# EFFICIENT PREFERENCE LEARNING
# WITH PAIRWISE CONTINUOUS OBSERVATIONS AND GAUSSIAN PROCESSES

*Bjørn Sand Jensen[1], Jens Brehm Nielsen[1,2] & Jan Larsen[1]*

[1]Technical University of Denmark,
Department of Informatics and
Mathematical Modeling,
Richard Petersens Plads B321
2800 Lyngby, Denmark

[2]Widex A/S,
Nymøllevej 6
3540 Lynge, Denmark

## ABSTRACT

Human preferences can effectively be elicited using pairwise comparisons and in this paper current state-of-the-art based on binary decisions is extended by a new paradigm which allows subjects to convey their degree of preference as a continuous but bounded response. For this purpose, a novel Beta-type likelihood is proposed and applied in a Bayesian regression framework using Gaussian Process priors. Posterior estimation and inference is performed using a Laplace approximation.

The potential of the paradigm is demonstrated and discussed in terms of learning rates and robustness by evaluating the predictive performance under various noise conditions on a synthetic dataset. It is demonstrated that the learning rate of the novel paradigm is not only faster under ideal conditions, where continuous responses are naturally more informative than binary decisions, but also under adverse conditions where it seemingly preserves the robustness of the binary paradigm, suggesting that the new paradigm is robust to human inconsistency.

***Index Terms***— Pairwise Comparisons, Continuous Response, Gaussian Processes, Laplace Approximation

## 1. INTRODUCTION

Traditionally, various aspects of human perception and cognition are assumed to be related to absolute psychological magnitudes or intensities. This includes the classical findings by Weber, Fechner and Stevens who, for example, investigated the perception of light intensity. However, recently Lockhead [1] has argued that every aspect of perception is relative, even those apparently absolute aspects investigated by Weber, Fechner and Stevens. In accordance with the theory in [1], we investigate human perception from a relative viewpoint and examine one such highly relative aspect, namely preference.

Formal treatment of relative aspects goes back to the ideas of Thurnstone [2] and the principle of comparative judgments.

In the present context it was revisited by Chu *et al.* [3] who formulated a Bayesian approach to preference learning using Gaussian Process (GP) priors. This formulation has initiated a number of related studies and applications, such as audiological preference [4], multi-subject food preference [5] and an extension for semi-supervised, active learning settings [6].

In this work we extend the likelihood model in [3] to support observations which in effect measure *the perceived degree to which one option is preferred over another*. This degree of preference can be obtained from a traditional paired comparison test, which implies that a subject is asked to give a subjective assessment of the degree to whether $A$ or $B$ is preferred over the other. Specifically, we model the observed degrees of preferences through a likelihood conditioned on a functional value difference and support inconsistent observations by applying a re-parameterized Beta distribution.

In a traditional setting, users would not be trusted to be able to quantify such an abstract and difficult aspect as degree of preference. Instead, we would rely on massive repetitions of a standard binary experiment to estimate the proportion of $A \succ B$ using this as an expression of the degree of any preferences. However, we want to exploit the extra information from continuous responses to get a faster method for preference elicitation without jeopardizing the robustness from standard binary responses. The hypothesis is that we are able to learn faster by (indirectly) observing the perceived probability of $A \succ B$ as opposed to a binary decision. Applying appropriate priors and noise modeling should ensure this to be true also under adverse conditions.

In order to examine this hypothesis, we apply the novel likelihood in a flexible Bayesian setup similar to [3] in which the prior on the underlying preference function is defined by a GP with a potentially complex covariance structure. The Laplace approximation is used for inference and model selection by *maximum-a-posteriori* (MAP) estimates. This provides a consistent probabilistic framework for making predictions and evaluating the predictive uncertainty. We use simulations with different synthetic noise scenarios in order

to compare a standard binary decision with the novel model. The performance of both methods is evaluated using the predictive performance.

## 2. MODELS FOR PAIRWISE OBSERVATIONS

In the previous section, we motivated pairwise comparisons from a cognitive perspective, yet pairwise comparisons can be considered more broadly. It is usually possible to describe any aspect of a pairwise comparison, such as preference, real difference, or perceived similarity in terms of a latent function [2].

In the following we will model the preference of two distinct inputs, $u \in \mathcal{X}$ and $v \in \mathcal{X}$, in terms of the difference between two functional values, $f(u)$ and $f(v)$. This implies a function, $f : \mathcal{X} \to \mathbb{R}$, which defines an internal, but latent absolute preference.

The general setup is as follows: We consider $n$ distinct inputs $x_i \in \mathcal{X}$ denoted $\mathcal{X} = \{x_i | i = 1, ..., n\}$, and a set of $m$ responses on pairwise comparisons between any two inputs in $\mathcal{X}$, denoted by

$$\mathcal{Y} = \{(y_k; u_k, v_k) | k = 1, ..., m\},$$

where $y_k \in \mathbb{Y}$. $u_k \in \mathcal{X}$ and $v_k \in \mathcal{X}$ are option one and two in the $k$'th pairwise comparison, respectively. The main topic of this paper is how the domain of the response variable influences the learning rate of the latent function $f$ in relation to the number of paired comparisons. As previously indicated, we will consider two cases:

- **binary** where $y_k = d_k, d_k \in \{-1, 1\}$
- **continuous and bounded** where $y_k = \pi_k, \pi_k \in ]0, 1[$.

In both cases we consider $y$ a stochastic variable, informally implying the definition of the conditional density given by $p(y_k | f_k(u_k), f(v_k))$, denoted by $p(y_k | \mathbf{f}_k)$ with $\mathbf{f}_k = [f(u_k), f(v_k)]^\top$.

### 2.1. Binary Response

When restricting the response variable to be a discrete, two-alternatives, forced choice, paired-comparison between the two presented options, we define the response variable as $d_k \in \{-1, 1\}$. A preference for either $u_k$ or $v_k$ is indicated by $-1$ or $+1$, respectively.

When considering noise on the forced decisions the resulting random variable can be modeled by a classic choice model such as the Logit or Probit [7, chapter 6]. In the current setting we restrict ourself to the Probit model mainly for analytical reasons.

Given a function, $f$, we can define the likelihood of observing a discrete choice $d_k$ directly as the conditional density.

$$p(d_k | \mathbf{f}_k, \boldsymbol{\theta}_{\mathcal{L}}) = \Phi\left(d_k \frac{f(v_k) - f(u_k)}{\sqrt{2}\sigma}\right), \quad (1)$$

where $\Phi(x)$ is the cumulative Gaussian (with zero mean and unity variance) and $\boldsymbol{\theta}_{\mathcal{L}} = \{\sigma\}$. This classic Probit likelihood is by no means a new invention and can be dated back to Thurstone and his fundamental definition of *The Law of Comparative Judgment*[2]. However, it was first considered with GPs in [3] and later in e.g. [5] and [6].

### 2.2. Continuous Response

The primary contribution of this paper is a novel response model allowing for more subtle judgments, where the response variable describes the degree to which the prevailing option is preferred.

For this purpose we formally define a continuous but bounded response $\pi \in ]0; 1[$ observed when comparing $u$ and $v$. The first option, $u$, is preferred for $\pi < 0.5$. The second option, $v$, is preferred for $\pi > 0.5$ and none is preferred for $\pi = 0.5$. Hence, the response captures both the choice between $u$ and $v$, and the degree of the preference.

Instead of using the Probit function directly as the choice model, it is used as a link function mapping from functional differences to continues bounded responses. More precisely, the Probit is used as a mean function for a Beta type distribution with parameterized shape parameters $\alpha$ and $\beta$, thus

$$p(\pi_k | \mathbf{f}_k) = \text{Beta}(\pi_k | \alpha(\mathbf{f}_k), \beta(\mathbf{f}_k)).$$

To express the shape parameters of the Beta distribution as a function of the Probit mean function $\mu(\mathbf{f}_k)$, we apply a well-known re-parametrization of the Beta distribution [8].

$$\alpha(\mathbf{f}_k) = \nu\mu(\mathbf{f}_k), \qquad \beta(\mathbf{f}_k) = \nu(1 - \mu(\mathbf{f}_k)), \quad (2)$$

where $\nu$ relates to the precision of the Beta distribution and is not parameterized by $f$. Finally, our novel likelihood depicted in Fig. 1 is described by

$$p(\pi_k | \mathbf{f}_k, \boldsymbol{\theta}_{\mathcal{L}}) = \text{Beta}(\pi_k | \nu\mu(\mathbf{f}_k, \sigma), \nu(1 - \mu(\mathbf{f}_k, \sigma))), \quad (3)$$

where $\boldsymbol{\theta}_{\mathcal{L}} = \{\sigma, \nu\}$ and $\mu(\mathbf{f}_k, \sigma)$ is given by

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

The precision term $\nu$ in Eq. (2) and Eq. (3) is inversely related to the observation noise on the continuous bounded responses. In general, $\nu$ can be viewed as a measure of how consistent the scale is used in a given comparison.

### 2.3. Gaussian Process Priors

At this point we have not specified any form, order or shape of $f$, but referred to $f$ as an abstract function. We maintain the abstraction by considering a non-parametric approach and use a Gaussian process (GP) to formulate our beliefs about $f$.
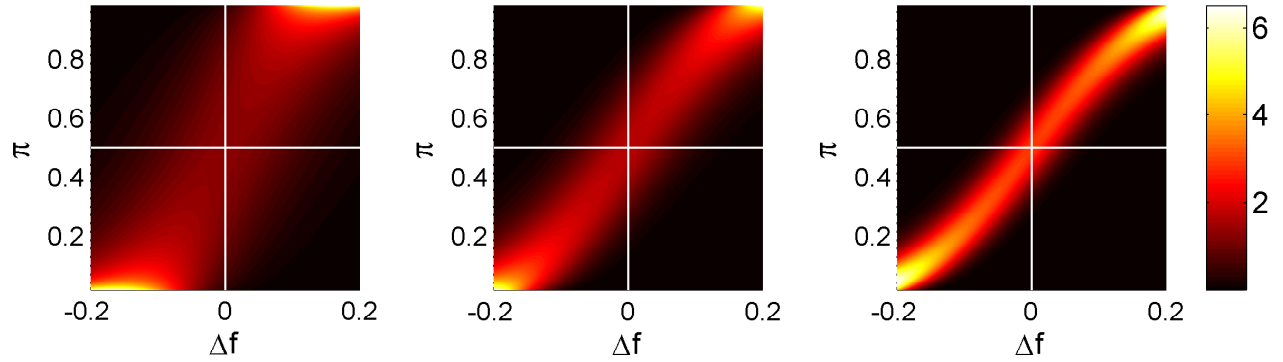
**Fig. 1**. Illustration of the proposed likelihood with $p(\pi_k|\mathbf{f}_k, \boldsymbol{\theta}_{\mathcal{L}})$ shown as a color level. The likelihood parameters $\boldsymbol{\theta}_{\mathcal{L}}$ are $\sigma = 0.1$ and left: $\nu = 3$, middle: $\nu = 10$ and right: $\nu = 30$

A GP is typically defined as "a collection of random variables, any finite number of which have a joint Gaussian distribution" [9]. Following [9] we denote a function drawn from a GP as $f(x) \sim \mathcal{GP}\left(\mathbf{0}, k(\cdot, \cdot)_{\boldsymbol{\theta}_c}\right)$ with a zero mean function, and $k(\cdot, \cdot)_{\boldsymbol{\theta}_c}$ referring to the covariance function with hyper-parameters $\boldsymbol{\theta}_c$, which defines the covariance between the random variables as a function of the inputs $\mathcal{X}$. The fundamental consequence of this formulation is that the GP can be considered a distribution over functions, i.e., $p(\mathbf{f}|\mathcal{X}, \boldsymbol{\theta}_c)$, with hyper-parameters $\boldsymbol{\theta}_c$ and $\mathbf{f} = [f(x_1), f(x_2), ..., f(x_n)]^T$, i.e., dependent on $\mathcal{X}$.

In a Bayesian setting we can directly place the GP as a prior on the function defining the likelihood. This leads us directly to a formulation given Bayes relation with $\boldsymbol{\theta} = \{\boldsymbol{\theta}_{\mathcal{L}}, \boldsymbol{\theta}_c\}$

$$p(\mathbf{f}|\mathcal{Y}, \mathcal{X}, \boldsymbol{\theta}) = \frac{p(\mathcal{Y}|\mathbf{f}, \boldsymbol{\theta}_{\mathcal{L}})\, p(\mathbf{f}|\mathcal{X}, \boldsymbol{\theta}_c)}{p(\mathcal{Y}|\boldsymbol{\theta}, \mathcal{X})}. \quad (4)$$

The prior $p(\mathbf{f}|\mathcal{X}, \boldsymbol{\theta}_c)$ is given by the GP and the likelihood $p(\mathcal{Y}|\mathbf{f}, \boldsymbol{\theta}_{\mathcal{L}})$ is either of the two likelihoods defined previously, with the assumption that the likelihood factorizes as usual, i.e., $p(\mathcal{Y}|\mathbf{f}, \boldsymbol{\theta}_{\mathcal{L}}) = \prod_{k=1:m} p(y_k|f(u_k), f(v_k), \boldsymbol{\theta}_{\mathcal{L}})$

The posterior of interest, $p(\mathbf{f}|\mathcal{Y}, \mathcal{X}, \boldsymbol{\theta})$, is directly defined when equipped with the likelihood and the prior, but it is unfortunately not of any known analytical form in either the binary nor the continuous case.

## 3. INFERENCE & PREDICTIONS

Since the likelihoods considered in this paper do not result in closed form solutions to the posterior in Eq. (4), we must resort to approximations, such as the Laplace approximation, Expectation Propagation or sampling. Since the main focus of this work is to examine the general properties of the likelihood proposed in Sec. 2.2, we use the well-know and relatively simple Laplace approximation. The required steps have previously been derived for the binary likelihood [3] (see [10]

for a detailed derivation), and in the following it will be derived for the proposed likelihood from Sec. 2.2.

### 3.1. Laplace Approximation

The main idea is to approximate the posterior by a single Gaussian distribution, such that $p(\mathbf{f}|\mathcal{Y}) \approx \mathcal{N}(\mathbf{f}|\hat{\mathbf{f}}, \mathbf{A}^{-1})$. Where $\hat{\mathbf{f}}$ is the mode of the posterior and $\mathbf{A}$ is the Hessian of the negative log-likelihood at the mode. The mode is found as $\hat{\mathbf{f}} = \arg\max_{\mathbf{f}} p(\mathbf{f}|\mathcal{Y}) = \arg\max_{\mathbf{f}} p(\mathcal{Y}|\mathbf{f})\, p(\mathbf{f})$.

The general solution to the problem can be found by considering the unnormalized log-posterior and the resulting cost function which is to be maximized, is given by

$$\psi(\mathbf{f}|\mathcal{Y}, \mathcal{X}, \boldsymbol{\theta}) = \log p(\mathcal{Y}|\mathbf{f}, \mathcal{X}, \boldsymbol{\theta}_{\mathcal{L}}) - \frac{1}{2}\mathbf{f}^T \mathbf{K}^{-1} \mathbf{f}$$
$$- \frac{1}{2}\log|\mathbf{K}| - \frac{N}{2}\log 2\pi. \quad (5)$$

where $\mathbf{K}_{i,j} = k(x_i, x_j)_{\boldsymbol{\theta}_c}$. We use a damped Newton method with soft linesearch to maximize Eq. (5). In our case the basic damped Newton step (with adaptive damping factor $\lambda$) can be calculated without inversion of the Hessian (see [10])

$$\mathbf{f}^{new} = \left(K^{-1} + \mathbf{W} - \lambda\mathbf{I}\right)^{-1}$$
$$\cdot \left[(\mathbf{W} - \lambda\mathbf{I}) - \mathbf{f} + \nabla\log p(\mathcal{Y}|\mathbf{f}, \mathcal{X}, \boldsymbol{\theta}_{\mathcal{L}})\right], \quad (6)$$

Using the notation $\nabla\nabla_{i,j} = \frac{\partial^2}{\partial f(x_i)\partial f(x_j)}$ we apply the definition $\mathbf{W}_{i,j} = -\sum_k \nabla\nabla_{i,j}\log p(y_k|\mathbf{f}_k, \boldsymbol{\theta}_{\mathcal{L}})$. We note that the term $\nabla\nabla_{i,j}\log p(y_k|\mathbf{f}_k, \boldsymbol{\theta}_{\mathcal{L}})$ is only nonzero when both $x_i$ and $x_j$ occur as either $v_k$ or $u_k$ in $\mathbf{f}_k$. In contrast to standard binary GP classification the Hessian $\mathbf{W}$ is not diagonal, which makes the approximation slightly more involved.

When converged, the resulting approximation is

$$p(\mathbf{f}|\mathcal{Y}, \mathcal{X}, \boldsymbol{\theta}) \approx \mathcal{N}\left(\mathbf{f}|\hat{\mathbf{f}}, \left(\mathbf{W} + \mathbf{K}^{-1}\right)^{-1}\right). \quad (7)$$

In the Beta case the required two first derivatives of the like-

lihood are given by:

$$\nabla_i \log p(\pi_k \mid \mathbf{f}_k, \boldsymbol{\theta}_{\mathcal{L}}) = \mathbb{I}(x_i) \cdot \nu \cdot \mathcal{N}(\mathbf{f}_k)$$

$$\cdot \left[ \log(\pi_k) - \log(1 - \pi_k) - \psi(\alpha) + \psi(\beta) \right] \text{ and } \quad (8)$$

$$\nabla\nabla_{i,j} \log p(\pi_k \mid \mathbf{f}_k, \boldsymbol{\theta}_{\mathcal{L}}) = -\mathbb{I}(x_i)\mathbb{I}(x_j) \cdot \nu^2 \cdot \mathcal{N}(\mathbf{f}_k),$$

$$\cdot \left[ \mathcal{N}(\mathbf{f}_k) \cdot \left( \psi^{(1)}(\alpha) + \psi^{(1)}(\beta) \right) + \frac{f(v_k) - f(u_k)}{2\nu\sigma^2} \right.$$

$$\left. \cdot \left( \log(\pi_k) - \log(1 - \pi_k) - \psi(\alpha) + \psi(\beta) \right) \right], \quad (9)$$

where we for convenience write $\alpha$ and $\beta$ without the dependency on $\mathbf{f}_k$ Eq. (2). $\psi(z)$ and $\psi^{(1)}(z)$ are the digamma function of zero'th and first order, respectively, $\mathcal{N}(\mathbf{f}_k) = \mathcal{N}\left( \frac{f(v_k) - f(u_k)}{\sqrt{2}\sigma} \Big| 0, 1 \right)$ and $\mathbb{I}(z)$ is an indicator function defined by

$$\mathbb{I}(z) = \begin{cases} 1 & \text{if } z = u_k \\ -1 & \text{if } z = v_k \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

We refer to [10] for a full derivation and for the required derivatives for the binary case as first described in [3].

## 3.2. Hyper-parameter Estimation

So far we have simply considered the hyper-parameters $\boldsymbol{\theta} = \{\boldsymbol{\theta}_{\mathcal{L}}, \boldsymbol{\theta}_c\}$ variables on which we can condition the primary posterior, and not worried about their values or distributions. In the following, we consider the hyper-parameters random variables on which we place a prior and the full posterior would be $p(\mathbf{f}, \boldsymbol{\theta}|\mathcal{Y})$. However, since the focus in this work is $p(\mathbf{f}|\mathcal{Y}, \mathcal{X}, \boldsymbol{\theta})$ we only use the prior on $\boldsymbol{\theta}$ to make point estimates of the hyper-parameters in terms of *maximum-a-posteriori* (MAP) estimates.

We obtain the MAP estimates by iterating between the Laplace approximation with fixed hyper-parameters, i.e. finding $p(\mathbf{f}|\mathcal{Y}, \mathcal{X}, \boldsymbol{\theta}^{\text{MAP}})$, followed by a maximization step in which $\boldsymbol{\theta}^{\text{MAP}} = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\mathcal{Y}, \mathcal{X}, \mathbf{f})$.

We first consider the standard evidence approach which seeks to optimize the marginal likelihood given by

$$p(\mathcal{Y}|\boldsymbol{\theta}, \mathcal{X}) = \int p(\mathcal{Y}|\mathbf{f}, \boldsymbol{\theta}_{\mathcal{L}}) p(\mathbf{f}|\mathcal{X}, \boldsymbol{\theta}_c) d\mathbf{f}$$

$$= p(\boldsymbol{\theta}|\mathcal{Y}, \mathcal{X}) p(\mathcal{Y}|\mathcal{X}) / p(\boldsymbol{\theta}|\mathcal{X}). \quad (11)$$

Our interest is in the posterior term, $p(\boldsymbol{\theta}|\mathcal{Y}, \mathcal{X})$, so considering Eq. (11) in terms of the log-posterior of $\boldsymbol{\theta}$ we obtain $\log p(\boldsymbol{\theta}|\mathcal{Y}, \mathcal{X}) = \log p(\boldsymbol{\theta}|\mathcal{X}) + \log p(\mathcal{Y}|\boldsymbol{\theta}, \mathcal{X}) - \log p(\mathcal{Y}|\mathcal{X})$, where $p(\boldsymbol{\theta}|\mathcal{X})$ is the prior and typical considered independent of $\mathcal{X}$. The evidence term, $\log p(\mathcal{Y}|\boldsymbol{\theta}, \mathcal{X})$, is analytical intractable in both likelihood cases, but we can approximate it using the existing Laplace approximation to obtain [10] $\log p(\mathcal{Y}|\boldsymbol{\theta}) \approx \log p(\mathcal{Y}|\hat{\mathbf{f}}, \boldsymbol{\theta}_{\mathcal{L}}) - \frac{1}{2}\hat{\mathbf{f}}^T\mathbf{K}^{-1}\hat{\mathbf{f}} - \frac{1}{2}\log|I + \mathbf{K}W|$. Now $\boldsymbol{\theta}^{\text{MAP}}$ is found by maximizing $\log p(\boldsymbol{\theta}|\mathcal{Y}, \mathcal{X})$ with respect to $\boldsymbol{\theta}$ and noting that $p(\mathcal{Y}|\mathcal{X})$ is independent of $\boldsymbol{\theta}$. We perform the optimization using a BFGS gradient method. The required derivatives and details are provided in [10].

The choice of particular priors is left for the simulations in Sec. 4, however, if $p(\boldsymbol{\theta})$ is the Uniform distribution, we obtain the traditional evidence optimization [9] as expected. It is noted that the complexity of the posterior inference is of the same order as standard GP regression described in [9].

## 3.3. Prediction

The main task is to estimate the latent function, $f$, with the end goal to do predictions of the observable variable $y$ for a pair of test inputs $r \in \mathcal{X}_t$ and $s \in \mathcal{X}_t$. In this paper, we are especially interested in the discrete decision, i.e., whether $r \succ s$ or $s \succ r$. This can be obtained from both likelihood models, thus allowing for direct comparison of the two formulations in terms of predictive performance.

We first consider the predictive distribution of $f$ which is required in both cases, and for notational convenience we omit the conditioning on $\mathcal{X}$ and $\mathcal{X}_t$. Given the GP, we can write the joint prior distribution between $\mathbf{f} \sim p(\mathbf{f}|\mathcal{Y}, \boldsymbol{\theta}^{\text{MAP}})$ and the test variables $\mathbf{f}_t = [f(r), f(s)]^T$ as

$$\begin{bmatrix} \mathbf{f} \\ \mathbf{f}_t \end{bmatrix} = \mathcal{N}\left( \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{K} & \mathbf{k}_t \\ \mathbf{k}_t^T & \mathbf{K}_t \end{bmatrix} \right), \quad (12)$$

where $\mathbf{k}_t$ is a matrix with elements $\mathbf{k}_{2,i} = k(s, x_i)_{\boldsymbol{\theta}_c^{\text{MAP}}}$ and $\mathbf{k}_{1,i} = k(r, x_i)_{\boldsymbol{\theta}_c^{\text{MAP}}}$ with $x_i$ being a training input. The conditional $p(\mathbf{f}_t|\mathbf{f})$ is obviously Gaussian as well and can be obtained directly from Eq. (12). The predictive distribution is given as $p(\mathbf{f}_t|\mathcal{Y}, \boldsymbol{\theta}^{\text{MAP}}) = \int p(\mathbf{f}_t|\mathbf{f}) p(\mathbf{f}|\mathcal{Y}, \boldsymbol{\theta}^{\text{MAP}}) d\mathbf{f}$. With the posterior approximated with the Gaussian from the Laplace approximation then $p(\mathbf{f}_t|\mathcal{Y}, \boldsymbol{\theta}^{\text{MAP}})$ will be Gaussian too and is given as $\mathcal{N}(\mathbf{f}_t|\mu^*, \mathbf{K}^*)$ with $\mu^* = [\mu_r^*, \mu_s^*]^T = \mathbf{k}_t\mathbf{K}^{-1}\hat{\mathbf{f}}$ and

$$\mathbf{K}^* = \begin{bmatrix} \mathbf{K}_{rr}^* & \mathbf{K}_{rs}^* \\ \mathbf{K}_{sr}^* & \mathbf{K}_{ss}^* \end{bmatrix} = \mathbf{K}_t - \mathbf{k}_t^T (\mathbf{I} + \mathbf{WK}) \mathbf{k}_t,$$

where $\hat{\mathbf{f}}$ and $\mathbf{W}$ are obtained from Eq. (7). With the predictive distribution for $\mathbf{f}_t$, the final prediction of the observed variable is available from

$$p(y_t|\mathcal{Y}, \boldsymbol{\theta}^{\text{MAP}}) = \int p(y_t|\mathbf{f}_t, \boldsymbol{\theta}_{\mathcal{L}}^{\text{MAP}}) p(\mathbf{f}_t|\mathcal{Y}, \boldsymbol{\theta}^{\text{MAP}}) d\mathbf{f}_t \quad (13)$$

If the likelihood is an odd function, as in both our cases, the binary preference decision between $r$ and $s$ can be made directly from $p(\mathbf{f}_t|\mathcal{Y})$. In contrast, evaluation of the integral in Eq. (13) is required for, e.g., soft decisions, reject options and sequential designs.

### 3.3.1. Binary Likelihood

If $p(\mathbf{f}_t|\mathcal{Y}, \boldsymbol{\theta}^{\text{MAP}})$ is Gaussian and we consider the Probit likelihood, the integral in Eq. (13) can be evaluated in closed form as a modified Probit function given by [3]

$$P(r \succ s|\mathcal{Y}) = \Phi((\mu_r^* - \mu_s^*)/\sigma^*) \quad (14)$$

with $(\sigma^*)^2 = 2\sigma^2 + \mathbf{K}_{rr}^* + \mathbf{K}_{ss}^* - \mathbf{K}_{rs}^* - \mathbf{K}_{sr}^*$
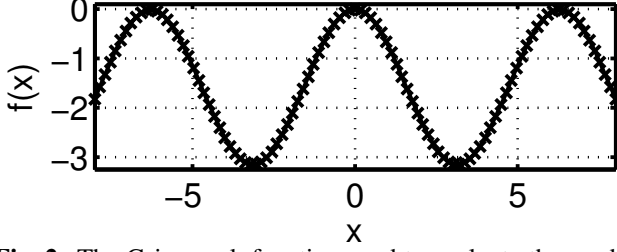
**Fig. 2**. The Griewangk function used to evaluate the predictive performance. Crosses indicate discrete samples. The center peak is slightly higher than the two others.

| Simulation | Data Noise | $\theta_{\mathcal{L}}$ | | $\theta_c$ | |
|---|---|---|---|---|---|
| | $\nu_D$ | $\sigma$ | $\nu$ | $\sigma_f$ | l |
| BR NoiseFree | No Noise | $\delta_1$ | | $\delta_{\text{ideal}}$ | $\delta_{\text{ideal}}$ |
| BR | $\{3, 10, 30\}$ | $\delta_1$ | | $\mathcal{U}_1$ | $\mathcal{U}_1$ |
| CBR NoiseFree | No Noise | $\delta_1$ | $\delta_{\to\infty}$ | $\delta_{\text{ideal}}$ | $\delta_{\text{ideal}}$ |
| CBR Ideal | $\{3, 10, 30\}$ | $\delta_1$ | $\delta_{\{3,10,30\}}$ | $\delta_{\text{ideal}}$ | $\delta_{\text{ideal}}$ |
| CBR | $\{3, 10, 30\}$ | $\delta_1$ | $\mathcal{G}(1,\eta)_{\{3,10,30\}}$ | $\mathcal{U}_1$ | $\mathcal{U}_1$ |

**Table 1**. Simulation conditions. $\delta_x$ is a point-mass, thus the parameter is constantly equal to $x$. The $\delta_{\text{ideal}}$ value is learned as $m \to \infty$. $\mathcal{U}_x$ is an uniform prior over $]0; \infty[$ with the parameter initialized to $x$. $\mathcal{G}(1,\eta)_x$ is a Gamma prior with inverse scale parameter $\eta = 0.05$ and initialization $x$.

### 3.3.2. Continuous Likelihood

In the continuous case the observed variable, $\pi$, does not directly define the discrete observation which is the main focus of this work. However, a binary preference can be derived from the continuous likelihood via the predictive distribution over $\pi$. With the suggested likelihood and mean function in Sec. 2.2 the probability of the binary choice is obtained as $P\left(r \succ s | \mathcal{Y}, \theta_{\mathcal{L}}\right) = \int_{\pi=0}^{\pi=1/2} p\left(\pi_t | \mathcal{Y}, \theta_{\mathcal{L}}\right) d\pi_t$, thus

$$P\left(r \succ s | \mathcal{Y}, \theta^{\text{MAP}}\right) =$$
$$\int p(\mathbf{f}_t | \mathcal{Y}, \theta^{\text{MAP}}) \text{Betacdf}\left(\frac{1}{2} \bigg| \alpha(\mathbf{f}_t), \beta(\mathbf{f}_t)\right) d\mathbf{f}_t \quad (15)$$

In the ideal case of a noise-free user, i.e., $\nu \to \infty$, the Beta distribution reduces to a point mass at the mean defined by the Probit function. Hence, in the limit of a completely consistent user, the predictions from Eq. (15) reduces to a classical choice model with predictions that follows Eq. (14).

## 4. EXPERIMENTAL RESULTS AND DISCUSSION

To study the performance of the models in a controlled setting, we use a synthetic dataset generated from the deterministic *Griewangk function* depicted in Fig. 2. We use the predictive performance of the binary decision to compare the learning rates of the binary response (BR) model as the baseline and the continuous bounded response (CBR) model. In each comparison, the two inputs are drawn randomly among 101 input points sampled uniformly from $x = [-8; 8]$.

The training points $\pi_k$ are drawn from a Beta distribution with the parameterization from Sec. 2.2 with the Probit link function in Eq. (4), $\sigma = 1$, and the Griewangk function values as the two inputs. The noise level on the training data is defined by the parameter $\nu_D$ corresponding to $\nu$ in the CBR model. The binary decision $d_k$ is determined by whether $\pi_k$ is smaller or larger than 0.5. For evaluation, we generate an independent binary test set located equidistantly in between the training points. Initial experiments showed that in order to get a robust predictive model for all noise level, it is important to learn the $\nu$ parameter in the CBR model. The

initial experiments also indicated that it is vital not to underestimate the noise, while an overestimation is not as crucial and provides overall good predictive performance. This suggests a prior with a monotonic increasing likelihood towards the highest noise level. A natural choice is a Gamma(1,$\eta$) prior with inverse scale parameter $\eta$.

The considered models, priors and parameters are listed in Table 1 where the covariance parameters, $\theta_c$, are applied in a GP prior with a covariance function defined by the *squared exponential* kernel $k_{SE}(x,x') = \sigma_f^2 \exp(-l^{-2}\|x - x'\|^2)$. When a specific prior is not a point-mass/constant indicated by $\delta_x$ in Table 1, the hyper-parameters are estimated (MAP) either for each training set size (realistic scenario) or for $m = 500$ (ideal scenario). The latter is indicated by $\delta_{\text{ideal}}$.

The learning curves from Fig. 3 show that under ideal conditions with nearly noise-free observations and a correct noise setting (Fig. 3, right plot) the CBR model outperforms the BR models as expected, since a continuous response will essentially provide more information from each experiment under ideal conditions than a binary response will. Also, in both high and moderate noise conditions (Fig. 3, left and middle plot) the CBR model with a correct noise setting (CBR Ideal) outperforms the corresponding BR model significantly in terms of learning rates and actually shows similar learning rates as the BR model under noise-free conditions. Finally and most importantly, the learning rates are only slightly lower when $\nu$ has been inferred from data via the MAP procedure (with different initializations) than when it is specified correctly, which suggests that the parameter inference framework with independent priors is robust in real-life-scenarios without ideal model and noise conditions.

We have focused on a controlled example to highlight properties of the model and inference, leaving a real-world validation for future work. Future work also includes the extension of the mean function, Eq. (4), using a mixture of Probit functions to account for different user behavior such as centering and contraction bias. For a real-world setting, a natural extension is a suitable active learning criteria, such as the *expected value of information* framework applied recently in e.g. [5] for the BR model.
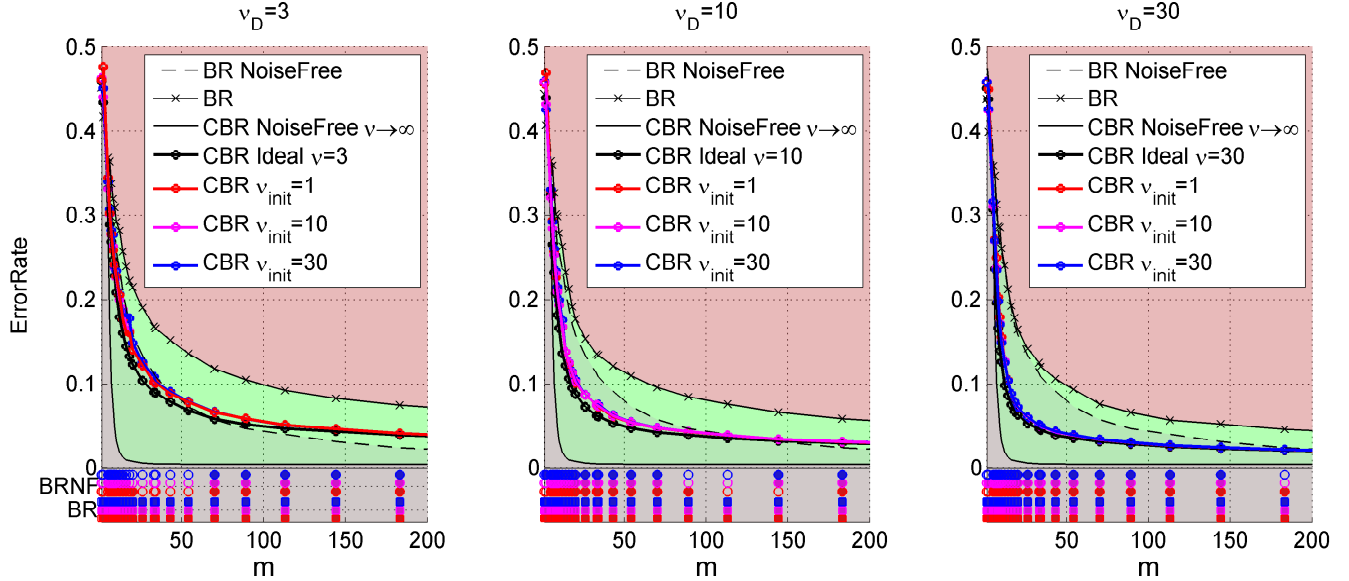
**Fig. 3**. Mean error test rates (MER) as a function of the number of experiments over 100 different realizations of the training set generated with different $\nu_D$. In the red and top green area MER are worse and better, respectively, than those obtained with the BR model on the noisy data. In the lower green area MER are also better than those obtained by the BR NoiseFree, and finally, the grey area corresponds to unrealistic MER better than those obtained with a CBR NoiseFree model with $\nu \to \infty$ evaluated with $\nu = 10^3$ on a noise-free data set. The six rows of markers indicate if the MER of the corresponding CBR model are significantly different from those resulting from the BR (squares) and from the BR NoiseFree (circles). If solid, the zero-hypothesis of the two means being equal is rejected at the 5% level using a paired t-test.

## 5. CONCLUSION AND PERSPECTIVES

We have proposed a new model for preference learning with Gaussian Process priors with the main purpose to increase the learning rate compared to the standard binary model applied in [3]. We have outlined a robust and flexible inference framework for the new model based on suitable priors and the Laplace approximation. Simulations were used to present properties and performance, which showed a significant information increase from each experiment under ideal conditions as expected but more importantly also under adverse conditions. The performance is especially increased in a certain window of opportunity.

## 6. REFERENCES

[1] 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.

[2] L. L. Thurstone, "A law of comparative judgement.," *Psychological Review*, vol. 34, 1927.

[3] 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.

[4] P. Groot, T. Heskes, T. Dijkstra, and J. Kates, "Predicting preference judgments of individual normal and hearing-impaired listeners with Gaussian Processes," *IEEE Transactions on Audio, Sound, and Language Processing*, 2010.

[5] 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.

[6] 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.

[7] R. D. Bock and J. V. Jones, "The measurement and prediction of judgment and choice.," 1968.

[8] S. Ferrari and F. Cribari-Neto, "Beta Regression for Modelling Rates and Proportions," *Journal of Applied Statistics*, vol. 31, no. 7, pp. 799–815, Aug. 2004.

[9] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, MIT Press, 2006.

[10] Bjørn Sand Jensen and Jens Brehm Nielsen, *Pairwise Judgements and Absolute Ratings with Gaussian Process Priors*, Technical Report, DTU Informatics, September 2011.