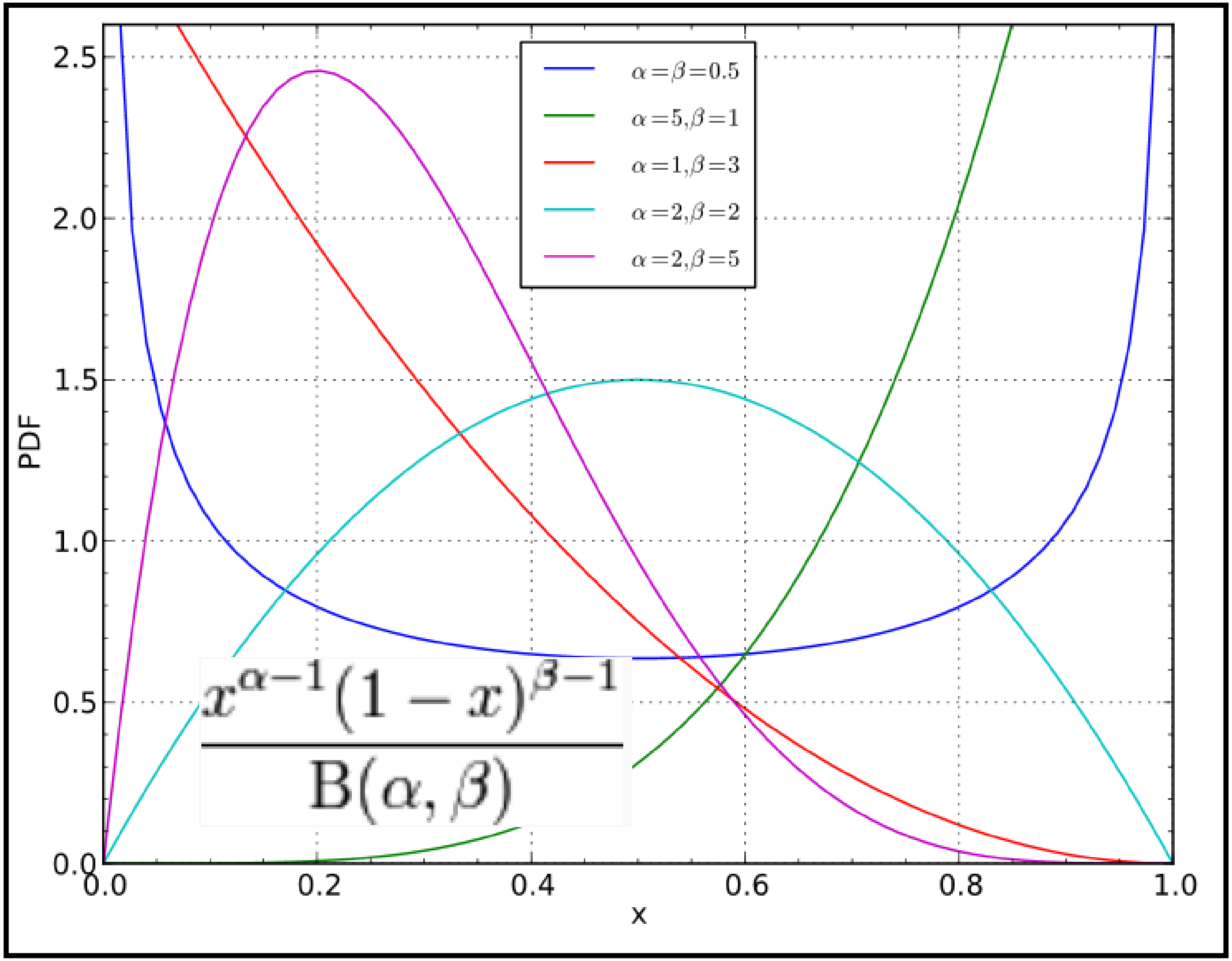
$$\alpha(\mathbf{f}_k) = \nu\mu(\mathbf{f}_k) \quad \beta(\mathbf{f}_k) = \nu(1-\mu(\mathbf{f}_k))$$

Mean function

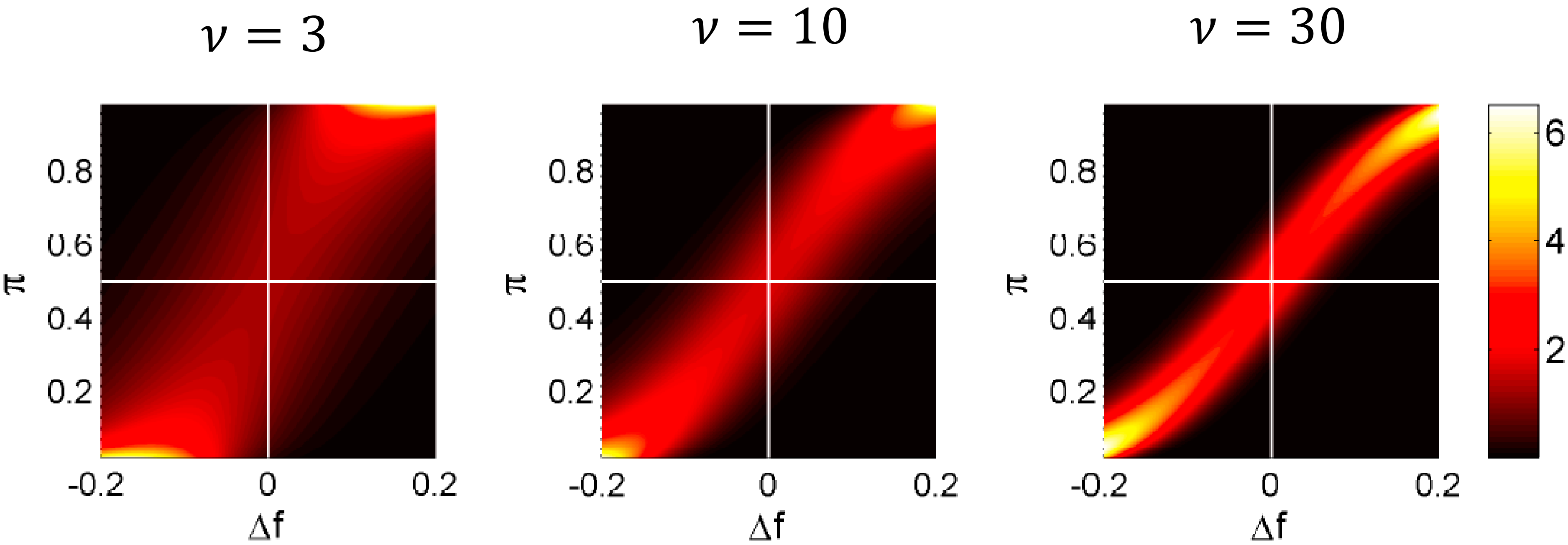
$$\mu\left(\mathbf{f}_k\right) = \Phi\left(\frac{f\left(v_k\right) - f\left(u_k\right)}{\sqrt{2}\sigma}\right)$$

$$\boxed{\mathcal{L}_{cont} \equiv p\left(\pi_k \middle| \mathbf{f}_k\right) = \mathrm{Beta}\left(\pi_k \middle| \nu\mu(\mathbf{f}_k), \nu(1-\mu(\mathbf{f}_k))\right)}$$

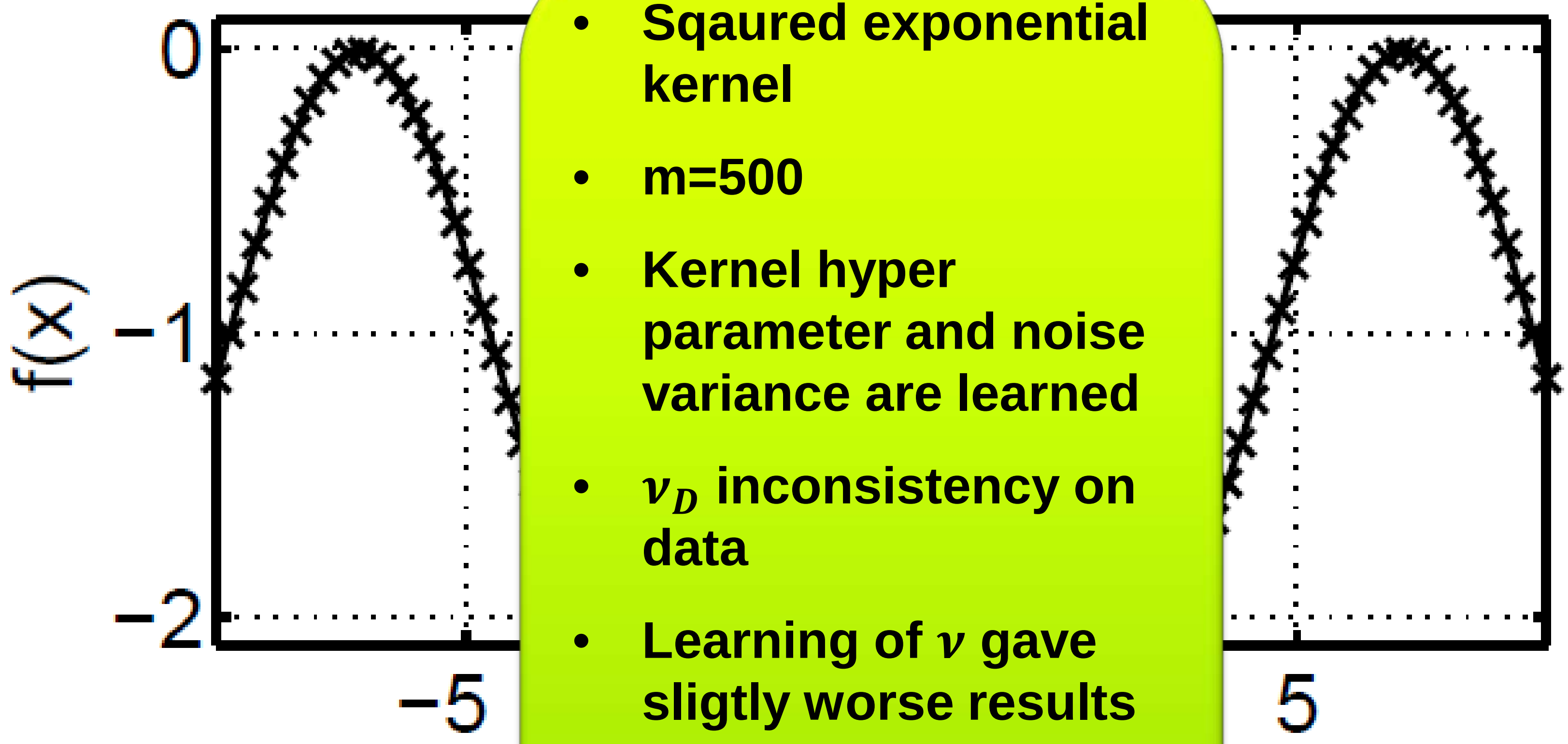
Beta distribution



Likelihood for $\sigma = 0.1$



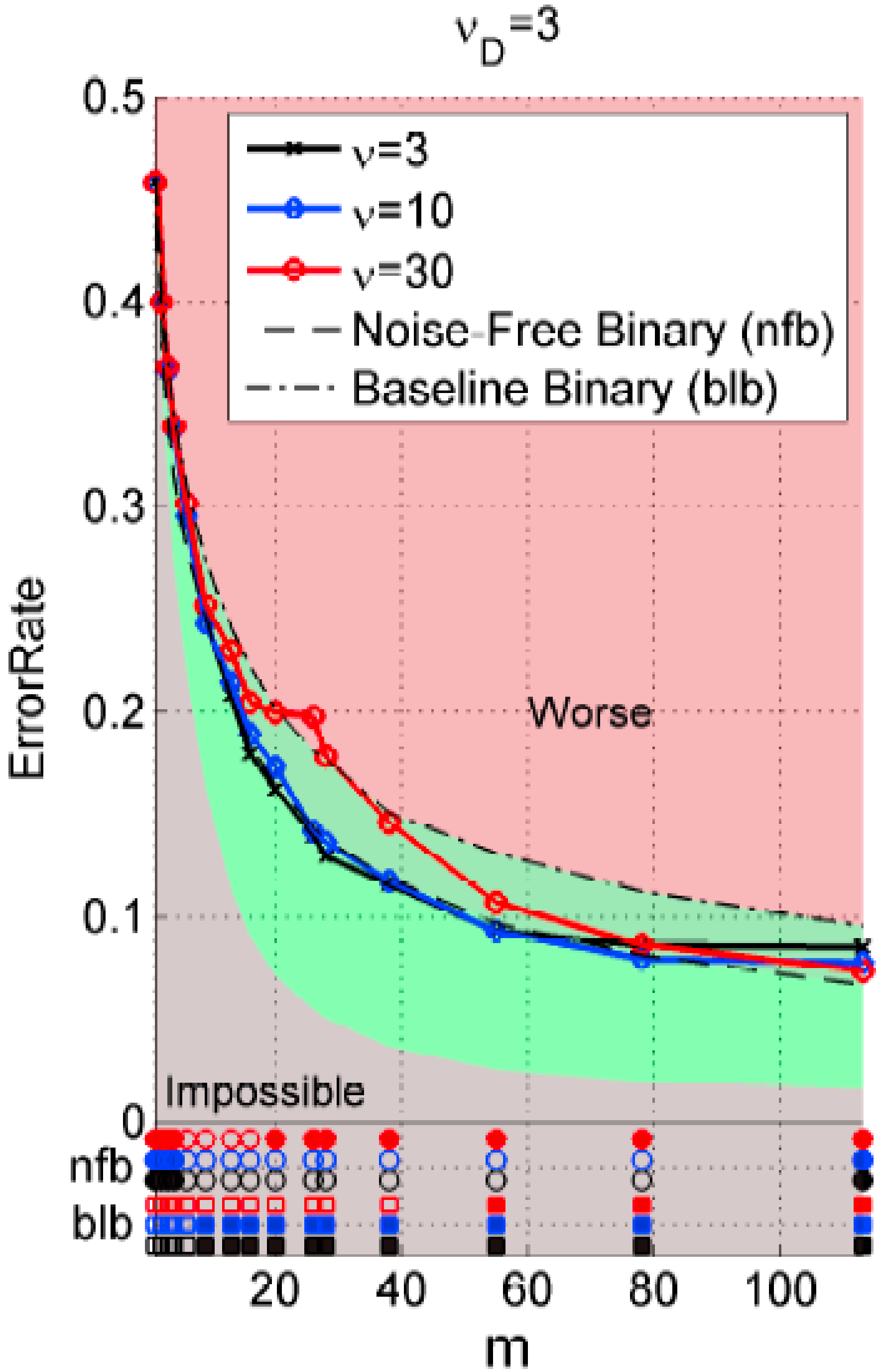
Syntetic experiment using the Griewangk function



- Sqaured exponential kernel
- m=500
- Kernel hyper parameter and noise variance are learned
- ν_D inconsistency on data
- Learning of ν gave sligtly worse results

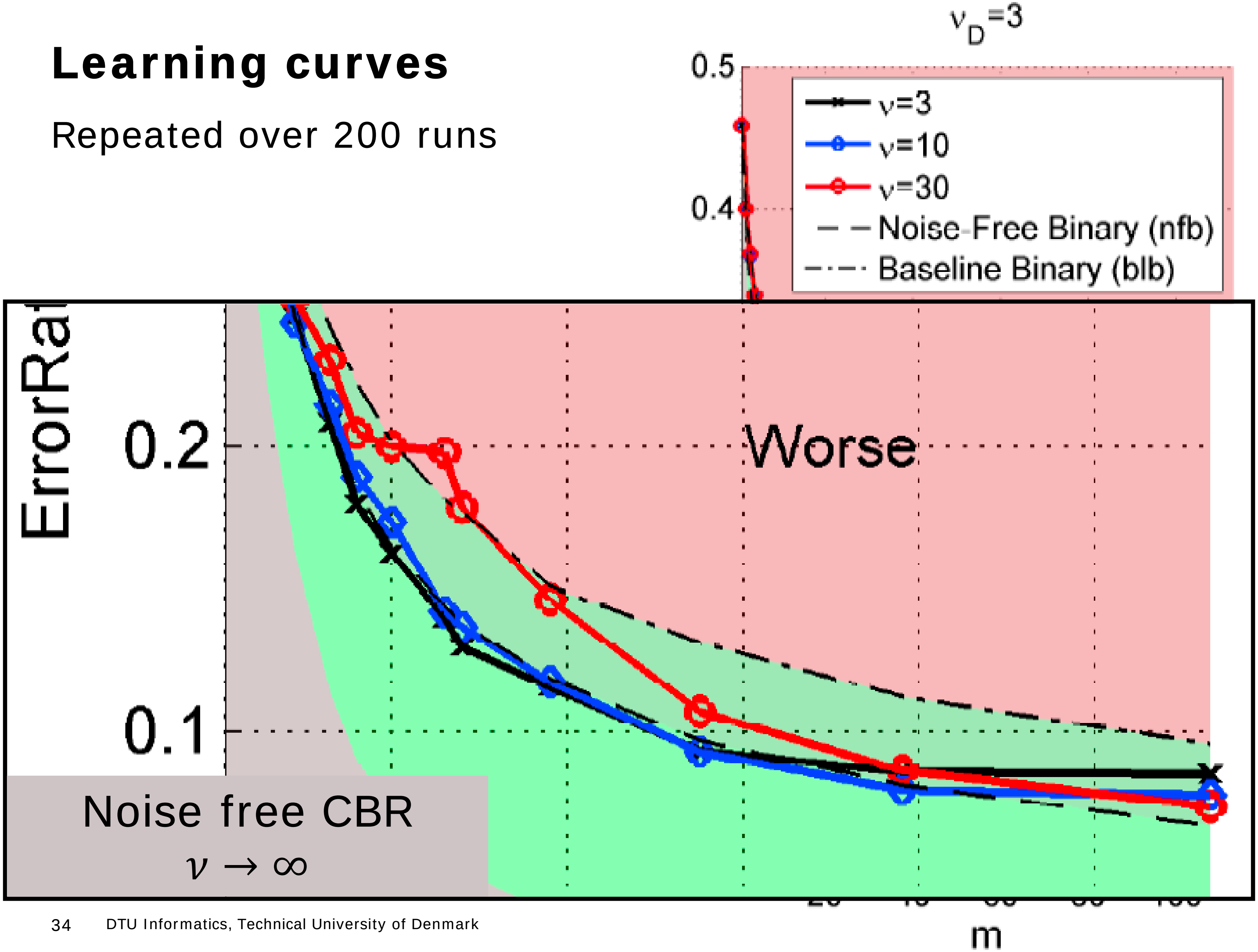
Learning curves

Repeated over 200 runs



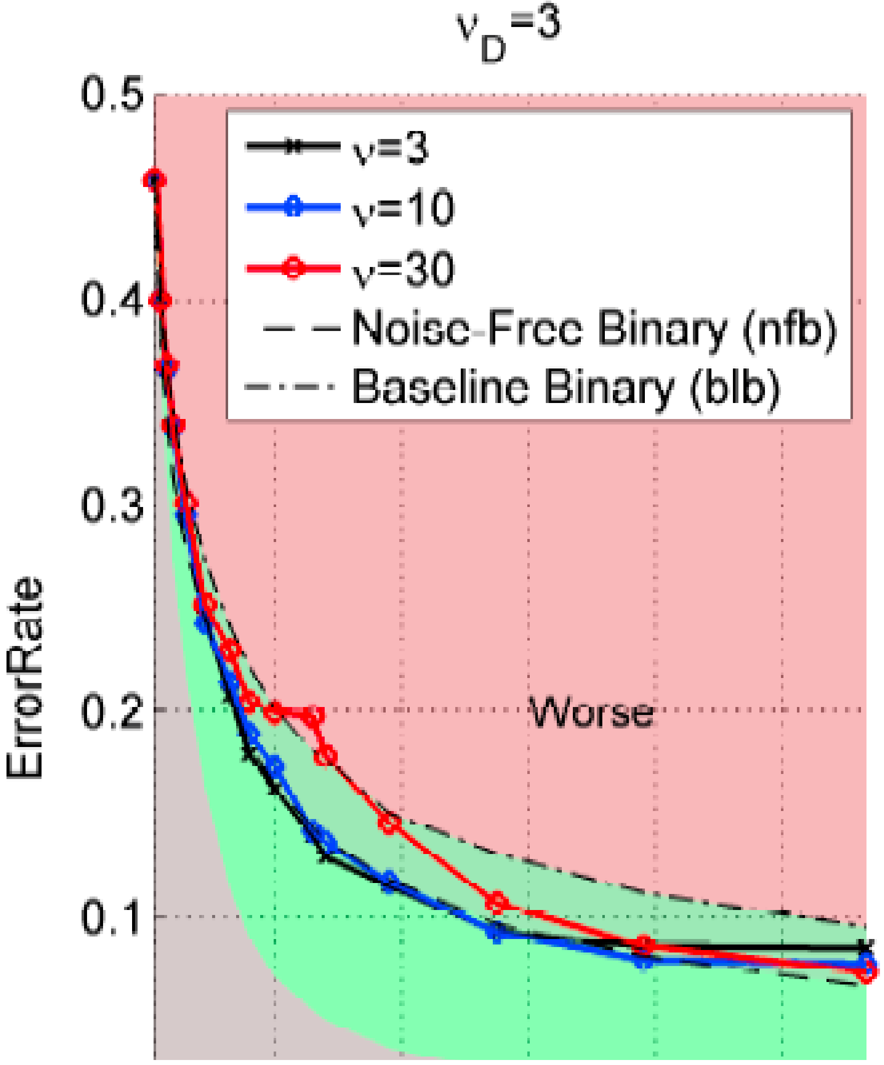
Learning curves

Repeated over 200 runs



Learning curves

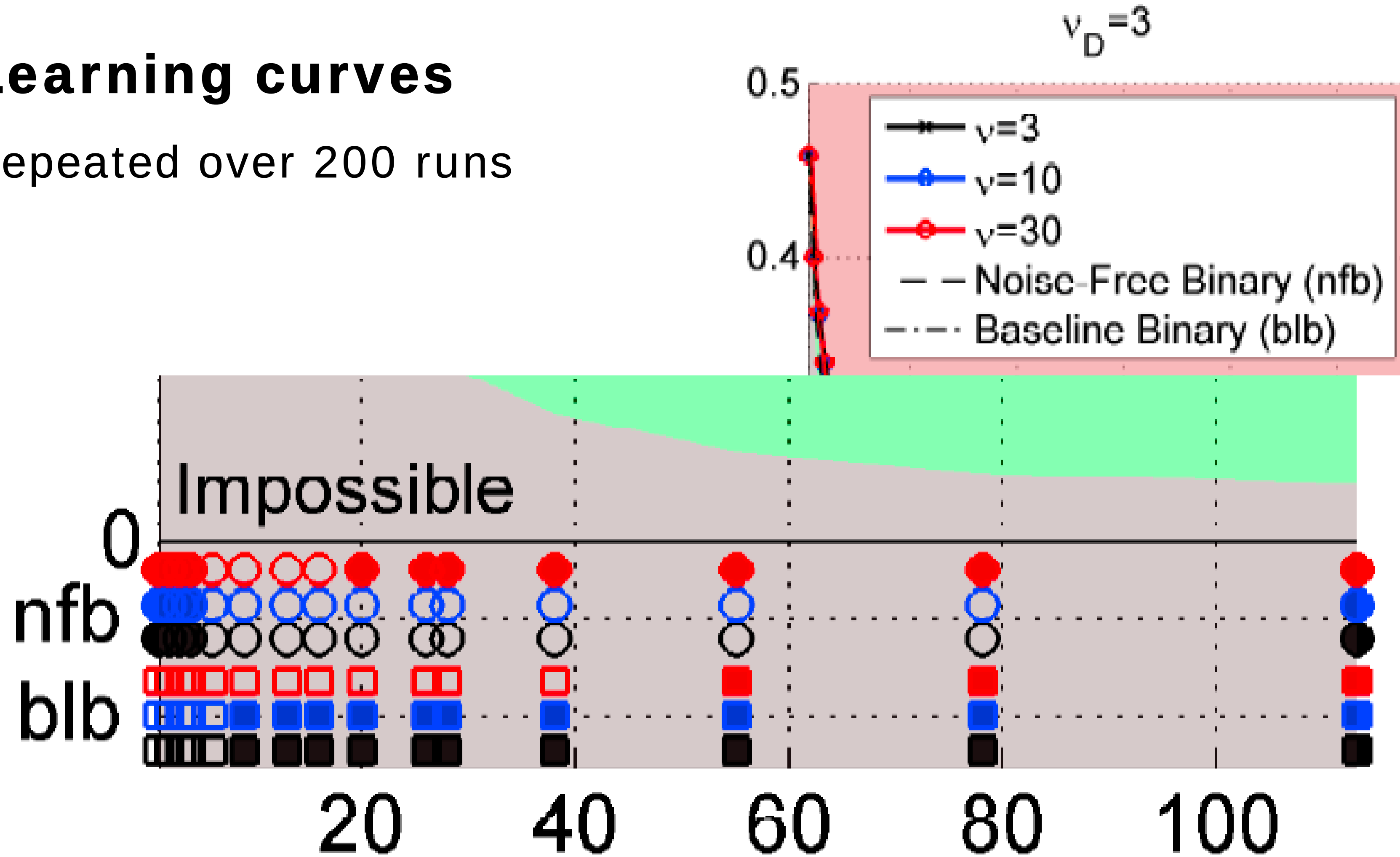
Repeated over 200 runs



Solid markers indicate significant difference of CBR model using a paried t-test with 5% significance level

Learning curves

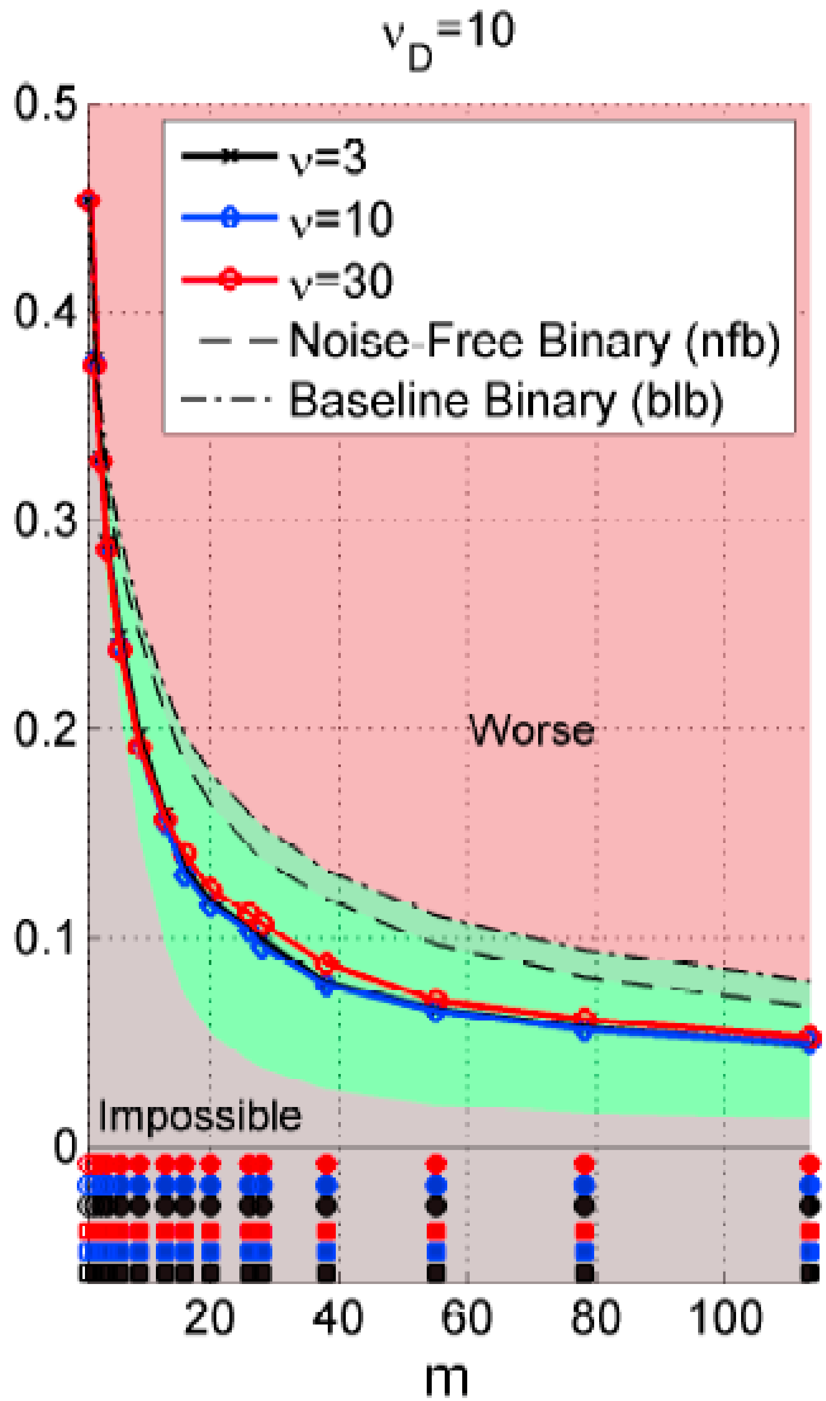
Repeated over 200 runs



Solid markers indicate significant difference of CBR model using a paried t-test with 5% significance level

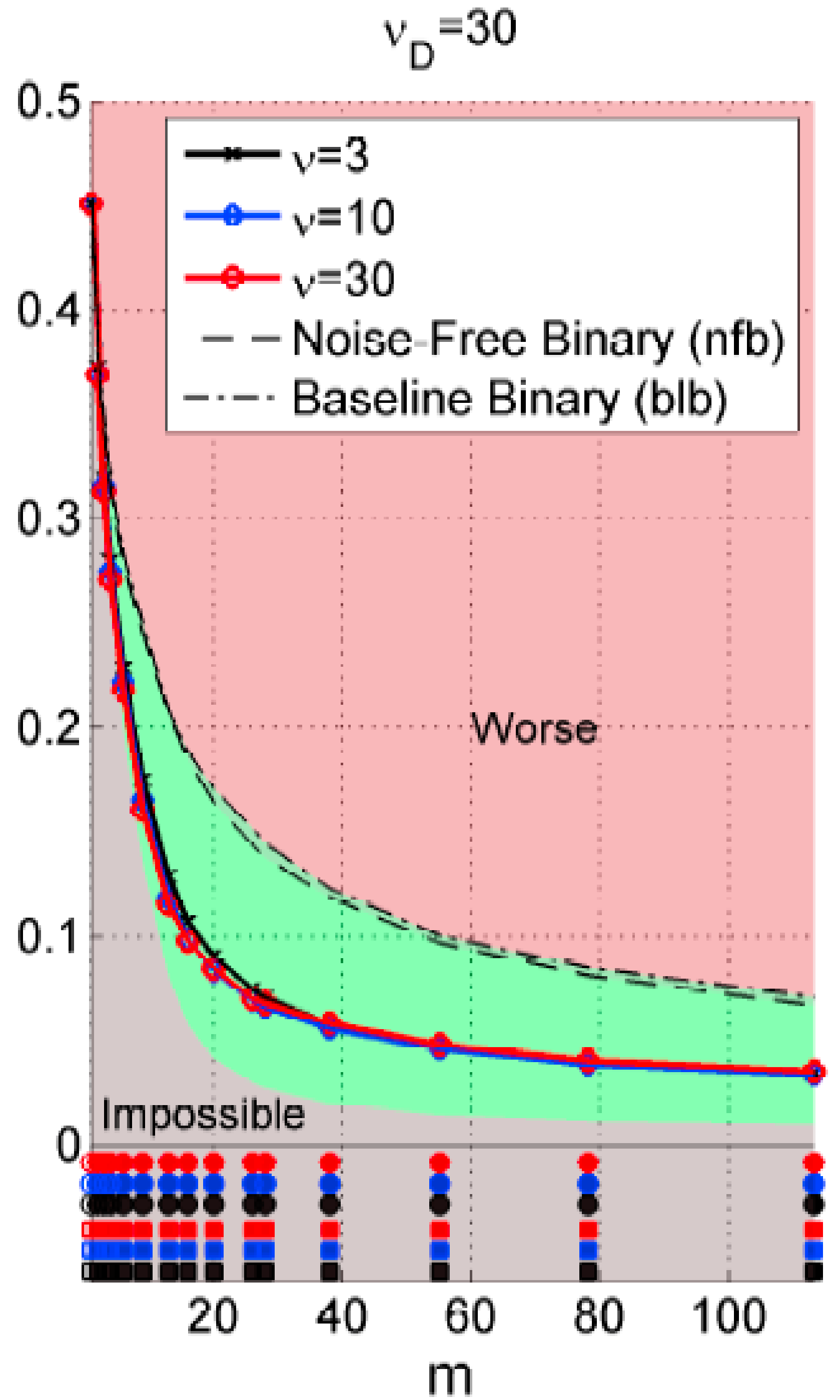
Learning curves

Repeated over 200 runs



Learning curves

Repeated over 200 runs



Linear model of difference in preference values

Standard linear, parametric model of the form $f(\mathbf{x}) = \mathbf{X}\mathbf{w}$

Pro

-Simple model, standard inference (IRLS, iterated reweighted least squares)

-Semi-flexible (e.g. polynomial regression)

Con

-Unregularized, i.e. unstable results for very few observation (or unconnected/noise free)

-Tricks (ad hoc) needed to handled special instance types such as distributions, Markov models and string objects.

-Standard tools does not account for uncertainty on $\mathbf{w}$'s

Gaussain Process preference function prior

GP function prior

Likelihood

Posterior we want to infer

$$p\left(\mathbf{f}|\mathcal{D}\right) = \frac{p\left(\mathcal{D}|\mathbf{f}\right)p(\mathbf{f})}{p\left(\mathcal{D}\right)}$$

No analytical form, hence, approximate inferece. We use Laplace approximation

Predicting preference

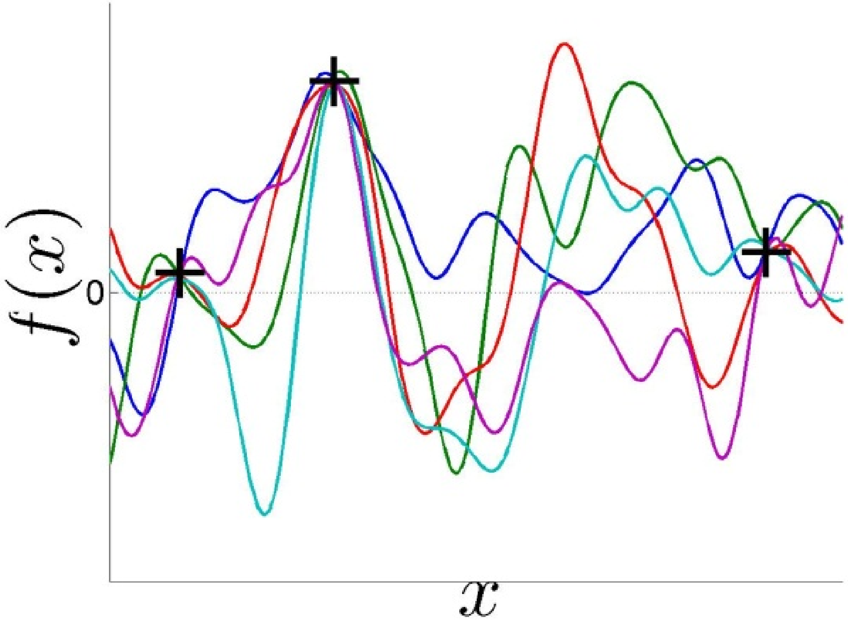
$$p\left(y_t|\mathcal{D}\right) = \int p\left(y_t|\mathbf{f}_t, \mathcal{D}\right) p\left(\mathbf{f}_t|\mathcal{D}\right) d\mathbf{f}_t$$

Non-Gaussian shape but for Probit likelihood analytical expression

Gaussian when using Laplace approximation

Gaussian Processes – the basics

- We can draw samples (entire functions) from the GP prior, as $f \sim \mathrm{GP}(0, K)$, where $[\boldsymbol{K}]_{i,j} = k(\mathbf{x}_i, \mathbf{x}_j)$
- As we continue to draw functions the mean function is zero everywhere, i.e. zero-mean GP



Gaussian Processes – sparse extension

Pseudo inputs

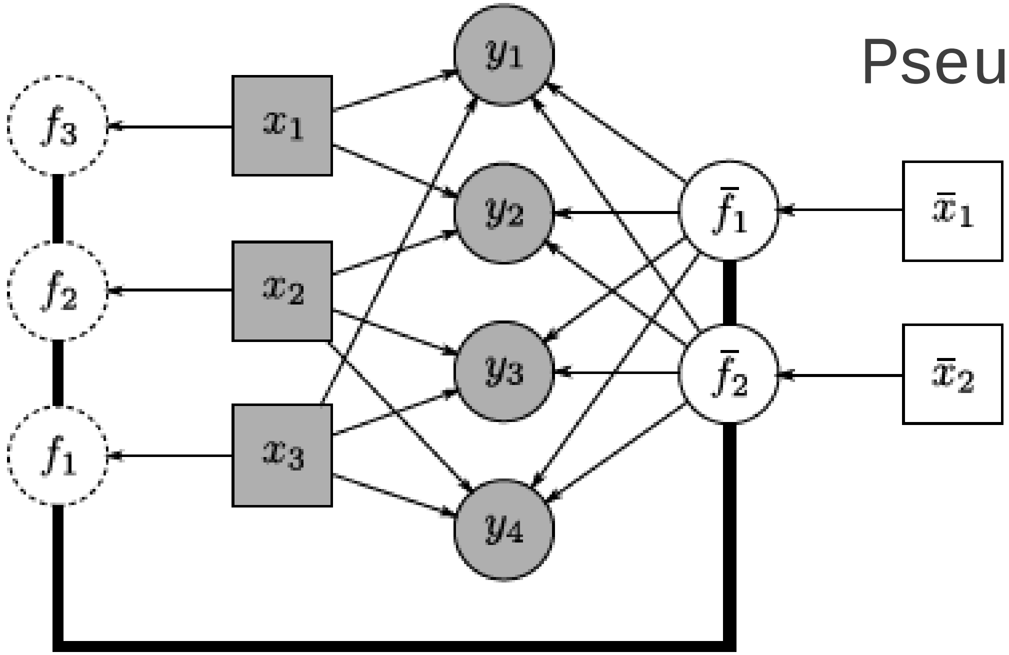


Figure 4: Graphical model

On Sparse Multi-Task Gaussian Process Priors for Music Preference Learning. Jens Brehm Nielsen, Bjørn Sand Jensen, and Jan Larsen. In NIPS CMPL workshop, pages 1-8, 2011.

Gaussian Processes – sparse extension

By defining a $\begin{bmatrix} \mathbf{f}_k \\ \bar{\mathbf{f}} \end{bmatrix} = \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{K}_{\mathbf{x}_k\mathbf{x}_k} & \mathbf{K}_{\bar{\mathbf{X}}\mathbf{x}_k}^\top \\ \mathbf{K}_{\bar{\mathbf{X}}\mathbf{x}_k} & \mathbf{K}_{\bar{\mathbf{X}}\bar{\mathbf{X}}} \end{bmatrix} \right)$, $l\}$ with $\mathbf{K}_{\mathbf{x}_k\mathbf{x}_k} = \begin{bmatrix} k(\mathbf{x}_{u_k}, \mathbf{x}_{u_k}) & k(\mathbf{x}_{u_k}, \mathbf{x}_{v_k}) \\ k(\mathbf{x}_{v_k}, \mathbf{x}_{u_k}) & k(\mathbf{x}_{v_k}, \mathbf{x}_{v_k}) \end{bmatrix}$, $\mathbf{K}_{\bar{\mathbf{X}}\mathbf{x}_k} = [\mathbf{k}_{u_k}, \mathbf{k}_{v_k}]$

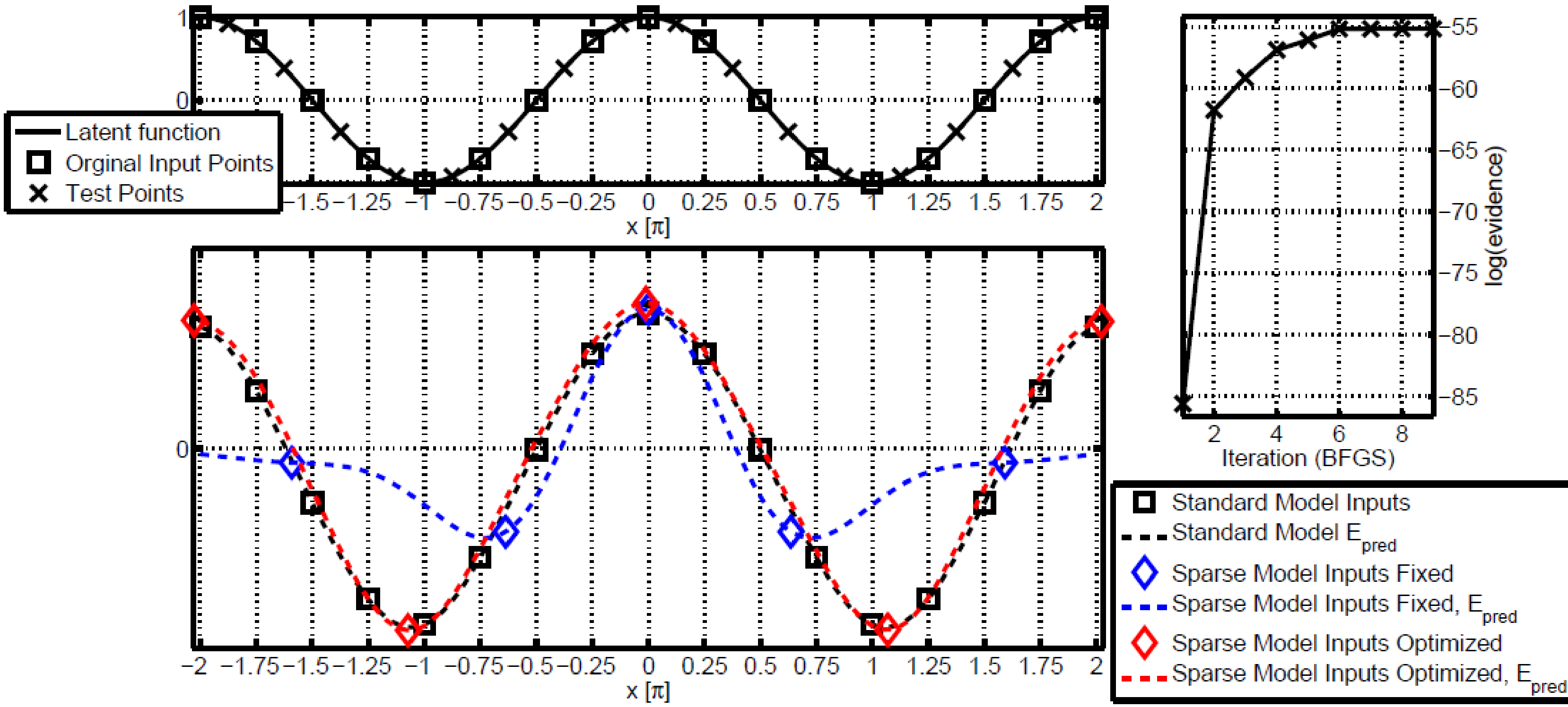
pric

No

vale

Thi

by (



On Sparse Multi-Task Gaussian Process Priors for Music Preference Learning. Jens Brehm Nielsen, Bjørn Sand Jensen, and Jan Larsen. In NIPS CMPL workshop, pages 1-8, 2011.

Gaussian Processes – multi-task

In multi-task learning the data we observe for different inputs is not necessarily related to the same task, e.g., user (task) rating different songs (inputs). Often, we can describe each task by a feature vector denoted $\boldsymbol{t} \in R^{d_t}$, e.g. (user age, mood, energy level etc.). Given a set of tasks

$$\mathcal{T} = \{\mathbf{t}_j | j = 1, \ldots, n_t\},$$

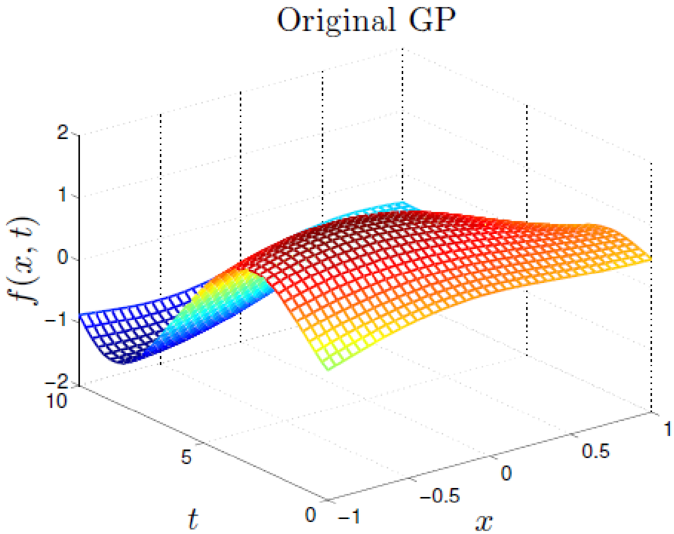
the multi-task kernel formulation from Bonilla *et al* (2008) is simply given by

$$\mathbf{K}_{\mathrm{MT}} = \mathbf{K}_{\mathcal{X}} \otimes \mathbf{K}_{\mathcal{T}}.$$

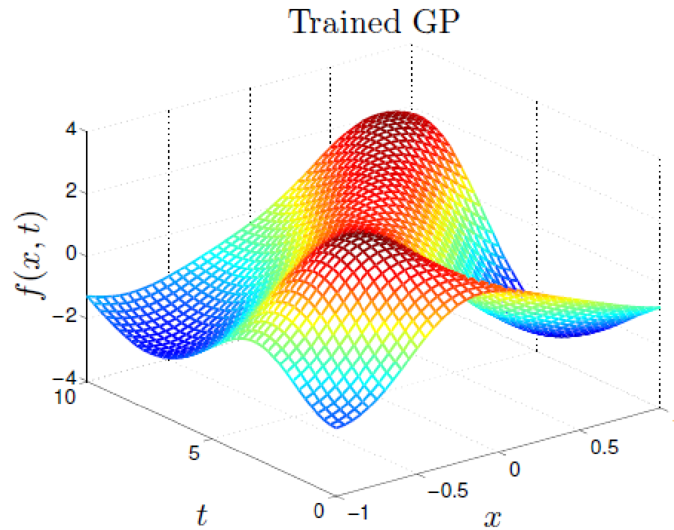
Gaussian Processes – multi-task

- This method fails in the pairwise case if the comparisons are made only between inputs for a given tasks (the case in the figure).
- Unfortunately, this is normally the case (e.g. one user compares two inputs.). The problem occurs because the data do not contain any information about the relationship across the tasks, hence the posterior is only influenced by the prior, specifying zero mean, which is obtained by shifting each of the task-specific functions to obtain zero mean over the input direction. The pairwise data is independent of such shifts.

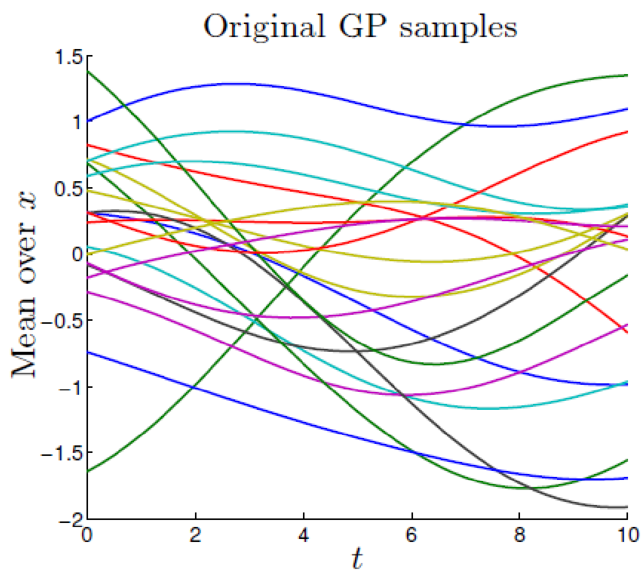
E. Bonilla, K. Ming and C Williams "*Multi-task Gaussian Process Prediction*" Advances in Neural Information Processing Systems 20, pp. 153–260, 2008



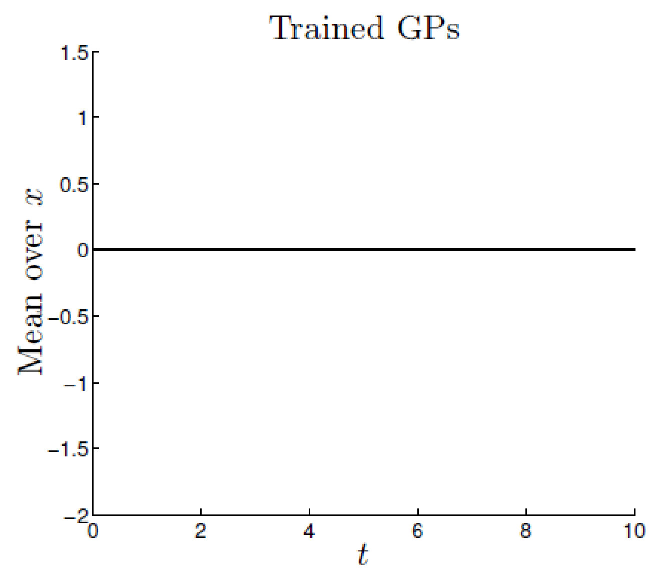
(a) GP prior sample



(b) Predictive GP mean



(c) Original means over input



(d) Resulting means over input

(Model Based) Sequential Design

- Goal:
  - 1) Unlimited resources (comparisons): Learn faster
  - 2) Limited resources (comparisons): Learn better (faster) with only the most informative experiment

Active learning by expected value of information EVOI

- Expected improvement $EI(\boldsymbol{x})$ is the expected improvement selecting input $\boldsymbol{x}$ relative to a current best perference function value
- MEI is the maximum of $EI(\boldsymbol{x})$ over $\boldsymbol{x}$

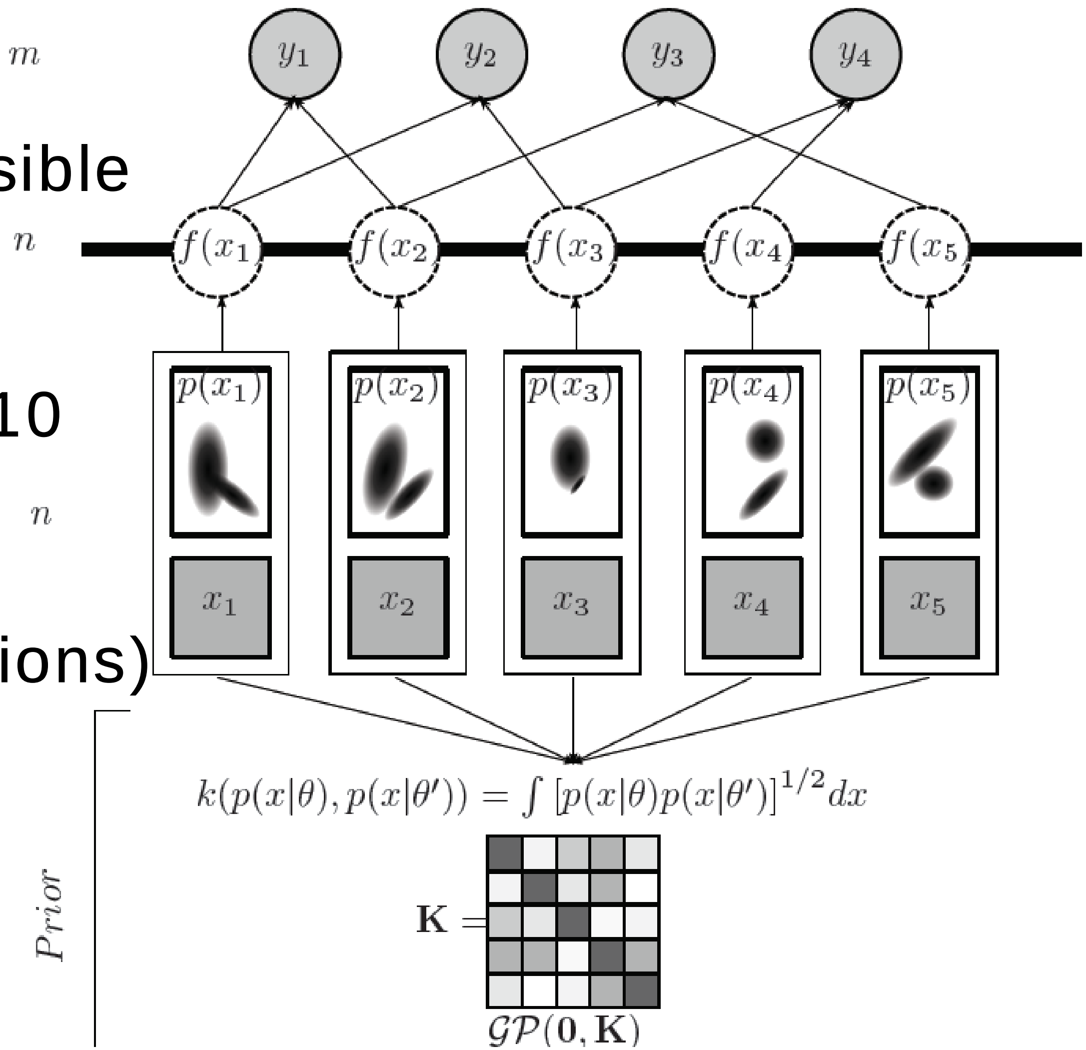
$$\begin{aligned}
\mathrm{EVOI}(\mathcal{D}, i, j) &= -\mathrm{MEI}(\mathcal{D}) + \left\langle \sum_{q_{ij}} p(q_{ij}|\mathrm{f}^*, \mathcal{D}) \mathrm{MEI}(\mathcal{D} \cup q_{ij}) \right\rangle_{p(\mathbf{f}^*|\mathcal{D})} \\
&= -\mathrm{MEI}(\mathcal{D}) + \left\langle p(x^{*i} \succ x^{*j}|\mathrm{f}^*, \mathcal{D}) \right\rangle_{p(\mathbf{f}^*|\mathcal{D})} \mathrm{MEI}(\mathcal{D} \cup \{x^{*i} \succ x^{*j}\}) \\
&\quad + \left\langle p(x^{*j} \succ x^{*i}|\mathrm{f}^*, \mathcal{D}) \right\rangle_{p(\mathbf{f}^*|\mathcal{D})} \mathrm{MEI}(\mathcal{D} \cup \{x^{*j} \succ x^{*i}\}),
\end{aligned}$$

E. Bonilla, S. Guo, and S. Sanner, “Gaussian Process preference elicitation,” in Advances in Neural Information Processing Systems 23, J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R.S. Zemel, and A. Culotta, Eds., pp. 262–270. 2010.

Applications

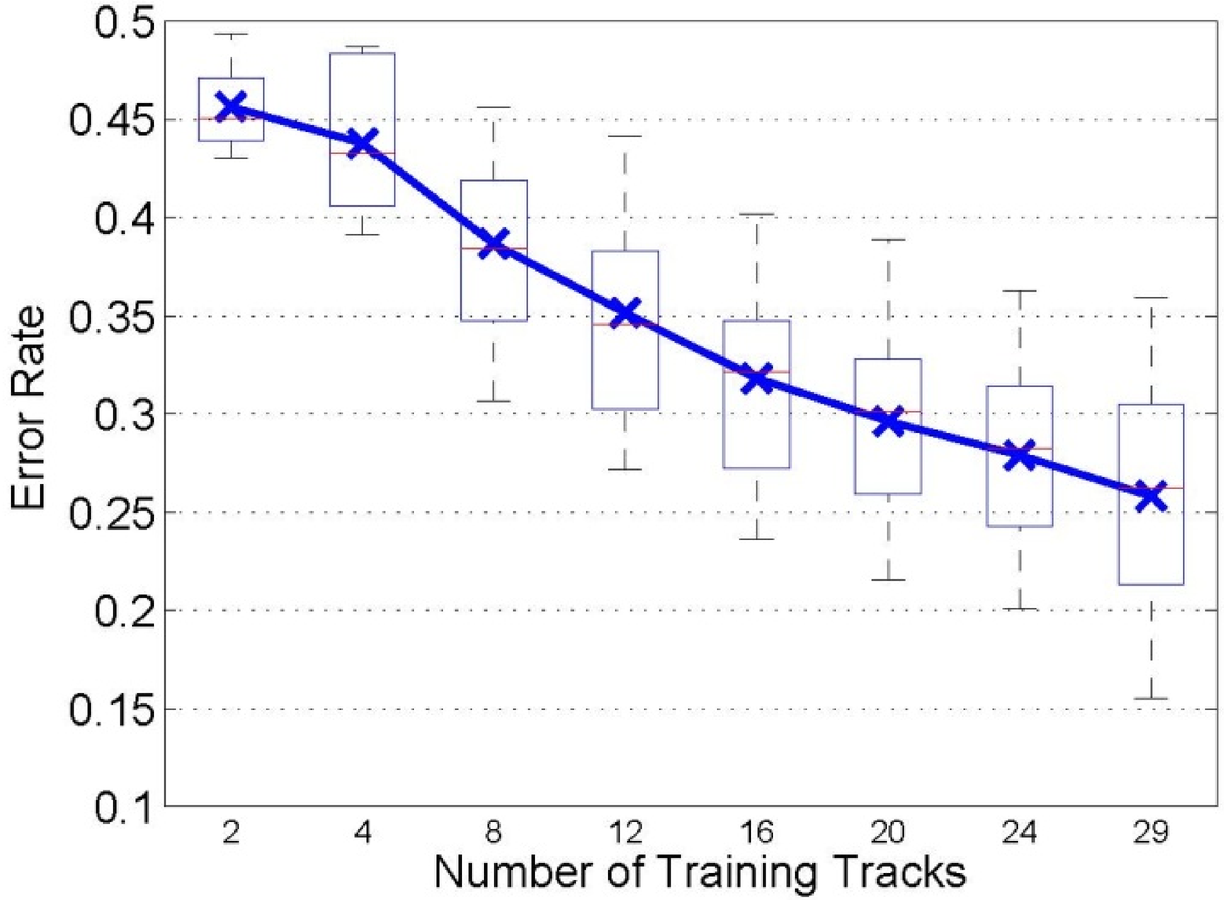
Music Preference

- 10 subjects
- 155 out of 450 possible pairs evaluated
- 3 genres (Classical, Rock, Heavy) with 10 tracks
- Standard features (MFCCs, 26 dimensions)
- Probit likelihood
- GP Prior: PPK: $k\big(p(x|\theta), p(x|\theta')\big) = \int \big(p(x|\theta)p(x|\theta')\big)^{1/2} dx$



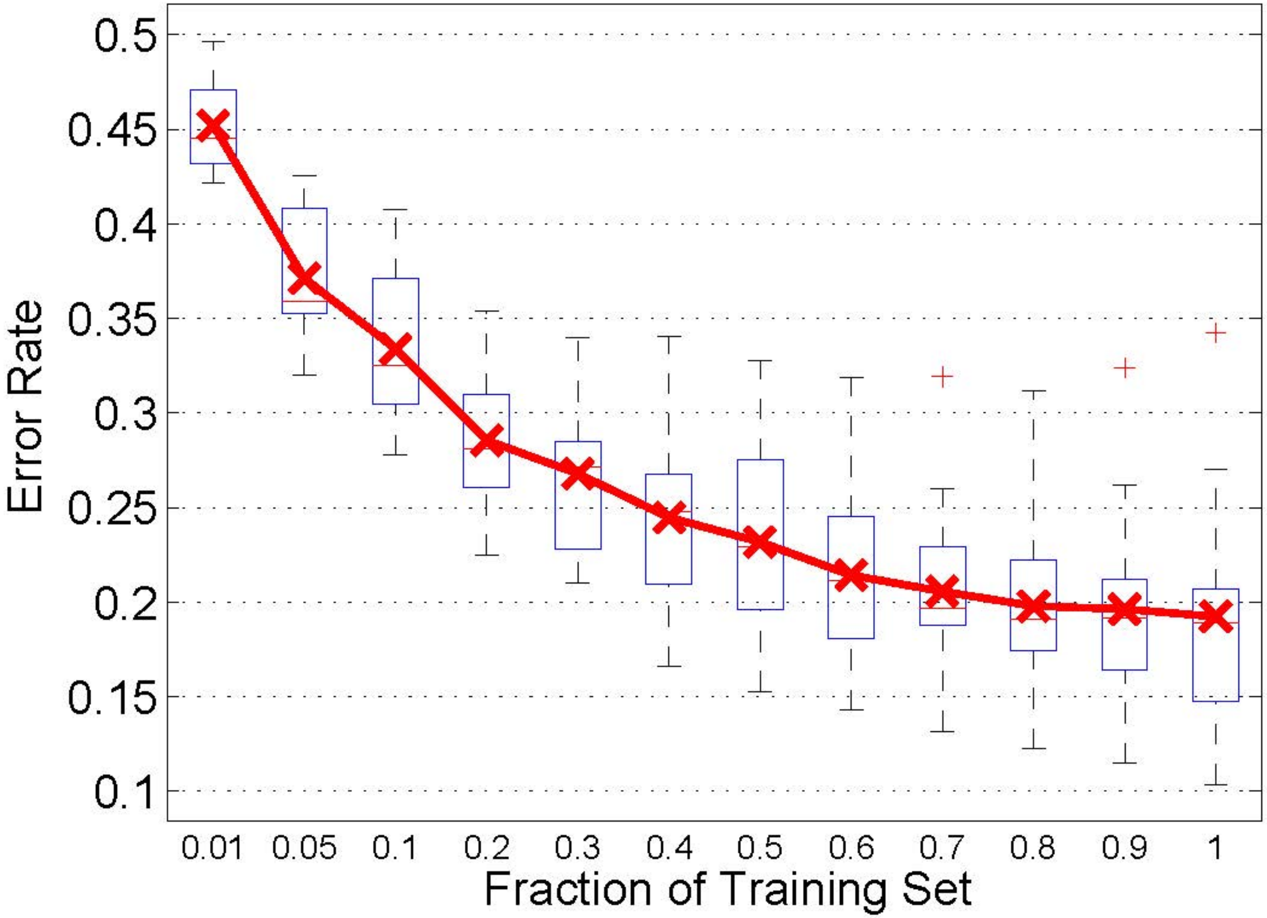
Music Preference

Leave one song out

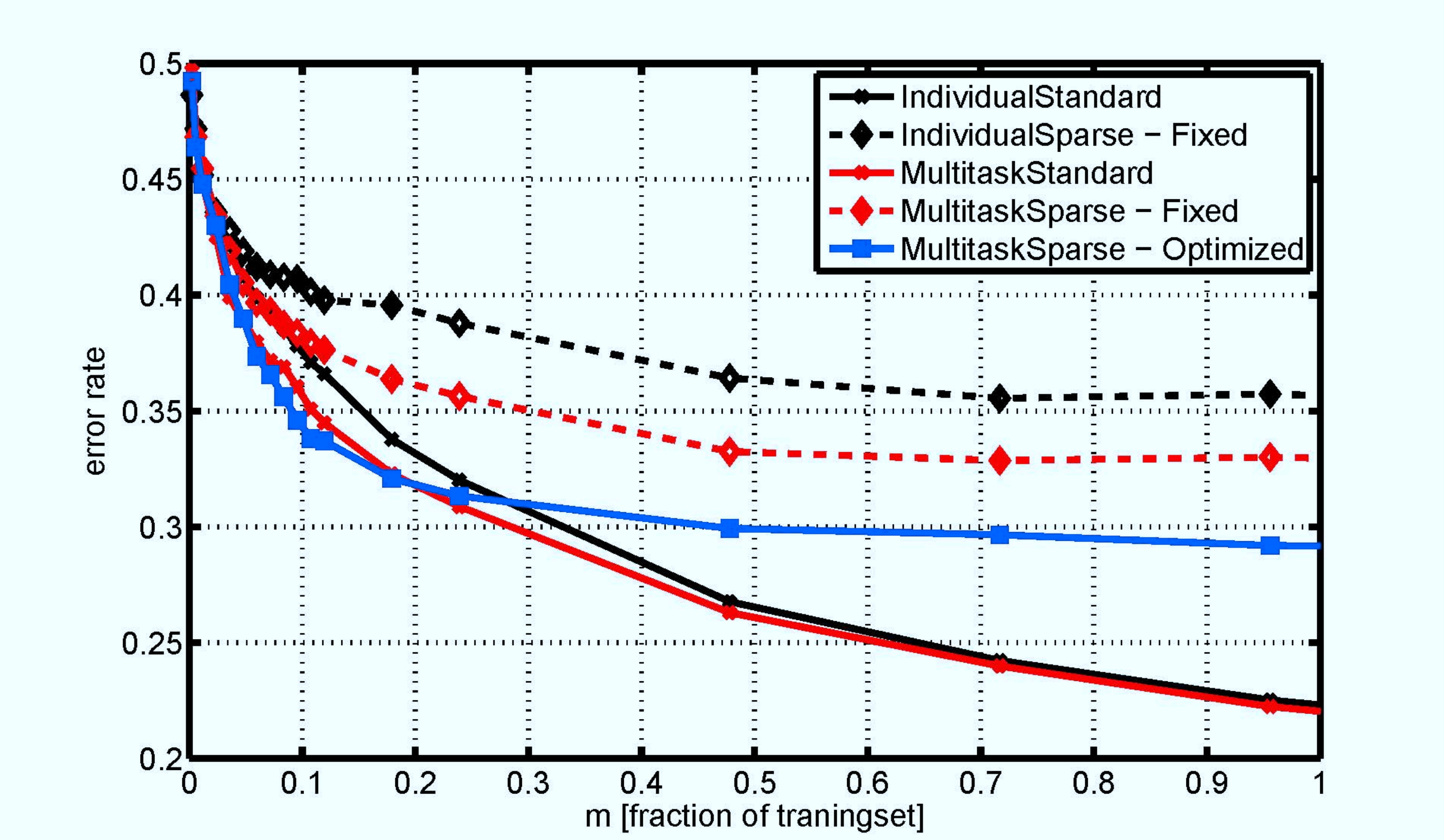


Music Preference

10 fold CV



Music Preference - sparse models



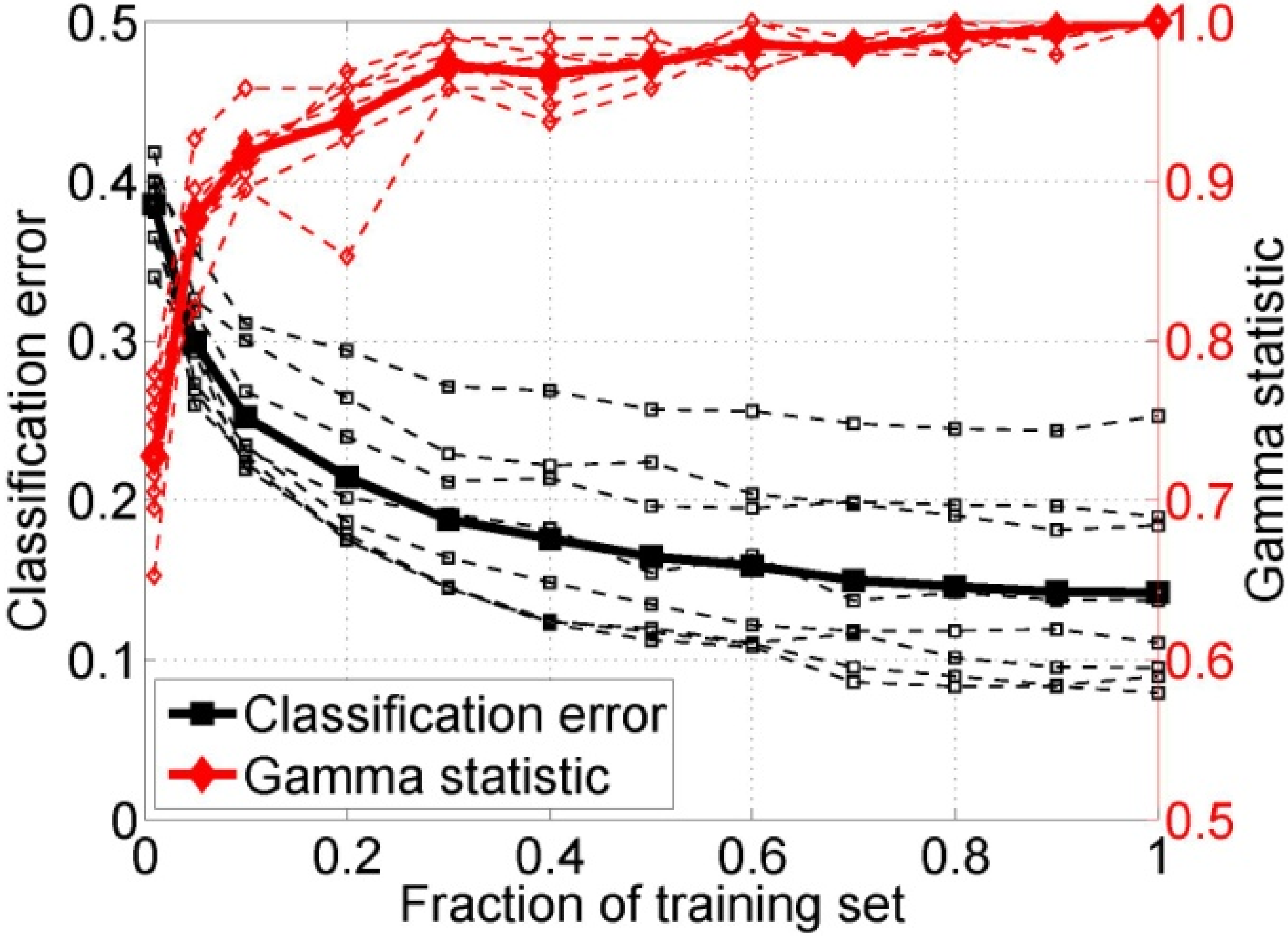
Sparse Multi-Task Gaussian Process Priors for Music Preference Learning. Jens Brehm Nielsen, Bjørn Sand Jensen, and Jan Larsen. In NIPS CMPL workshop, pages 1-8, 2011.

Music Emotions

- Same framework as for music preference
- USPOP2002 dataset
- 8 subjects (2 females, 6 males)
- Valence and Arousal preference scored individually in random order
- All 190 possible pairs are evaluated

Modeling Expressed Emotions in Music using Pairwise Comparisons, Jens Madsen, Jens Brehm Nielsen, Bjørn Sand Jensen, and Jan Larsen, CMMR 2012

Music Emotions: Arousal



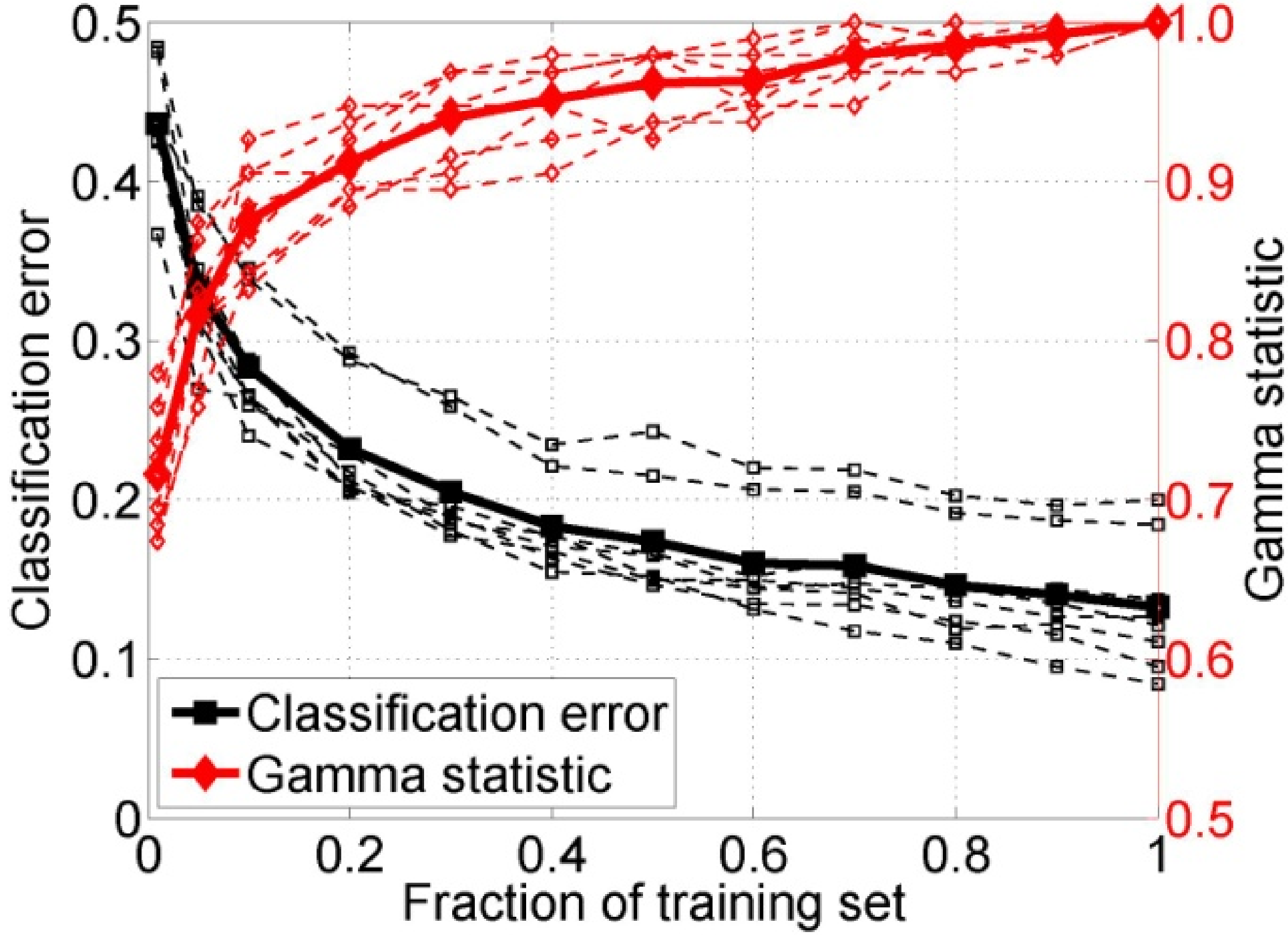
Gam

Ns-N

Ns:

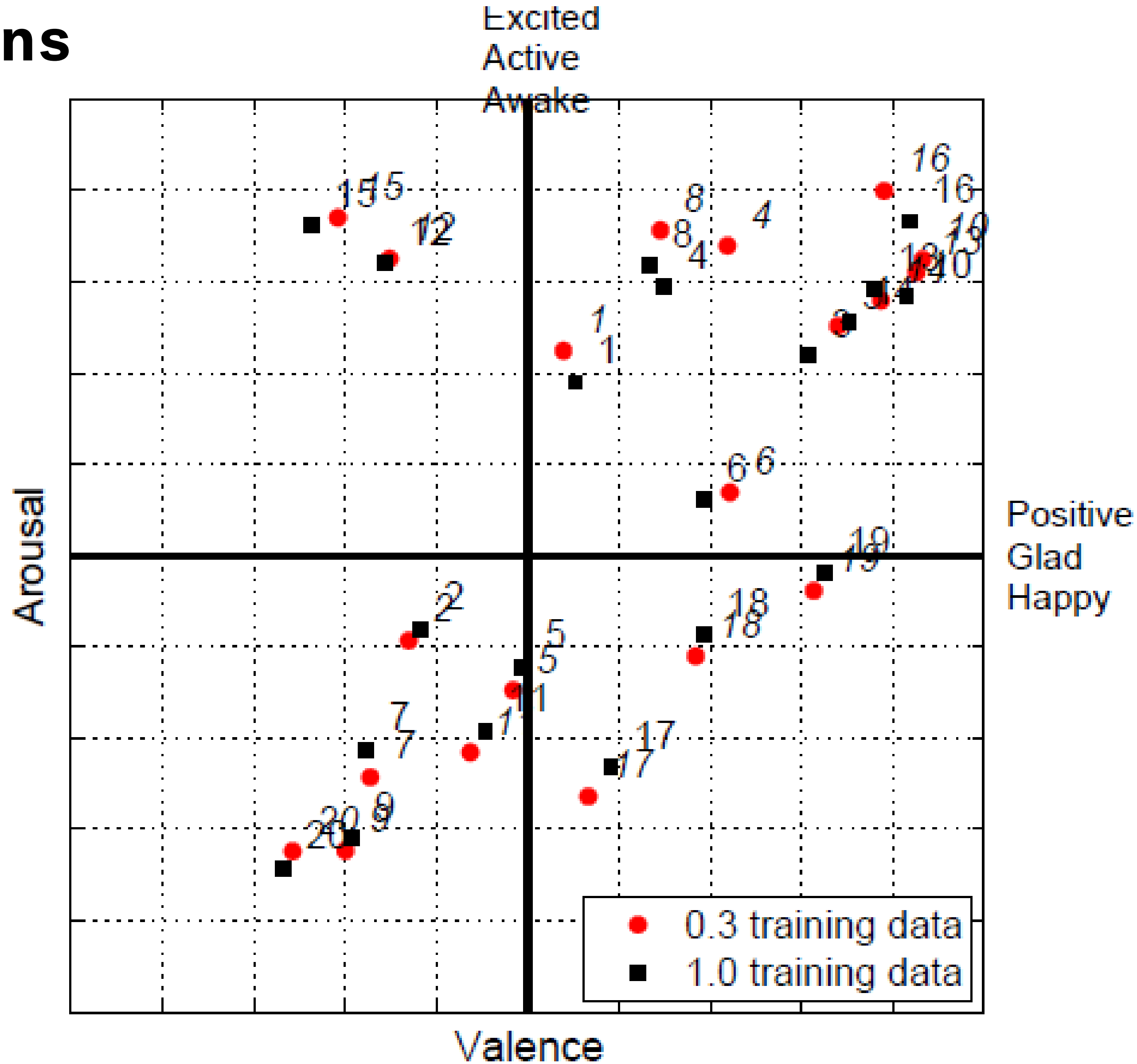
Nd:

Music Emotions: Valence



Music Emotions

No.	Song name
1	311 - T and p combo
2	A-Ha - Living a boys adventure
3	Abba - Thats me
4	Acdc - What do you do : money honey
5	Aaliyah - The one i gave heart to
6	Aerosmith - Mother pop
7	Alanis Morissette - Thes the thoughts
8	Alice Cooper - Im your
9	Alice in Chains - Killer i
10	Aretha Franklin - A cha
11	Moby - Everloving
12	Rammstein - Feuer frei
13	Santana - Maria caracol
14	Stevie Wonder - Anothe
15	Tool - Hooker with a pe
16	Toto - We made it
17	Tricky - Your name
18	U2 - Babyface
19	Ub40 - Version girl
20	Zz top - Hot blue and righteous



AV values obtained by normalizing preference functions

Hearing Aid Optimization

- Industrial PhD project width Widex on optimizing heading aids

Conclusions and outlook

- New modeling frameworks for pairwise elicitation of music preference (music, exoressed emotion in songs)
- Sparsification helps reducing the number of evaluations as well as more elaborate (discrete choice with confidence rating and degree of difference models)
- Active learning scheme to be incorporated
- Multi-user and hierarchical models
- Compare pairwise with direct scaling (in emotion elicitation)
- Evaluate on large real data sets in higher input dimensions is to be done