No-regret Algorithms for Capturing Events in Poisson Point Processes

Mojmír Mutný [1] **Andreas Krause** [1]

Abstract

Inhomogeneous Poisson point processes are widely used models of event occurrences. We address *adaptive sensing of Poisson Point processes*, namely, maximizing the number of captured events subject to sensing costs. We encode prior assumptions on the rate function by modeling it as a member of a known *reproducing kernel Hilbert space* (RKHS). By partitioning the domain into separate small regions, and using heteroscedastic linear regression, we propose a tractable estimator of Poisson process rates for two feedback models: *count-record*, where exact locations of events are observed, and *histogram* feedback, where only counts of events are observed. We derive provably accurate anytime confidence estimates for our estimators for sequentially acquired Poisson count data. Using these, we formulate algorithms based on optimism that provably incur sublinear count-regret. We demonstrate the practicality of the method on problems from crime modeling, revenue maximization as well as environmental monitoring.

1. Introduction

Poisson point processes are widely used for modeling discrete event occurrences, with a rich history in probability theory (Kingman, 2002). Their modeling power is demonstrated by a multitude of applications in engineering (Snyder and Miller, 2012), crime rate modeling (Shirota et al., 2017), measuring genetic biodiversity (Diggle et al., 2013), environmental monitoring (Heikkinen and Arjas, 1999) and pollutant modeling (Diggle, 2013). In this work, we design a sequential algorithm for *adaptive sensing of spatio-temporal Poisson processes*. Our approach adaptively focuses on particular regions of the space – *sensing regions* – in order to maximize the number of events captured over time subject to the costs of sensing. Crucially, we allow the rate of event occurrence to *vary smoothly* over the sensed domain. In particular, we assume the rate function can be modeled as a member of a *reproducing kernel Hilbert space (RKHS)* with a known kernel k.

As a running example, consider surveying spontaneous taxi requests on a street, e.g., by a visual gesture. Recorded requests over multiple days are depicted in Fig. 1a, along with the inferred intensity. The rate function in this case can be modeled as smooth function. For simplicity, suppose that the intensity is time invariant. As operators, we might be interested in servicing as many requests as possible by allocating taxi cars to different areas. In this application, a *sensing action* corresponds to such an allocation of sending taxis to a certain region of a city and waiting there for requests. The *sensing regions* are sets arising from hierarchically splitting the domain (cf. Fig. 1a), corresponding roughly to small city blocks. The cost of each action – sensing region – can be reasonably chosen to be proportional to the area of the sensing as it takes more taxi drivers to cover a larger area.

2. Background

Let $\mathcal{D} \subset \mathbb{R}^d$ be a compact subset. An *inhomogeneous spatio-temporal Poisson point process* $\mathfrak{P}$ is a random process such that for any subset $A \subset \mathcal{D}$, $N(A)$ denotes the random variable representing the number of events in A for a unit duration and $n(A)$ a realization of this random variable. If $A, B \subset \mathcal{D}$ and $A \cap B = \emptyset$ then $N(A)$ and $N(B)$ are independent. In addition, the process has an associated *intensity function* $\lambda(x)$, where the number of events sensed during unit time in A is distributed as,

$$N(A) \sim \text{Poisson}\left(\int_A \lambda(x)dx\right). \quad (1)$$

These conditions fully specify a *Poisson process* (Snyder and Miller, 2012).

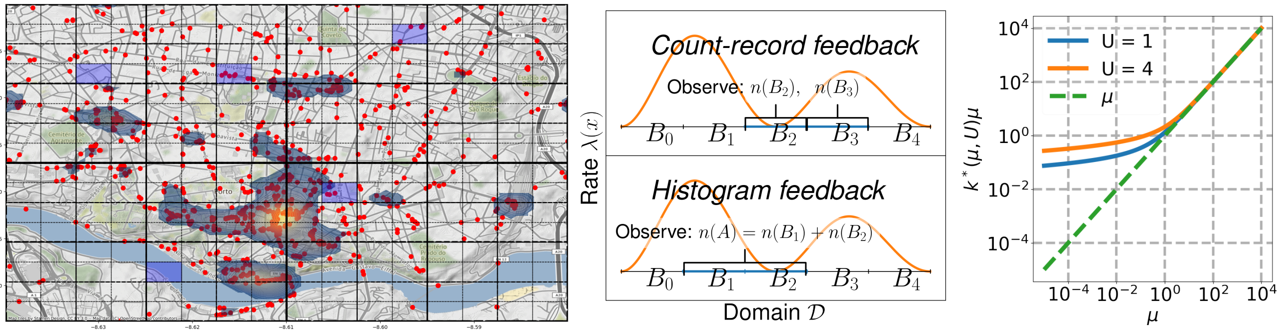
2.1. Sensing Problem: Capturing Events

We consider the problem of adaptive sensing of Poisson processes – capturing of as many events of $\mathfrak{P}$ as possible subject to the cost of sensing. We act sequentially by picking *sensing regions*, obtaining some *feedback* about the process within them at some *cost*, with the goal minimize *count-regret*. Below, we elaborate on these aspects.

Actions In our abstract framework, *sensors* can be placed to monitor a subset of the domain $\mathcal{D}$ for a certain duration

[1]Department of Computer Science, ETH Zurich, Zurich, Switzerland. Correspondence to: Mojmír Mutný <mojmir.mutny@inf.ethz.ch>.

Proceedings of the 38^{th} International Conference on Machine Learning, PMLR 139, 2021.



(a) (b) (c)

Figure 1. a) The estimated rate of taxi requests along with raw observational data in red, as well as a quadtree sensing structure generating rectangular sensing regions. The blue regions are examples of sensing regions - only events inside it are observed. b) An example of binning with $\mathcal{B}$: An illustrative rate function is depicted in orange. The difference between the two feedback models is demonstrated in selected regions. c) This figure refers to Section 5. We depict the values of scaled variances due to the heavier tail of Poisson random variables, as a function of the true variance. To construct estimators with confidence set guarantees we need to normalize the residuals of linear regression by scaled variances instead of true variances as is customary. The relation depicts scaled variances $k^*\mu$ as function of true variance μ. We see that as μ gets larger, no scaling is necessary since Poisson tail behaves Gaussian-like. This scaling depends on norm of feature vectors U.

Δ. We refer to such a sensing event as an *action*. The set of actions is parametrized by a collection of sets $\mathcal{A}$, where to simplify the exposition we assume $|\mathcal{A}| < \infty$.

Feedback models In each iteration, we pick a sensing action A, i.e., a measurement region and a duration Δ. We then receive feedback about the events which occurred in A only. We distinguish two important feedback modes:

- *Count-record feedback*: We obtain the *number* of events $n(A)$, and the *locations* of events $\{x_i\}_{i=1}^{n(A)}$ in A during period Δ.
- *Histogram feedback*: We only receive the *number* of events in A, $n(A)$ for the duration period Δ. [1]

The count-record feedback has been investigated in the context of geostatistics with passively collected data, and health surveillance (Diggle et al., 2013). The histogram feedback can be motivated, e.g., by limitations of the sensing technology, or by privacy reasons, as sometimes it might not be desirable to reveal, e.g., locations of sick individuals. In the taxi pickups example, the taxi driver might not be collecting the exact locations of passengers, but rather just counting the number of passengers in the current city block.

Costs Each of the actions A has an associated *cost* $w_t(A, \Delta)$, where the cost function $w_t : \mathcal{A} \to \mathbb{R}$ is known and potentially time dependent. A reasonable assumption in practice is that the cost is time uniform and $w(A, \Delta) = w(|A|, \Delta)$ depends on the volume (area) of the sensed set and the duration of sensing. We will fix $\Delta_t = \Delta$ to be a fixed minimum sensing time duration. Should a longer sensing duration be desired, the action can be repeated if the costs permit it. Now, we mention three important classes of costs: a) *uniform costs:* $w(A) \propto |A|$, a natural notion modeling that the cost of monitoring a sensing region A is proportional to its size. b) *Concave costs:* $w(A) \propto g(|A|)$, where g is monotone concave. This class of costs allows to model, e.g., fixed setup costs, and puts preference over larger sets. Without concave cost structure, sensing the smallest regions of $\mathcal{A}$ is always preferred over any other strategy using larger sets. When including, e.g., a fixed cost to sensing, it can be beneficial to first sense large sets before narrowing down to promising regions. Lastly, we introduce c) *one-off costs:* $w_t(A) = |A| + \infty \times \mathbb{I}_{A = A_i i \leq t}$, where $\mathbb{I}$ is an indicator function of whether we sensed that region already. The cost is infinite for any region that we already sensed. This covers a practically important scenario, where regions cannot be sensed multiple times.[2]

Count-Regret We want to spend our budget in the most effective way by minimizing the following *count-regret*,

$$R_T := \sum_{t=1}^{T} \Delta_t \Big(w_t(A_t, \Delta_t) \frac{\mathbb{E}[N(A_t^*)]}{w_t(A_t^*, \Delta_t)} - \mathbb{E}[N(A_t)] \Big), \tag{2}$$

where R_T is a function of $\{(A_t, \Delta_t)_{t=1}^T\}$. This regret represents the average number of missed events due to investment of the budget $\{w_t(A_t)\}_{t=1}^T$ to potentially suboptimal actions, in terms of their expected count to cost ratio. The above expression (2) simplifies significantly when cost is proportional to the sensed duration Δ_t – which is why we use it in our work. In case of the *one-off* costs,

[1] Terminology from Snyder and Miller (2012).

[2] A similar cost is studied in the context of valuable item discovery in Vanchinathan et al. (2015), but with a different regret measure and sub-Gaussian noise.

the best sensing action at time t, A_t^* depends on the prior choices of the algorithm. In that sense, with this cost, we measure whether eventually the future choices are made in a cost-effective way. For uniform and concave costs it is time invariant, i.e., $A_t^* = A^*$.

Challenges The goal of this work is to design a sequential algorithm that is *no-regret*, meaning that $R_T/T \to 0$ as $T \to \infty$, and exploits the regularity that the RKHS assumption imposes. There are three technical challenges that are specific to our setting: **1)** the intensity function λ needs to be positive everywhere in $\mathcal{D}$ – a constraint we need to efficiently enforce. **2)** The Poisson random variables are naturally heteroscedastic, since their mean and variance are equal – which we need to take into account. **3)** In order to prove the no-regret property, we need to design confidence sets that deal with the tail of the Poisson distribution which is heaver than that of any sub-Gaussian random variable, for which adaptive confidence sets are known.

2.2. Contributions

1) We show that approaches to *enforce positivity* via finite constraints such as positive bases or positive trace-regression can be treated jointly in our framework of Poisson sensing. **2)** By carefully designing an appropriate super-martingale process, we are the first to derive *valid anytime confidence sets for Poisson responses for heteroscedastic linear regression*. **3)** Using these estimators and confidence sets, we formulate an optimistic algorithm CAPTURE-UCB. It can be used with *histogram* as well as *count-record* feedback, and both *provably suffer count-regret bounded by* $\tilde{\mathcal{O}}(\sqrt{T})$ (up to log factors) for finite rank kernels, and we propose an extension for general kernels. **4)** We demonstrate the *broad applicability* of our algorithms on several practically relevant problems.

3. Modeling and Algorithmic protocol

Binning To address the problem, we adopt a regression perspective, where we regress on the observed counts that fall into separate parts of the domain. This treatment does not permit us to directly use the exact locations of the event occurrences in the *count-record* feedback. Instead, we partition $\mathcal{D}$ to a set $\mathcal{B}$ of small disjoint regions, obtained, e.g., by hierarchically splitting the domain (as in Fig. 1a). Then, we model the counts falling into each $B \in \mathcal{B}$ as $N(B) \sim \text{Poisson}(\int_B \lambda(x)dx)$. With *count-record* feedback, we obtain multiple observations in each sensing round, for each $B \subset A$ where $B \in \mathcal{B}$. To denote the observed number of counts, instead of random variables, we use the notation $n(A) = \sum_{B \subset A} n(B \cap A)$.

Count-record feedback is often referred to as *semi-bandit feedback* in bandit literature, which provides superior information to just observing $n(A)$, often denoted *bandit feedback* to differentiate them. In Figure 1b, we illustrate the two different feedback models with partition $\mathcal{B}$. The size of the partition $\mathcal{B}$ influences the information we can extract from the location data, but as we will see in later sections it cannot be chosen arbitrarily small. For now, assume a partition is given.

Estimators Suppose that there are t sensing sessions during which we sensed regions A_j and observed counts $n(A_j)$ for $j \in [t]$. With the count-record feedback, we use the regression estimate $\hat{\lambda}$ as:

$$\operatorname*{arg\,min}_{\lambda \in \mathcal{H}_k, \lambda \geq 0} \sum_{j=1}^{t} \sum_{B \subset A_j} \frac{(n(B \cap A_j) - \int_B \lambda(x)dx)^2}{\sigma_j(B)^2} + \frac{\gamma}{2} \|\lambda\|_k^2,$$

where $\lambda \geq 0$ denotes the positivity constraint, $\|\lambda\|_k$ is the associated norm to $\mathcal{H}_k$, γ is the penalization constant, and σ_j are the estimated variances that we will describe in later sections. In the absence of the location feedback, we only obtain tuples of A and their counts $n(A)$. The estimator $\hat{\lambda}$ is the solution to:

$$\operatorname*{arg\,min}_{\lambda \in \mathcal{H}_k, \lambda \geq 0} \sum_{i=1}^{t} \frac{(\int_{A_i} \lambda(x)dx - n(A_i))^2}{\sigma_i^2} + \frac{\gamma}{2} \|\lambda\|_k^2. \quad (3)$$

Poisson processes have an associated likelihood. The *penalized likelihood* estimator might thus seem to be an appropriate estimator, since it enforces the property that the mean prediction of $N(A)$ is equal to its variance. However, as we explain in Section 5.2, confidence sets for these estimators are difficult to control to facilitate regret analysis.

Algorithm Protocol We design an algorithm that first estimates the rate function $\hat{\lambda}$ using the estimator in (3). Then, based on a utility function, it chooses the most suitable sensing region and subsequently receives feedback in form of observations. As we are interested in maximizing the number of counts subject to the cost, we adopt the principle of *optimism* (Auer, 2002; Srinivas et al., 2010), and design an algorithm which optimizes a $(1 - \delta)$ confidence upper bound on the number of counts per cost. We call the algorithm CAPTURE-UCB and report it in Algorithm 1. In order to apply the algorithm, we need to perform efficient estimation, as well as construct $1 - \delta$ confidence sets, which we address in the following sections.

Organization We organize our contributions as follows: in Section 4, we explain how we efficiently enforce positivity; in Section 5, we design a novel way to construct confidence set for heteroscedasdic linear regression with Poisson counts; and lastly in Section 6, we provide a regret analysis of CAPTURE-UCB followed by experimental results.

Throughout this work, we utilize a technical assumption that the kernel can be approximated by a finite rank kernel up

Algorithm 1 CAPTURE-UCB

Require: cost function w_t, partition (schedule) $\mathcal{B}$, $t = 1$
1: **while** $t \leq T$ **do**
2: Estimate $\hat{\lambda}$ as in (3)
3: $A_t = \arg\max_{A \in \mathcal{A}} \frac{\text{ucb}_t(A)}{w_t(A)}$ as in (10)
4: Sense A_t for time Δ
5: Receive: $\begin{cases} n(B \cap A_t) \text{ for } B \in \mathcal{B} & \text{if } \textit{count-record} \\ n(A_t) & \text{if } \textit{histogram} \end{cases}$
6: $t = t + 1$.
7: **end while**

to an arbitrary precision. We make this assumption for two reasons: a) in order to enforce positivity, we use estimators which are of this form; and b) the mathematical tools used to prove confidence sets require Euclidean structure.

General RKHSs can be treated by picking a sufficiently large basis which incurs negligible approximation error in comparison to the desired optimization accuracy. We provide a more detailed discussion in Section 6.

Assumption 1 (Finite rank representation)**.** *We assume that $\lambda \in \mathcal{H}_k$, where $\lambda : \mathcal{D} \to \mathbb{R}^+$, where $\mathcal{D} \subset \mathbb{R}^d$, λ has bounded norm $\|\lambda\|_k \leq u$ for a known value u. Also, $\lambda(x) \geq l > 0$, where can l be arbitrarily small. Further, there exists a basis $\{\phi_k\}$ s.t. kernel k has finite rank as $k(x, y) = \sum_{k=1}^m \phi(x)_k \phi_k(y)$, and $\sqrt{\sum_k^m \phi_k(x)^2} \leq U$ for all x with U known.*

4. Efficient Estimators

Efficient estimators of the rate function need to impose the positivity constraint in a tractable form, allow us to efficiently perform the integrations $\int_A \lambda(x)dx$, and allow us to construct and optimize over confidence sets. We present two variants that fulfill this goal that we found practical.

4.1. Linear Model with Positive Basis

Our first estimator category are linear models, i.e., $f \approx \sum_{i=1}^m \phi_i(x)^\top \theta_i$, where $f \in \mathcal{H}_k$ can be represented up to a truncation error by positive basis functions $\phi(x) \geq 0$ for all $x \in \mathcal{D}$ and positive parameters $\theta_i \geq 0$. They satisfy that as $m \to \infty$ the approximation error decreases. With this basis, the positivity can be simply enforced by the constraint $\theta \geq 0$. There are many examples of such bases with Bernstein polynomials and positive Hermite splines being perhaps the oldest (Papp and Alizadeh, 2014). In this work, we use the positive basis due to Maatouk and Bay (2016); Cressie and Johannesson (2008) which we call *triangle* basis.

Triangle basis This representation is merely a finite basis approximation to the original kernel k, which can be made as accurate as we need by increasing the basis size m. We follow the exposition of López-Lopera et al. (2019). Given a domain $\mathcal{D} = [-1, 1]$, we create a finite dimensional approximation $\lambda = \Phi(x)^\top \theta$ where $\Phi(\cdot) \in \mathbb{R}^m$. Let us define the individual triangle basis functions ϕ_j for $j \in [m]$:

$$\phi_j(x) = \begin{cases} 1 - |\frac{(x-t_j)m}{2}| & \text{if } |\frac{(x-t_j)m}{2}| \leq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

In order to enforce the same norm constraint under the approximated basis, $\Phi(x)^\top \Phi(y)$ needs to be close to the kernel $k(x, y)$. This can be enforced by transforming the above basis functions to obtain $\Phi(x) = \mathbf{\Gamma}^{1/2}\phi(x)$, with $\mathbf{\Gamma}_{ij} = k(t_i, t_j)$, where t_i for $i \in [m]$ corresponds to the nodes. This basis has the property that at $x = t_i$, $i \in [m]$, the approximation is exact while in between the nodes t, it linearly interpolates. The constraints $l \leq \lambda(x) \leq u$ can be represented as linear constraints on θ: $l \leq \mathbf{\Gamma}^{1/2}\theta \leq u$. Cartesian products of the basis generalize this construction to arbitrary dimension, albeit increasing the basis size exponentially with the dimension. Due to the simple form of the basis, the integrals $\varphi_A := \int_A \Phi(x)dx$ can be easily evaluated.

4.2. Positive Semi-definite Trace Regression Model

Alternatively, we can stipulate $\lambda(x) = \text{Tr}(\psi(x)\psi(x)^\top \mathbf{\Theta})$, where $\mathbf{\Theta}$ is a linear map $\mathcal{H}_{\tilde{k}} \to \mathcal{H}_{\tilde{k}}$ and $\psi(x) : \mathbb{R}^d \to \mathcal{H}_{\tilde{k}}$ and $\psi(x)^\top$ its adjoint. The positivity of the responses is enforced by requiring that $\mathbf{\Theta}$ is a positive semi-definite operator, a model recently introduced by Marteau-Ferey et al. (2020). To reconcile this assumption with RKHS membership notice that λ is a member of $\mathcal{H}_k$, where $k = \tilde{k}^2$ for finite rank kernels (see Lemma 12 in Appendix).

Employing a finite basis for the kernel, we can represent $\lambda(x) = \psi(x)^\top \mathbf{\Theta} \psi(x)$, where $\mathbf{\Theta} \succeq 0$. Notice that, as with the positive basis, the expected counts can be expressed in a closed form as $\mathbb{E}[N(A)] = \text{Tr}\left(\left(\int_A \psi(x)\psi(x)^\top dx\right) \mathbf{\Theta}\right) := \text{Tr}\left(\Psi_A \Psi_A^\top \mathbf{\Theta}\right)$. The advantage of this formalism is that the basis ψ does not need to be positive as in the previous case; the disadvantage is that it increases the parameters quadratically to m^2. Hence, optimal description bases such as Mercer decomposition, or in the case of stationary kernels, optimal Fourier bases are appropriate choices which have very low m. Since the trace is a linear operator, it can be represented as a matrix operator via vectorization, $\text{Tr}\left(\Psi_A \Psi_A^\top \mathbf{\Theta}\right) = \text{vec}(\Psi_A \Psi_A^\top)^\top \text{vec}(\mathbf{\Theta})$ and its treatment is identical to the above estimator with the only difference in the form of the constraint.

5. Confidence Sets

We now derive ellipsoidal adaptive confidence sets for the estimator in (3). The analysis easily extends to the *count-record* estimator by substitution. We rely on the theory of self-normalized martingales, and identify a super-martingale M_t that upper-bounds the error process generating residuals in our estimation problem and subsequently apply Ville's inequality:

$$P\left(\sup_t \log(M_t) \geq \log\left(\frac{\mathbb{E}[M_t]}{\delta}\right)\right) \leq \delta. \quad (5)$$

This is a common technique for establishing adaptive confidence sets (de la Peña et al., 2009; Howard et al., 2018).

Optimality Conditions Let us first start by showing that under the conditions of Assumption 1 and using the shorthand $\varphi_{A_i} := \Delta \int_{A_i} \phi(x)dx$, the optimality conditions of the estimator (3) can be written in the following form:

$$(\mathbf{X}\mathbf{\Sigma}^{-1}\mathbf{X}^\top + \mathbf{I}\gamma)\hat{\theta} - \mathbf{X}^\top\mathbf{\Sigma}^{-1}n + \xi(x)^\top\bar{\mathbf{\Lambda}} = 0 \quad (6)$$

$$\xi^\top(x)\bar{\mathbf{\Lambda}}\theta = 0 \text{ for all } x \in \mathcal{D}. \quad (7)$$

Specifically for *histogram feedback*, $\mathbf{X}_{ij} = (\varphi_{A_i})_j$, $\mathbf{\Sigma}$ is diagonal with σ_i^2 on diagonal, $n_i = n(A_i)$. The operator $\bar{\mathbf{\Lambda}} : \mathbb{R}^m \to \mathbb{R}$, and $\xi(x)$ the dual variable are associated with satisfaction of the constraint $\lambda(x) \geq 0$. To facilitate our analysis, we relate $\bar{\mathbf{\Lambda}}$, an infinite dimensional operator, to $\mathbf{\Lambda}$, a finite dimensional one, by the following assumption. Both estimators in Section 4 satisfy these conditions as it is either evident or we show in Appendix A.4.

Assumption 2 (Constraints)**.** *The optimality condition in* (7) *can be imposed by $p < \infty$ linear constraints. Namely, $\xi^\top\mathbf{\Lambda}\theta = 0$, where $\xi \in \mathbb{R}_+^p$ and $\mathbf{\Lambda} \in \mathbb{R}^{p\times m}$.*

5.1. Self-normalized Approach: Super-martingale

By manipulating (6) and using the model $n = \mathbf{X}\theta + \epsilon$, we see that in order to derive ellipsoidal confidence sets, we need to control the magnitude of the self-normalized noise term $\|S_t\|_{\mathbf{V}_t^{-1}}$, where

$$S_t := \sum_{i=1}^{t-1} \frac{\varphi_{A_i}}{\sigma_i^2}\epsilon_i \text{ and } \mathbf{V}_t = \sum_{i=1}^{t-1} \frac{\varphi_{A_i}\varphi_{A_i}^\top}{\sigma_i^2} + \gamma\mathbf{I}. \quad (8)$$

The random variable ϵ_i is zero mean Poisson noise and σ_i^2 are the scaling variances in heteroscedastic linear regression. If $\exp(\|S_t\|_{\mathbf{V}_t^{-1}})$ were a super-martingale, we could use Ville's inequality (5) to obtain high probability bounds. In order to show this, we split the problem into two parts. First we prove that for $M_t(x) := \exp(\langle x, S_t\rangle - \frac{1}{2}\|x\|_{\mathbf{V}_t}^2)$ for any x s.t. $\|x\| \leq 1$ is a super-martingale, and by choosing an extremal x over the set $\|x\| \leq 1$, $M_t(x)$ can be related to $\exp(\|S_t\|_{\mathbf{V}_t^{-1}})$. This step is usually addressed by pseudo-maximization or a covering argument.

We present the solution to the first problem. It turns out that $M_t(x)$ is a super-martingale as long as we choose the variances in our estimation, σ_i^2, in a careful way:

Lemma 1. *Let $\epsilon_i = z_i - \mu_i$, where $z_i \sim$ Poisson(μ_i), where μ_i is the rate, $\|\varphi_{A_i}\|_2 \leq U \ \forall\ i \in [t]$, then there exists a finite $k_i(U/\mu_i) \geq 1$ such that when $\sigma_i^2 \geq k_i\mu_i$, and $\mathbb{E}[\sigma_i^2|\mathcal{F}_{i-1}] = \sigma_i^2$, $M_t(x)$ is super-martingale for $\|x\|_2 \leq 1$, with $M_0 = 1$. Let $c_i = \frac{U}{\mu_i}$*

The smallest such k_i^* can be found by solving:

$$\min_{k\geq 1} k \text{ s.t. } \frac{1}{2k}\left(-c_i^2 - 2c_i\right) + \left(\exp\left(\frac{c_i}{k}\right) - 1\right) \leq 0.$$

Notice that each variance is scaled individually, depending on μ_i. In Figure 1c, we can see the value of k^* for different U over the ground-truth variance μ (purple). For large enough μ, no increase is needed, thus $k^* \approx 1$. This is intuitive, since the excess kurtosis of the Poisson distribution decreases as $\frac{1}{\mu}$. In Appendix B, we provide a detailed analysis of limiting values of k^* detailing the intuition from Fig. 1c.

The above Lemma is the key to finding anytime confidence sets via pseudo-maximization. The difference to the classical treatment (see Szepesvari and Lattimore (2019)) arises because our super-martingale $M_t(x)$ is defined only for *bounded* $\|x\|_2 \leq 1$ instead on the whole $\mathbb{R}^m$ – a challenge already encountered by de la Peña et al. (2009) and Faury et al. (2020). By using arguments as in Faury et al. (2020), we can show anytime confidence sets for our estimator.

Theorem 1. *Let $\delta \in (0, 1)$, under Assumption 1 and 2, the solution to* (6) *satisfies for all $t \geq 0$*

$$\left\|\hat{\theta} - \theta\right\|_{\mathbf{V}_t}^2 \leq \beta_t(\delta) := \text{bias}(\xi, \gamma) + \zeta_t(\delta) \quad (9)$$

with probability $1 - \delta$, where $\mathbf{V}_t = \sum_{i=1}^t \frac{\varphi_{A_i}\varphi_{A_i}}{\sigma_i^2} + \gamma\mathbf{I}$,

$$\zeta_t(\delta) = \left(\frac{\sqrt{\gamma}}{2} + \frac{2}{\sqrt{\gamma}}\log\left(\frac{1}{\delta}\frac{\det(\mathbf{V}_t)^{1/2}}{\det(\gamma\mathbf{I})^{1/2}}\right) + \frac{2m}{\sqrt{\gamma}}\log(2)\right)^2,$$

$\text{bias}(\xi, \gamma) = \gamma\|\theta\|_2^2 + \left\|\xi^\top\mathbf{\Lambda}\right\|_{\mathbf{V}_t^{-1}}^2$, σ_t^2 is such that $\sigma_t^2 \geq k^*\mu_t$ as in Lemma 1, $\mathbb{E}[\sigma_t^2|\mathcal{F}_{t-1}] = \sigma_t^2$.

In contrast to the sub-Gaussian concentration results of Abbasi-Yadkori et al. (2011), the constant in the confidence parameter $\beta(\delta)$ has an explicit dimensionality dependence and an extra power of $\log\det(\mathbf{V}_t)$. In fact, a more general result can be shown where these two factors can be balanced slightly, which we do not explore here, but provide in Appendix. We conjecture that the explicit dependence on m is an artefact of the anytime proof technique. We provide more discussion in Section 6 regarding the necessity of the dependence on m. In order to recover overall scaling of $\sqrt{\beta(\delta)}$, which appears in the regret proofs with sub-Gaussian noise assumption, quantitatively similar to Abbasi-Yadkori et al. (2011)'s $\mathcal{O}(\sqrt{m}(\log t))$, we need to set $\gamma = m$. Proofs can be found in Appendix B.

Notice that the μ_i, which are needed in order to rescale the confidence ellipsoid $\mathbf{V}_t$, are *not known*. Therefore in order to use the above theorem, we need to estimate them. As the mean and variance prediction are identical for Poisson responses, and the scaled variance σ_i is an increasing function of μ_i (Lemma 7 in Appendix), we take a pragmatic approach and use the upper bound on the mean to estimate the scaling variance:

$$\sigma_t^2(A) = \text{ucb}_{t-1}(A)k^*(\text{ucb}_{t-1}(A)).$$

At time $t=0$, we use the worst case estimates $\text{ucb}_0(A) = u\int_A dx$. Notice that the sequence of variances needs to be deterministic when conditioned on the past data $\mathcal{F}_{t-1}$. The above upper-bound fulfills this condition since the estimates from previous round $(t-1)$ are used. The positivity constraints $\mathbf{\Lambda}$ appears as bias similarly as other sources of bias appear in related works studying confidence sets (Mutný et al., 2020).

5.2. Likelihood and Likelihood Ratios

Confidence sets for adaptively collected data with known likelihood function can be calculated with a similar procedure as above by noting that the likelihood ratio is a super-martingale under the true distribution (Robbins et al., 1972). Consequently, the $(1-\delta)$ confidence set $C(\theta)$ can be constructed as $C(\theta) = \left\{\theta : \sum_{i=1}^t \log\left(\frac{p(\hat{\theta}_{i-1},(\varphi_{A_i},n_i))}{p(\theta,(\varphi_{A_i},n_i))}\right) \leq \log\left(\frac{1}{\delta}\right)\right\}$, where p is the likelihood and $\hat{\theta}_i$ is the estimate at time i. While appealing, these confidence sets are difficult to approximate or cover in terms of a tractable convex set, e.g., an ellipsoid. For example, a common approach to approximate these confidence sets is to use Taylor's theorem and expand around the maximum likelihood estimate up to the second order. This is known as Laplace approximation, in the context of Bayesian estimators. In particular, short calculation reveals that the approximate covariance is: $\mathbf{V}_t = \sum_{i=1}^t n_i \varphi_{A_i}\varphi_{A_i}^\top/(\varphi_{A_i}^\top\theta)^2$. This approximation, however suffers from a pathology: in region A with very small rate $\varphi_A^\top\theta$, the counts are mostly zero, but observing $n_i = 0$ does not increase $\mathbf{V}_t$ and hence does not shrink the confidence set. Therefore, on problems with small counts (small l), it is reasonable to expect that this approach fails as the algorithm believes the uncertainty is high despite sensing multiple times – a scenario which we observe in numerical experiments in Section 7. If we were to lower bound the likelihood ratio, and hence maintain coverage, the dependence on n_i stays. Nevertheless, as long as $l > 0$, an approach based on such confidence sets would converge, albeit slowly, as it does not extract information from not observing events $(n_i = 0)$.

Penalized likelihood does not in fact differ much from least-squares regression. Its optimality conditions have exactly the same form as Equation (6), with the notable difference that instead of σ_i^2 on diagonal of $\mathbf{\Sigma}$, they contain empirical variances $\varphi_{A_i}^\top\hat{\theta}$. In light of this, it might make sense to use the confidence sets derived for the regression estimator with the likelihood estimator to achieve best of both worlds in practice: an approach we explore numerically.

6. Regret Analysis

Upper Confidence Bound In order to run CAPTURE-UCB, we need to calculate upper confidence bounds for each sensing region; for the estimators presented in Section 4, we can do this by solving a convex program: [3]:

$$\begin{aligned} \text{ucb}_{t-1}(B),\ \text{lcb}_{t-1}(B) &= \max_\theta \varphi_B^\top,\ \min_\theta \varphi_B^\top\theta \\ \text{subject to } & (\theta - \hat{\theta}_{t-1})^\top \mathbf{V}_{t-1}(\theta - \hat{\theta}_{t-1}) \leq \beta_t(\delta) \\ & l \leq \mathbf{\Gamma}^{1/2}\theta \leq u \ \ (\text{or } M(\theta) \succeq 0). \end{aligned} \tag{10}$$

where $\mathbf{V}_{t-1} = \sum_{i=1}^{t-1} \frac{1}{\sigma_t{}^2(A_i)}\varphi_{A_i}\varphi_{A_i}^\top + \gamma\mathbf{I}$ and $\beta(\delta)$ is the confidence parameter (9).

6.1. Information Gain

Our regret bound that we establish in Section 6.2 will depend on the representational capacity of $\mathcal{H}_k$, which can be loosely interpreted as its dimensionality of the Hilbert space. We will state our results in terms of *information gain* (Srinivas et al., 2010), a common notion in the kernelized bandits literature, which also appears in our confidence parameter $\beta_t(\delta)$ as $\log(\det(\mathbf{V}_t)/\det(\mathbf{I}\gamma))$. It is more commonly defined in its dual formulation due to Weinstein–Aronszajn identity,

$$\gamma_T(\mathcal{A}) = \log\det(\mathbf{K}_T + \mathbf{\Sigma}_T) - \log\det(\mathbf{\Sigma}_T)^{4}$$

where $(\mathbf{K}_T)_{i,j} = \int_{A_j}\int_{A_i} k(x,z)dxdz = \kappa(A_i, A_j)$, and diagonal $(\mathbf{\Sigma}_T)_{ii} = \gamma\sigma_i^2$. Notably, in our setting, the kernel function measures similarity between Borel sets. This is in contrast to point-wise similarity as common for kernel methods. The argument of the information gain indicates that the form of γ_T differs for count-record (with $\mathcal{B}$) and histogram (with $\mathcal{A}$) feedback due to different inputs to the kernel matrix k.

To give some intuition behind this quantity, for any finite rank RKHS, as in Assumption 1, we can immediately give a bound

$$\gamma_T \leq m\log\left(1 + \frac{Tu}{\gamma\rho^2}\right) \tag{11}$$

with $\rho^2 = \min_t \sigma_t^2$. This means that it has logarithmic dependence on T and linear dependence on m. While this bound is satisfactory for finite rank RKHS – or linear bandits – with fixed features $\Phi(x)$ or $\Psi(x)$, it might be less so for kernelized bandits where m can grow arbitrarily large.

[3] The operator M takes vector into a matrix form to be consistent with trace regression formulation.

[4] The γ_T with the time subscript should not be confused with γ without a subscript, which denotes regularization parameter.

In light of this, we show that by using either the *triangle basis* or *QFF* of Mutný and Krause (2018) for stationary kernels, γ_T, as function of m has a finite limit related to the original kernel k. The formal statement of this result along with its proof can be found Appendix in C.2 and C.4.

Secondly, previous results bounding γ_T assume point evaluations, and do not extend trivially to integral actions. We show that these can be related to the standard bounds on γ_T at cost of changing the regularization parameter (which can also lead to improvements). Importantly, this change does not impact the asymptotic scaling in T for stationary kernels, e.g., for the squared exponential kernel used in our experiments, $\gamma_T \leq \mathcal{O}((\log T)^d)$, (more details in C.4 and C.5).

6.2. Regret Bound

Classical bandit proof techniques yields the following result as a consequence of Theorem 1.

Theorem 2. *For any $\delta \in (0, 1)$, under Assumptions 1, 2, and cost function w such that the best sensing action is time invariant,* CAPTURE-UCB *suffers for $T > 0$:*

$$R_T \leq \begin{cases} \mathcal{O}(\sqrt{\rho_H T \Delta \beta_T(\delta, \mathcal{A}) \gamma_T(\mathcal{A})}) & \textit{if histogram} \\ \mathcal{O}(\sqrt{\rho_C T \Delta \beta_T(\delta, \mathcal{B}) \gamma_T(\mathcal{B})}) & \textit{if count-record} \end{cases}$$

with probability $1 - \delta$. Hereby, $\beta(\delta)$ is a function of γ_T and is as in Theorem 1, and $\rho_H = \max_{t \in T} \sigma_t^2(A_t)$, $\rho_C = \max_{t \in T} \max_{B \in A_t} \sigma_t^2(B)$.

The values of ρ_H and ρ_C are finite and depend only on U and u. As $l \to 0$ (lower bond on the intensity) ρ is well-behaved an decreases towards zero, where the decrease depend only on U and u as one can see in Lemma 4 in Appendix.

Inspecting the direct dependence of β_T with $\gamma = m$, we conclude that the overall regret is $\mathcal{O}(\sqrt{T}(\gamma_T/\sqrt{m} + \sqrt{m})\sqrt{\gamma_T})$ involving the basis size m. Using (11), we can refine this bound and express it only in terms of m and T as $\mathcal{O}(m\sqrt{T}\log T)$, i.e., the same as for sub-Gaussian random noise up to a logarithmic factor of T. The effect of Poisson noise is captured in the value of ρ_C and ρ_H, which are the scaled variances and they depend only on the values of U, u and Δ in the worst case.

Extension to general RKHS While the rate of regret R_t is optimal (up to logarithmic factors) for any finite rank bandit algorithm, m does not always represent the true intrinsic statistical difficulty of the problem. The appropriate complexity measure instead is γ_T. It remains an important open problem to express the results of Theorem 1 solely in terms of γ_T. Instead, we discuss an approach based on finite approximations. Namely, for any time t, suppose we want to optimize the regret up to precision $\epsilon = 1/t^2$. This is sufficient, since we cannot hope to have better estimation accuracy than $t^{-1/2}$. Using the Mercer decomposition of the kernel, we use for computation $\kappa = \Phi(x)^\top \Phi(x)$, where, for $\Phi(x) \in \mathbb{R}^m$, we construct a truncated version, where only eigenvalues above ϵ are considered. To construct the Mercer decomposition, we require a measure, which in this case is the uniform measure over the action set. Let us denote the number of eigenvalues above ϵ by m_ϵ. This number is upper-bounded by twice the effective dimension (Bach, 2017), which is itself upper-bounded by γ_t. We can then apply the confidence results with m_{1/t^2} playing effectively role of γ_t, since it can be bounded by it (with regularization constant ϵ). Under this supposition, the overall regret for $t \leq T$ is bounded by $\mathcal{O}(\sqrt{T}\gamma_T)$, again recovering the best known complexity for kernelized bandits with sub-Gaussian noise assumption. Notice that we had to drop the anytime assumption, since the regularization parameter needs to scale with T.

The previous theorem is proved under the assumption that the true function can be represented via finite rank approximation. We did not empirically study the effect of the approximation error, but this error can be reduced arbitrarily by increasing m, and as we already established that γ_T of the approximated RKHS converges to the true γ_T from below, the bound remains valid.

6.3. Refinement problem: Count-record

The regret of the count-record feedback depends on the domain partition $\mathcal{B}$ we choose. If the partition is very small, then so are the counts falling into the individual sets B. We know that the concentration properties of the Poisson distribution lead to larger corrections for moderate variances, but overall they decrease to zero, albeit slowly. In other words, $\rho_C = \max_{t \in T} \max_{B \in A_t} \sigma_{t-1}^2(B)$ decreases with decreasing size of partitioning even if the correction k^* grows. The overall variance $\sigma^2(B) = k^* \mathbb{E}[N(B)]$ converges to 0 as the variance of $N(B)$ goes to zero (Lemma 4), which is necessarily an effect of partitioning more finely. On the other hand, $\gamma_T(\mathcal{B}) \geq \gamma_T(\mathcal{A})$ for the same set of chosen actions, since the elements of $(\mathbf{K}_T)_{ij}$ are $\kappa(B_i, B_j)$ for $B_i, B_j \in \mathcal{B}$ instead of the actions A_i, A_j as for histogram feedback, and these can be shown to be smaller (formal statement in Appendix C.9). Hence, one should choose a level of partition such that the two aspects are balanced – local information and adaptivity. We call this the *refinement problem*. Unfortunately, as our assumptions are very general, optimal partitioning depends non-trivially on the kernel k, the action set $\mathcal{A}$ and the past data. While it is theoretically conceivable to search over all $\mathcal{B}$ for each $A \in \mathcal{A}$ to find the smallest ucb, this procedure is practically infeasible. Perhaps a practical balance is to calculate a sequence of bounds for each A based on sequential splitting of the domain (quadtree in two dimensional case), and stop splitting once ucb does not decrease, or to fix a sufficiently fine binning, an approach we use in our experiments and in theory. The exact analysis of this *refinement problem* on the regret is an interesting direction for future research.

7. Experiments and Applications

In our experiments, we investigate both feedback models as well as different cost models: *uniform cost* $w(A) = |A|$, *concave cost* $w(A) = |A| + 0.01$, and *one-off cost*. We compare: ϵ-*greedy*, CAPTURE-UCB, and the Thompson sampling algorithm of Grant et al. (2019), which does not exploit correlation between sensing regions. We also report a variant of our approach called Information Directed Sampling CAPTURE-IDS, which we explain in Appendix A.2. It has a similar flavor as UCB but uses a different utility function. Additionally, we compare with algorithms based on penalized likelihood: UCB-$\mathcal{L}$ and IDS-$\mathcal{L}$ using the same confidence set as our heteroscedasdic estimator. We also report UCB-LAP – a UCB algorithm with quadratic approximation of the likelihood ratio confidence set. This approximation is equivalent to Laplace approximation in the Bayesian formulation with a Gaussian prior. The algorithm ϵ-*greedy* simply plays the current best estimate with $1 - \epsilon$ probability, and with $\epsilon = \mathcal{O}(t^{-1/2})$ probability a random sensing region is selected. To predict the best region, ϵ-greedy uses the RKHS model.

Confidence parameter We relax the theoretical requirements of the confidence sets, and simply use $\beta = 4$, as well as *tweak the past*, by which we mean we tweak the past variances with the new estimates of λ, causing σ_t^2 to be no longer adapted. This is not warranted by theory but nearly always improves the performance in practice. We study these two aberrations from the theory on a *toy-problem*: $\lambda(x) = 4\exp(-(x+1))\sin(2x\pi)^2$ with $\delta = 0.5$. In Fig. 2a), we see that both changes improve the finite time performance with the most impact due to different choice of β. While *tweaking* improves, the overall behavior is similar. This result is consistent with the observations from the bandit literature, where β is viewed as a tuning parameter. The remaining experiments are with *tweaking*, $\beta = 4$ and $\gamma = 1$. Additionally, the Poisson counts are superimposable – this means that sensing the same regions multiple times can be merged together as they contain the same amount of information. While the likelihood estimator automatically implements this feature, we need to enforce it with regression perspective, which can be done in connection with *tweaking*.

Summary There are three main overall messages from the benchmarks: 1) All algorithms utilizing the RKHS structure substantially outperform algorithms that do not use this structure, such as Grant et al. (2019). 2) The confidence sets stemming from Laplace approximation fail, as evident in, e.g., Fig. 2. 3) Both feedback models, as well as all cost models do not have qualitative performance effect on the algorithm and are handled equally well. Now we discuss the modeling and setup for each experiment separately.

7.1. Benchmarks

Taxis Suppose we have to dispatch 5 taxi cars each to a distinct block of a city, and the goal is maximize the number of serviced passengers. We use histogram feedback and hierarchically split the domain as in Fig. 1a. We use the *triangle basis* with squared exponential kernel, uniform costs, $\Delta = 60$ min, and report the results in Fig. 2b).

Crime surveillance We use a crime dataset released by the San Francisco police department to estimate the intensity of burglary occurrences over space (see Fig. in Appendix D). The sensing duration is set to $\Delta = 30$ days corresponding roughly to placing a mobile camera to that region. We report the results with count-record feedback, triangle basis and squared exponential kernel in Fig. 2c).

Environmental Monitoring We use two datasets from Baddeley et al. (2015), one containing locations of an African tree shrub *Beilschmiedia*, and one containing *Gorilla* nesting locations in a Cameroon rain forest. These examples can be motivated by species monitoring from satellites or an airplane. Each sensing action corresponds to zooming onto a specific region in order to identify the objects of interest. In the first case, we implement the one-off constraint and use the trace regression estimator, where the features Ψ are modeled as Fourier features of the squared exponential kernel. The squared exponential kernel takes as input the value of slope $s_{(x,y)}$ and height $h_{(x,y)}$ at location (x, y), since we believe these to be predictive of habitat. We report captured events with count-record feedback in Fig. 2d). In Fig. 2e), we report the regret on the *Gorillas* dataset with count-record feedback, concave costs, squared exponential kernel and triangle basis. More details along with the fitted models are presented in Appendix D.

8. Related Work

The intensity of Poisson point processes is typically estimated either via smoothing kernels (Berman and Diggle, 1989) or via kernelized estimators (Lloyd et al., 2015). Kernelized estimators need to enforce the positivity constraint either via a link function (Adams et al., 2009; John and Hensman, 2018), by enforcing it at sufficiently many points in the domain (Agrell, 2019; Aubin-Frankowski and Szabó, 2020) or using specially constructed positive bases such as positive splines, Bernstein polynomials (Papp and Alizadeh, 2014) or the simpler non-polynomial *triangle basis* (Maatouk and Bay, 2016; López-Lopera et al., 2018). Positive bases have been used for Cox process inference by López-Lopera et al. (2019) and Alizadeh et al. (2008). We use estimators based on positive bases or using constrained formulations as in (Marteau-Ferey et al., 2020) or (Aubin-Frankowski and Szabó, 2020).

Confidence sets Confidence sets for adaptively collected data with sub-Gaussian noise dependence are analyzed via self-normalized processes (de la Peña et al., 2009), as

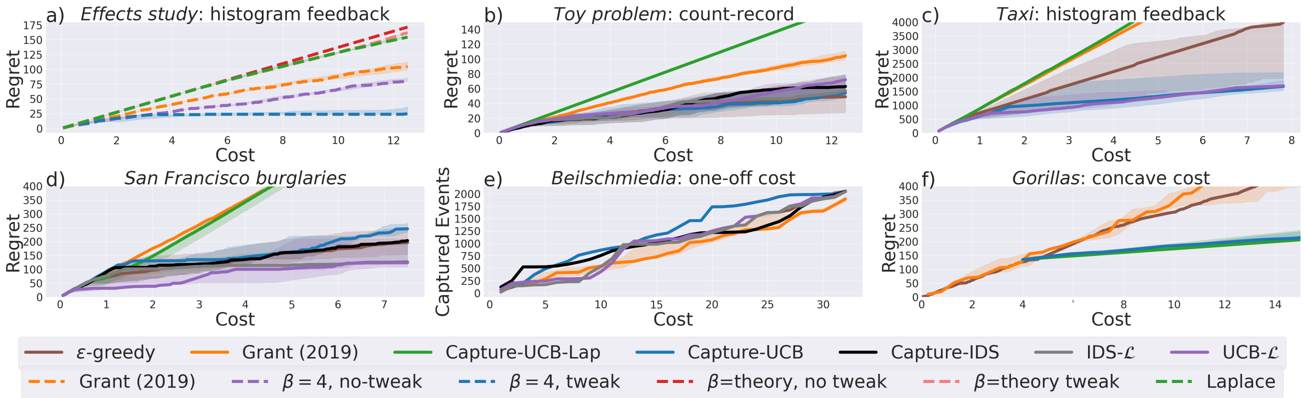


Figure 2. Numerical experiments on our benchmarks: Each experiment was repeated $5 - 10$ times, when randomization is in either the algorithm or output. Standard quantiles are depicted as shaded regions. Approaches utilizing the RKHS structure are superior for all datasets. The algorithms with Laplace approximation fail on problems with regions with low intensity. The effect of peeking and width of confidence sets is investigated in a), where importance of smaller β is more prominent than the effect of *tweaking*. In f) notice that due to concave costs, some algorithms start with high cost since they sensed larger regions – this is expected and desired.

in the seminal work of Abbasi-Yadkori et al. (2011) for linear bandits. This theory can be extended to other noise generating processes by identifying the right non-negative super-martingale (Howard et al., 2018). Faury et al. (2020) present adaptive confidence bounds using the self-normalized technique for bounded noise. We also exploit self-normalized martingales with similar techniques.

Adaptive Sensing Grant et al. (2019; 2020) study adaptive sensing of Poisson point processes with sensing costs. They bin the domain, and sequentially refine estimators of bins with truncated gamma priors to design a regret minimizing algorithm. Importantly, the regret authors study is different to *count-regret* studied here – the cost is additive in their work while here the cost enters as a ratio. Crucially, their bins are considered to be independent, while we exploit correlation via the RKHS assumption. The binning is done in order to address the refinement problem as their collection action sets $\mathcal{A}$ is not finite. Grant and Szechtman (2020) study a related problem constrained to sense a region of increasing volumes.

After binning, regression based estimators tailored for heavy tailed noises, e.g., based on median on means are applicable. However, these methods do not adapt to a particular noise model nor they adapt to heteroscedascity (Medina and Yang, 2016; Shao et al., 2018) of the Poisson distribution, and have worst-case flavor. In the bandit literature, there are methods that adapt to *a priori known* heteroscedascity with sub-Gaussian noice (Kirschner and Krause, 2018), or to semi-bandit feedback (György et al., 2007; Kirschner et al., 2020).

Generalized Linear Models For count data, the textbook approach (Cameron and Trivedi, 2013) suggests log-linear models: $\mathbb{E}[N(A)] = \exp(\eta^\top \psi(A))$ – an instance of generalized linear models (GLMs) McCullagh (2018). However, this approach cannot be easily reconciled with the linear Poisson process structure and the RKHS assumption. We explain this in more detail in Appendix A.1. Without the RKHS assumption and instead using the log-linear model with specially constructed $\psi(A)$, Filippi et al. (2010); Jun et al. (2017) propose bandit algorithms for GLMs. However they do not cover Poisson noise, since they assume bounded and sub-Gaussian noise, respectively.

9. Future Work and Conclusion

Future Works There are several issues that we did not address in this paper that warrant further study and open new avenues of research. Among many others, we did not allow for Δ to be time varying nor did we analyze the approximation properties of finite bases. Perhaps the biggest open question is the derivation of *anytime* confidence sets independent of finite rank m only scaling with respect to γ_T.

Conclusion We introduced the problem of sensing inhomogeneous Poisson point processes. We derived adaptive confidence sets for heteroscedasdic linear regression for Poisson count data with properly scaled variances. Applying these, we proposed algorithms that minimize the *count regret* based on optimism and information directed sampling. We proved that our optimistic algorithm is no-regret, and demonstrate the applicability of the approach on several real-world problems. We discussed several practical implementation improvements, and compared to algorithms based on Laplace approximation confidence sets.

Acknowledgement

This research was supported by the SNSF grant 407540_167212 through the NRP 75 Big Data program and by the Swiss National Science Foundation through the NCCR Catalysis.

References

Y. Abbasi-Yadkori, D. Pál, and C. Szepesvári. Improved algorithms for linear stochastic bandits. In *Advances in Neural Information Processing Systems*, pages 2312–2320, 2011.

R. P. Adams, I. Murray, and D. J. MacKay. Tractable nonparametric bayesian inference in poisson processes with gaussian process intensities. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 9–16, 2009.

C. Agrell. Gaussian processes with linear operator inequality constraints. *arXiv preprint arXiv:1901.03134*, 2019.

F. Alizadeh, J. Eckstein, N. Noyan, and G. Rudolf. Arrival rate approximation by nonnegative cubic splines. *Operations Research*, 56(1):140–156, 2008.

P.-C. Aubin-Frankowski and Z. Szabó. Hard shape-constrained kernel machines. *arXiv preprint arXiv:2005.12636*, 2020.

P. Auer. Using confidence bounds for exploitation-exploration trade-offs. *Journal of Machine Learning Research*, 3(Nov):397–422, 2002.

F. Bach. On the equivalence between kernel quadrature rules and random feature expansions. *Journal of Machine Learning Research*, 18(21):1–38, 2017.

A. Baddeley, R. E., and T. R. *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC Press, 2015.

M. Berman and P. Diggle. Estimating weighted integrals of the second-order intensity of a spatial point process. *Journal of the Royal Statistical Society: Series B (Methodological)*, 51(1):81–92, 1989.

A. C. Cameron and P. K. Trivedi. *Regression analysis of count data*, volume 53. Cambridge university press, 2013.

N. Cressie and G. Johannesson. Fixed rank kriging for very large spatial data sets. *J. R. Statist. Soc. B*, 70(1):209–226, 2008.

V. H. de la Peña, M. J. Klass, and T. L. Lai. Theory and applications of multivariate self-normalized processes. *Stochastic Processes and their Applications*, 119(12):4210 – 4227, 2009. ISSN 0304-4149. doi: https://doi.org/10.1016/j.spa.2009.10.003. URL http://www.sciencedirect.com/science/article/pii/S0304414909001689.

P. Diggle. *Statistical analysis of spatial point patterns*. 2013.

P. J. Diggle, P. Moraga, B. Rowlingson, B. M. Taylor, et al. Spatial and spatio-temporal log-gaussian cox processes: extending the geostatistical paradigm. *Statistical Science*, 28(4):542–563, 2013.

L. Faury, M. Abeille, C. Calauzenes, and O. Fercoq. Improved optimistic algorithms for logistic bandits. In *ICML'20: Proceedings of the 37th International Conference on International Conference on Machine Learning*, 2020.

S. Filippi, O. Cappe, A. Garivier, and C. Szepesvári. Parametric bandits: The generalized linear case. In *Advances in neural information processing systems*, pages 586–594, 2010.

J. A. Grant and R. Szechtman. Filtered poisson process bandit on a continuum. *arXiv preprint arXiv:2007.09966*, 2020.

J. A. Grant, A. Boukouvalas, R.-R. Griffiths, D. S. Leslie, S. Vakili, and E. M. De Cote. Adaptive sensor placement for continuous spaces. *arXiv preprint arXiv:1905.06821*, 2019.

J. A. Grant, D. S. Leslie, K. Glazebrook, R. Szechtman, and A. N. Letchford. Adaptive policies for perimeter surveillance problems. *European Journal of Operational Research*, 283(1):265–278, 2020.

A. György, T. Linder, G. Lugosi, and G. Ottucsák. The online shortest path problem under partial monitoring. *Journal of Machine Learning Research*, 8(Oct):2369–2403, 2007.

E. Hazan, A. A., and S. Kale. Logarithmic regret algorithms for online convex optimization. (69(2-3)), 2007.

J. Heikkinen and E. Arjas. Modeling a poisson forest in variable elevations: a nonparametric bayesian approach. *Biometrics*, 55(3):738–745, 1999.

S. R. Howard, A. Ramdas, J. McAuliffe, and J. Sekhon. Time-uniform chernoff bounds via nonnegative supermartingales. 2018.

S. John and J. Hensman. Large-scale Cox process inference using variational Fourier features. In J. Dy and A. Krause, editors, *ICML 2018*, volume 80 of *Proceedings of Machine Learning Research*, pages 2362–2370, Stockholmsmässan, Stockholm Sweden, 10–15 Jul 2018. PMLR. URL http://proceedings.mlr.press/v80/john18a.html.

K.-S. Jun, A. Bhargava, R. Nowak, and R. Willett. Scalable generalized linear bandits: Online computation and hashing. *arXiv preprint arXiv:1706.00136*, 2017.

M. Kanagawa, P. Hennig, D. Sejdinovic, and B. K. Sriperumbudur. Gaussian processes and kernel methods: A review on connections and equivalences. *arXiv preprint arXiv:1807.02582*, 2018.

Kingman. Poisson processes. 2002.

J. Kirschner and A. Krause. Information directed sampling and bandits with heteroscedastic noise. *arXiv preprint arXiv:1801.09667*, 2018.

J. Kirschner, T. Lattimore, and A. Krause. Information directed sampling for linear partial monitoring. *Conference on Learning Theory (COLT 2020)*, 2020.

C. Lloyd, T. Gunter, M. A. Osborne, and S. J. Roberts. Variational inference for gaussian process modulated poisson processes. In *Proceedings of the 32Nd International Conference on International Conference on Machine Learning - Volume 37*, ICML'15, pages 1814–1822. JMLR.org, 2015. URL http://dl.acm.org/citation.cfm?id=3045118.3045311.

A. F. López-Lopera, F. Bachoc, N. Durrande, and O. Roustant. Finite-dimensional gaussian approximation with linear inequality constraints. *SIAM/ASA Journal on Uncertainty Quantification*, 6(3):1224–1255, 2018. doi: 10.1137/17M1153157. URL https://doi.org/10.1137/17M1153157.

A. F. López-Lopera, S. John, and N. Durrande. Gaussian process modulated cox processes under linear inequality constraints. *arXiv preprint arXiv:1902.10974*, 2019.

H. Maatouk and X. Bay. Gaussian process emulators for computer experiments with inequality constraints. 2016.

U. Marteau-Ferey, F. Bach, and A. Rudi. Non-parametric models for non-negative functions. 2020.

P. McCullagh. *Generalized linear models*. Chapman and Hall, 2018.

A. M. Medina and S. Yang. No-regret algorithms for heavy-tailed linear bandits. In *International Conference on Machine Learning*, pages 1642–1650, 2016.

M. Mutný and A. Krause. Efficient high dimensional bayesian optimization with additivity and quadrature fourier features. In *Neural and Information Processing Systems (NeurIPS)*, December 2018.

M. Mutný, K. Johannes, and A. Krause. Experimental design for orthogonal projection pursuit regression. *AAAI2020*, 2020.

D. Papp and F. Alizadeh. Shape-constrained estimation using nonnegative splines. *Journal of Computational and Graphical Statistics*, 23(1):211–231, 2014. doi: 10.1080/10618600.2012.707343. URL https://doi.org/10.1080/10618600.2012.707343.

H. Robbins, D. Siegmund, et al. A class of stopping rules for testing parametric hypotheses. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 4: Biology and Health*. The Regents of the University of California, 1972.

D. Russo and B. Van Roy. Learning to optimize via posterior sampling. *Mathematics of Operations Research*, 39(4):1221–1243, 2014.

B. Schölkopf, R. Herbrich, and A. Smola. A generalized representer theorem. In *Computational learning theory*, pages 416–426. Springer, 2001.

H. Shao, X. Yu, I. King, and M. R. Lyu. Almost optimal algorithms for linear stochastic bandits with heavy-tailed payoffs. In *Advances in Neural Information Processing Systems*, pages 8420–8429, 2018.

S. Shirota, A. E. Gelfand, et al. Space and circular time log gaussian cox processes with application to crime event data. *The Annals of Applied Statistics*, 11(2):481–503, 2017.

D. L. Snyder and M. I. Miller. *Random point processes in time and space*. Springer Science & Business Media, 2012.

N. Srinivas, A. Krause, S. M. Kakade, and M. Seeger. Gaussian process optimization in the bandit setting: No regret and experimental design. *International Conference on Machine Learning*, 2010.

C. Szepesvari and T. Lattimore. *Bandit Algorithms*. 2019.

S. Vakili, K. Khezeli, and V. Picheny. On information gain and regret bounds in gaussian process bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 82–90. PMLR, 2021.

H. P. Vanchinathan, A. Marfurt, C.-A. Robelin, D. Kossmann, and A. Krause. Discovering valuable items from massive data. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1195–1204, 2015.