Online Linear Regression with Paid Stochastic Features

Nadav Merlis[1], **Kyoungseok Jang**[2], **Nicolò Cesa-Bianchi**[3]

[1]Technion, [2]Chung-Ang University, [3]University of Milan & Politecnico di Milano
nmerlis@technion.ac.il, ksjang@cau.ac.kr, nicolo.cesa-bianchi@unimi.it

Abstract

We study an online linear regression setting in which the observed feature vectors are corrupted by noise and the learner can pay to reduce the noise level. In practice, this may happen for several reasons: for example, because features can be measured more accurately using more expensive equipment, or because data providers can be incentivized to release less private features. Assuming feature vectors are drawn i.i.d. from a fixed but unknown distribution, we measure the learner's regret against the linear predictor minimizing a notion of loss that combines the prediction error and payment. When the mapping between payments and noise covariance is known, we prove that the rate $\sqrt{T}$ is optimal for regret if logarithmic factors are ignored. When the noise covariance is unknown, we show that the optimal regret rate becomes of order $T^{2/3}$ (ignoring log factors). Our analysis leverages matrix martingale concentration, showing that the empirical loss uniformly converges to the expected one for all payments and linear predictors.

1 Introduction

In online linear regression, a learner sequentially observes i.i.d. feature vectors $x_t \in \mathbb{R}^d$ and aims to correctly predict outputs $y_t = x_t^\top \theta^*$, knowing neither θ^* nor the distribution from which the feature vectors are drawn. When measured by the quadratic prediction error, this is potentially the most canonical online regression setting, and several extensions to more complex losses and regression models have been studied in the past. Although in the vast majority of works the features x_t are assumed to be perfectly observable, in practice, we usually only have access to a noisy version of x_t. For example, features obtained from experiments often have measurement errors, and contextual information on users only relies on a random sample of their behavior. Sometimes the features are even intentionally corrupted (e.g., via locally differentially private mechanisms) to protect user privacy.

Motivated by scenarios where the learner can reduce the noise intensity by allocating more resources or more work, we study a variant of online linear regression where the features are stochastic and corrupted by additive noise, and the learner may decide to pay to reduce the noise level so that a higher payment corresponds to a smaller noise covariance (w.r.t. the positive definite order). When conducting experiments, this payment can be related to renting better lab equipment or hiring more experienced staff, but also to extending the measurement duration and/or averaging more raw samples. With user data, this noise reduction can be achieved by collecting more statistics before making a decision. In privacy-protected settings, the learner may offer payments to compensate users for limited privacy loss (either directly or by offering free services)—see, e.g., (Wang, Ying, and Zhang 2018). While reducing the feature noise naturally facilitates prediction, it also causes some penalty to the learner (either time, money, or given services). Thus, the learner should optimally balance the prediction error and the cost due to payments.

We measure a predictor's performance by a loss combining the prediction error with the payment. The goal of the learner is to minimize regret, defined as the difference between the learner's expected cumulative loss and the cumulative loss of the best linear predictor, i.e., the one that optimally balances prediction errors and payments.

We first tackle the setting where the learner knows how the noise covariance changes as a function of the payment. This can be the case, for instance, when the noise is generated by a known differentially private mechanism. We show how to use this knowledge to effectively utilize samples collected at one payment level to estimate the optimal prediction at any other payment level. This information sharing leads to a major acceleration in the learning process and obviates the need for payment exploration. In particular, we devise an empirical loss estimator that uniformly converges to the expected one across all potential payment levels and predictions (including payments that were never offered by the algorithm). We utilize this result to prove that an algorithm that greedily minimizes this empirical loss to determine the payment level and linear predictor achieves a regret bound of $\widetilde{\mathcal{O}}(\sqrt{T})$ after T prediction rounds. We complement this result with a lower bound, proving that the dependence on T in this bound can only be improved by logarithmic factors. Note that the optimal rate in this setting is exponentially worse than the $\Theta(\ln T)$ bound for the noise-free case.

We then move on to the case where the noise covariances are not known in advance. This could happen, for example, when noise reduction is the result of an extended data collection period, especially if the noise is temporally correlated. In this scenario, we can no longer use our uniform estimation

scheme and must incorporate a payment exploration mechanism into our algorithm. Inspired by UCB-like algorithms (Auer 2002), we limit the potential payments to a discretized grid and introduce an optimistic loss estimator, which encourages the algorithm to choose less-played payment levels. We show that by carefully tuning the grid size, this approach yields a regret bound of $\widetilde{\mathcal{O}}\big(T^{2/3}\big)$. We prove that, similarly to the case of known covariances, the dependence on T in this bound can only be improved by logarithmic factors. Our lower bound construction carefully designs a set of noise covariance profiles to induce losses that imitate hard instances in Lipschitz bandits (Kleinberg 2004).

Outline. The rest of the paper is organized as follows. We first position our work in light of the existing literature in Section 2. We then formalize our setting and describe our assumptions in Section 3. In Section 4, we present our loss estimator when the noise covariance profile is known in advance; we show how to use this estimator to derive an order-optimal algorithm. Then, Section 5 extends these results to situations when the noise covariance profile is not given to the learner. We conclude the paper with a summary and a discussion on the many potential future directions in Section 6.

2 Related Work

Errors-in-variables models. Linear regression with noisy covariates (also known as errors-in-variables, or EIV) is a classical topic in statistics—see, e.g., (Fuller 2009) and (Agarwal and Singh 2021) for recent applications to differential privacy. Most results in EIV concern the problem of estimating θ^* under different norms, typically in a high-dimensional setting (large d) under assumptions of sparsity (θ^* has few non-zero components) or low-rank (the covariates admit a low-dimensional representation), see (Rosenbaum and Tsybakov 2010; Loh and Wainwright 2011; Chen and Caramanis 2013; Rosenbaum and Tsybakov 2013; Kaul and Koul 2015; Belloni, Rosenbaum, and Tsybakov 2017; Datta and Zou 2017) and references therein. A related line of work studies EIV under Bayesian assumptions (Reilly and Patino-Leal 1981; Ungarala and Bakshi 2000; Figueroa-Zúñiga et al. 2022). Agarwal et al. (2023) consider EIV in an estimation error task where covariates are chosen adaptively (as in active learning or in multi-armed bandits). Closer to our work is the paper by Agarwal et al. (2019), where they study a linear regression model with fixed design and missing or corrupted variables. Unlike us, they bound the prediction error in a transductive setting, where the covariates on which predictions are evaluated are available at training time.

Local DP. A prominent situation where features are noisy is when they are protected by a local differentially private mechanism (Kasiviswanathan et al. 2011). Two well-known approaches to create this protection are to add the raw features either Laplace or Gaussian noise (Yang et al. 2024). Many works study how to learn when features are private, including in the context of linear regression (Duchi, Jordan, and Wainwright 2013; Smith, Thakurta, and Upadhyay 2017; Wang and Xu 2019; Fukuchi, Tran, and Sakuma 2017). However, their goal is mostly to estimate the true parameters of the model given the noisy features; in contrast, we want to learn how to use online noisy data to output good predictions. As we will show, the best predictors are not the true parameters, but rather depend on the noise level.

Online learning with noisy features. Closely related to ours is the work by Cesa-Bianchi, Shalev-Shwartz, and Shamir (2011), who studied online linear regression with square loss with noisy features and known/unknown covariances. However, there are several crucial differences: they do not have a notion of cost, and their notion of regret compares to the best linear predictor on the noise-free features. Moreover, their focus is on a setting in which the learner may obtain more than one noisy realization of the same feature vector. A second closely related work is the one by Van Der Hoeven et al. (2023). They prove a $d^2(\ln T)\sqrt{T}$ regret bound in a binary classification setting in which: the learner can pay to reduce the noise level, the covariates (called experts in their setting) are binary, and noise and payments apply independently on each coordinate. However, because of their focus on classification with independent feature noise, their techniques are not directly applicable to our regression setting with correlated noise. Other works assume noisy features, but in the partial feedback setting, include (Kim et al. 2023) and (Zheng et al. 2020).

Privacy and incentives. Another aspect of our work is the learner's capability to pay for noise reduction. In the privacy literature, this is often addressed as a problem of mechanism design: an analyst designs a procedure that incentivize individuals to sell their private data, aiming to obtain the most accurate estimation at a fixed payment budget (or paying a minimal cost for a target accuracy level) (Ghosh and Roth 2011; Nissim, Orlandi, and Smorodinsky 2012; Nissim, Vadhan, and Xiao 2014; Wang, Ying, and Zhang 2018; Fallah et al. 2024). Hsu et al. (2014) also studied the problem in which an individual is offered a payment to participate in a study and is willing to join it only if the payment covers its privacy loss. They show how an analyst running the experiment should tune the payment and privacy levels according to various considerations, including accuracy. All these works focus on payments that minimize the estimation error of the true parameters. Instead, our goal is to choose payments that balance the prediction error and the feature costs for every new incoming user. Also, our payment model is different: in our setting, the noise level is fixed (and potentially unknown) given the payment.

3 Setting

We study the problem of online linear regressions when the learner observes noisy features and can pay to reduce the noise level. At each round $t \in [T]$, the environment privately generates the true features $x_t \in \mathbb{R}^d$ independently from a fixed unknown distribution D_x and the output $y_t = x_t^\top \theta^*$ for some unknown $\theta^* \in \mathbb{R}^d$.[1] The learner decides to pay a cost ('payment') $c_t \in [0, 1]$ and in return observes noisy features $\hat{x}_t(c_t) = x_t + n_t(c_t)$, where $n_t(c) \sim D_n(c)$ is a zero-mean

[1] All results can be easily extended to noisy linear outputs, as we discuss in Remark 3 and Section B.2.

random variable independent of x_t. For brevity, we often omit the dependence on the cost c_t and denote the noisy features by $\hat{x}_t$. Given this observation, the learner tries to predict $\hat{y}_t(\hat{x}_t)$ so that it is as close as possible (in the squared distance) to y_t, based on the noisy features $\hat{x}_t$ and the data observed in the previous rounds. After the prediction, the learner observes the real value y_t. We focus on the class of linear predictors $\hat{y}_t(\hat{x}_t) = \hat{x}_t^\top \nu$. We denote $\mathbb{E}[x_t] = \bar{x}$, and also denote the covariance matrices of samples from D_x and $D_n(c)$ by $\Sigma_x = \mathbb{E}\left[x_t x_t^\top\right]$ and $\Sigma_n(c) = \mathbb{E}\left[n_t(c) n_t(c)^\top\right]$, respectively. Lastly, we denote the covariance of $\hat{x}_t$ given a cost c by $\Sigma_{\hat{x}}(c) = \mathbb{E}\left[\hat{x}_t(c)\hat{x}_t(c)^\top\right] = \Sigma_x + \Sigma_n(c)$.[2]

Additional Notations. We define the quadratic form of a vector $x \in \mathbb{R}^d$ w.r.t. a positive (semi-)definite matrix $A \in \mathbb{R}^{d\times d}$ as $\|x\|_A^2 = x^\top A x$. We also denote by $\|x\|$, the ℓ_2 norm of a vector $x \in \mathbb{R}^d$, and by $\|A\|_{op} = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}$, the operator norm of a matrix $A \in \mathbb{R}^{d\times d}$. We use $A \succ 0$ to denote positive definite matrices, and similarly write $A \succeq 0$ if A is positive semi-definite and $B \succeq A$ when $B - A \succeq 0$. We denote the observed history of the decision process up to the beginning of round $t+1$ (including the choice of payment but not the features) by $I_t^o = \{c_1, \hat{x}_1, \hat{y}_1, y_1, \ldots, c_t, \hat{x}_t, \hat{y}_t, y_t, c_{t+1}\}$ and the entire past data by $I_t = I_t^o \cup \{x_s, n_s\}_{s\in[t]}$.

Regularity assumptions. We assume that increasing the payment decreases the noise level, i.e., if $c_1 \leq c_2$, it holds that $\Sigma_n(c_2) \preceq \Sigma_n(c_1)$. Intuitively, it is equivalent to requiring that the variance of the noise measured in any arbitrary direction cannot rise when we increase the payment. We also assume that $\Sigma_{\hat{x}}(1) \succ 0$; by the covariance monotonicity assumption, this also implies that $\Sigma_{\hat{x}}(c) \succ 0$ for all $c \in [0,1]$. We further assume that both the features x_t and the noise $n_t(c)$ are conditionally R^2-subgaussian random vectors, that is, for any $u \in \mathbb{R}^d$, we have $\mathbb{E}\left[e^{u^\top (x_t - \bar{x})} | I_{t-1}\right] \leq e^{\frac{\|u\|^2 R^2}{2}}$ and $\mathbb{E}\left[e^{u^\top n_t} | I_{t-1}\right] \leq e^{\frac{\|u\|^2 R^2}{2}}$. Finally, we assume that both $\|\theta^*\|$ and $\|\bar{x}\|$ are upper bounded by a known $S > 0$.

Optimality Criterion. We compare ourselves to a linear regressor which pays a cost that optimally balances the prediction error and the payments for the features. In particular, for a given $\lambda > 0$, we define the loss for a fixed cost level c and linear predictor ν as

$$\ell(c,\nu) = \mathbb{E}_{x\sim D_x, n\sim D_n(c)}\left[\left((x+n)^\top \nu - x^\top \theta^*\right)^2\right] + \lambda c.$$

With a slight abuse of notation, for any given $S > 0$, we define the optimal linear predictor and loss given cost c as

$$\nu^*(c) = \arg\min_{\nu:\|\nu\|\leq S}\{\ell(c,\nu)\}, \qquad \ell^*(c) = \min_{\nu:\|\nu\|\leq S}\{\ell(c,\nu)\},$$

and denote the optimal cost and loss by

$$c^* \in \arg\min_{c\in[0,1]}\{\ell^*(c)\},$$
$$\ell^* = \min_{\nu:\|\nu\|\leq S, c\in[0,1]}\{\ell(c,\nu)\} = \min_{c\in[0,1]}\{\ell^*(c)\}.$$

[2] Since $x_t, \hat{x}_t$ are not zero mean, $\Sigma_x, \Sigma_{\hat{x}}(c)$ are sometimes called their autocorrelation.

We remark that $\nu^*(c), \ell^*(c)$ and ℓ^* all depend on S; we omit this dependence to simplify notations.

Remark 1. *Our choice to limit ourselves to linear predictions is due to both computational and statistical constraints. Specifically, a nonlinear optimal predictor requires computing $\mathbb{E}[x_t|\hat{x}_t]^\top \theta^*$. This estimator heavily depends on the entire feature and noise distributions, rendering the estimation problem extremely complicated and usually requires intractable computations. We adopt the solution of Van Der Hoeven et al. (2023) to this issue, and instead compare ourselves to the natural tractable benchmark. We also note that for some distributions, linear predictors are the best possible benchmark, as we later prove through our lower bounds.*

The loss and the optimal linear predictor can also be specified using the covariance matrices as follows:

Claim 1. *It holds that*

$$\ell(c,\nu) = \nu^\top \Sigma_{\hat{x}}(c)\nu - 2\nu^\top \Sigma_x \theta^* + \theta^{*\top} \Sigma_x \theta^* + \lambda c.$$

Moreover, let $\bar{\nu}(c) = (\Sigma_{\hat{x}}(c))^{-1}\Sigma_x\theta^*$. If $\|\bar{\nu}(c)\| \leq S$ then $\nu^*(c) = \bar{\nu}(c)$ and

$$\ell^*(c) = \theta^{*\top}\Sigma_x\theta^* - \theta^{*T}\Sigma_x\Sigma_{\hat{x}}(c)^{-1}\Sigma_x\theta^* + \lambda c.$$

Since we do not assume that the noise covariance is Lipschitz in the payment, the losses might not be continuous. Nonetheless, the monotonicity assumption induces a near-Lipschitz behavior:

Claim 2. *The loss $\ell^*(c)$ is λ-one-sided Lipschitz: for any $0 \leq c_1 \leq c_2 \leq 1$, it holds that $\ell^*(c_2) \leq \ell^*(c_1) + \lambda(c_2 - c_1)$.*

In words, the loss cannot significantly increase from a slight rise in the payment. Intuitively, this holds since increasing the cost always reduces the prediction error; therefore, the loss increment is always upper bounded by the increase in the payment term λc. This only linearly affects the loss, thus implying one-sided Lipschitzness. As a consequence, our algorithms could limit themselves to a finite grid of costs, knowing that even if the optimal cost is not on the grid, there is a slightly higher payment of similar loss (Tullii et al. 2024). The proofs for both claims can be found at Section A.

Remark 2. *Monotonicity of the noise covariances and Lipschitzness of the noise covariances are two incomparable assumptions – neither implies the other. Nonetheless, we rely on the monotonicity only to prove Claim 2, as well as arguing that $\Sigma_{\hat{x}}(c) \succ 0$ for all $c \in [0,1]$. Since Claim 2 also immediately holds if $\Sigma_n(c)$ is Lipschitz in c, under the assumption that $\Sigma_{\hat{x}}(c) \succ 0$, all our positive results would also hold for Lipschitz covariances.*

3.1 Regret Definition

As the objective, algorithms are similarly measured by a loss that combines their prediction error and feature payments. We focus on minimizing the expected cumulative loss, also equivalent to minimizing the regret:

$$\mathrm{Reg}(T) = \mathbb{E}\left[\sum_{t=1}^{T}\left(\left(\hat{y}_t(\hat{x}_t) - y_t\right)^2 + \lambda c_t\right)\right] - T\ell^*.$$

The regret can be decomposed as follows (see Section A for the proof).

Lemma 1. *For any arbitrary history-dependent prediction rule $\hat{y}_t : \mathbb{R}^d \mapsto \mathbb{R}$ and any arbitrary sequence $\nu_1, \ldots, \nu_T \in \mathbb{R}^d$, both potentially depend on I^o_{t-1}, it holds that*

$$\mathrm{Reg}(T) = \mathbb{E}\left[\sum_{t=1}^{T}\Big(\big(\hat{y}_t(\hat{x}_t) - y_t\big)^2 - \Big(\big(\hat{x}_t^\top \nu_t - y_t\big)^2\Big)\Big)\right] + \mathbb{E}\left[\sum_{t=1}^{T}(\ell(c_t, \nu_t) - \ell^*)\right].$$

In particular, if the algorithm only outputs linear predictions $\hat{y}_t(\hat{x}_t) = \hat{x}_t^\top \nu_t$, then

$$\mathrm{Reg}(T) = \mathbb{E}\left[\sum_{t=1}^{T}(\ell(c_t, \nu_t) - \ell^*)\right].$$

The first term encapsulates the regret of the prediction rule compared to some dynamic linear predictor, while the second term measures this linear predictor compared to the best linear predictor. In the rest of the paper, we restrict ourselves to linear predictions (as discussed in Remark 1), and so we focus on bounding the loss difference.

4 Online Linear Regression with Known Noise Covariances

In this section, we start from the favorable setting in which the learner has perfect knowledge of the noise covariances $\Sigma_n(c)$ for all $c \in [0, 1]$. In particular, to output a good prediction at the right cost, the learner only misses the feature covariance Σ_x and the linear map θ^*. Since neither of these quantities depends on the payment c_t, our aim is to *share information* between different cost observations, thus improving the regret. To this end, we take a closer look at the expected loss $\ell(c, \nu)$ in Claim 1, noticing that only a single term depends on the noise covariance: a term of the form $\nu^\top \Sigma_{\hat{x}}(c)\nu$. Given noisy independent feature samples at cost c, this term can be estimated empirically via $\hat{\Sigma}_{\hat{x}}(c) = \sum_{s=1}^{t} \hat{x}_s(c)\hat{x}_s(c)^\top$, whose expectation is $t\Sigma_{\hat{x}}(c)$. However, we only observe the noisy features for payments we played $\{c_s\}_{s \le t}$, so we cannot directly use this estimator for other unplayed payments. To circumvent this, we utilize our knowledge of the noise covariance and include a correction term that *shifts* the noise covariance from c_s to c. In particular, we suggest estimating the covariance with

$$\hat{\Sigma}_{\hat{x}}(c) = \sum_{s=1}^{t}\big(\hat{x}_s(c_s)\hat{x}_s(c_s)^\top + \Sigma_n(c) - \Sigma_n(c_s)\big).$$

This estimator uses *all past samples* while enjoying the desired expectation

$$\begin{aligned}\mathbb{E}\Big[\hat{\Sigma}_{\hat{x}}(c)\Big] &= \mathbb{E}\left[\sum_{s=1}^{t}\big(\hat{x}_s(c_s)\hat{x}_s(c_s)^\top + \Sigma_n(c) - \Sigma_n(c_s)\big)\right] \\ &= \mathbb{E}\left[\sum_{s=1}^{t}\big((\Sigma_x + \Sigma_n(c_s)) + \Sigma_n(c) - \Sigma_n(c_s)\big)\right] \\ &= \mathbb{E}\left[\sum_{s=1}^{t}(\Sigma_x + \Sigma_n(c))\right] = t\Sigma_{\hat{x}}(c).\end{aligned}$$

Inspired by this, we add a correction term of the form $\nu^\top(\Sigma_n(c) - \Sigma_n(c_s))\nu$ to the natural empirical loss and define the following loss estimator:

$$\hat{L}^{kc}_t(c, \nu) = \sum_{s=1}^{t}\left(\Big(\hat{x}_s^\top \nu - y_s\Big)^2 + \nu^\top(\Sigma_n(c) - \Sigma_n(c_s))\nu + \lambda c\right).$$

We indeed show that this empirical loss uniformly converges to the expected loss $\ell(c, \nu)$ and is almost convex:

Proposition 2. *With probability at least $1 - 3\delta$, for all $t \ge 1$, $c \in [0, 1]$ and $\nu \in \mathbb{R}^d$ s.t. $\|\nu\| \le S$, it holds that*

$$\left|\hat{L}^{kc}_t(c, \nu) - t\ell(c, \nu)\right| \le 9S^2\sqrt{8t\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}$$

where

$$\begin{aligned}\beta_t &= R^2\left(d + 2\sqrt{d\ln\frac{3t(t+1)}{\delta}} + 2\ln\frac{3t(t+1)}{\delta}\right) \\ &\quad + S^2\left(1 + 2\sqrt{\frac{\ln\frac{3t(t+1)}{\delta}}{d}}\right) = \widetilde{\mathcal{O}}(R^2 d + S^2).\end{aligned}$$

Furthermore, under the same event, the regularized loss

$$\hat{L}^{kc,R}_t(c, \nu) = \hat{L}^{kc}_t(c, \nu) + \gamma_t\|\nu\|^2 \tag{1}$$

for $\gamma_t = 2\sqrt{8t\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}$ is quadratic and strictly convex in ν for all $t \ge 1$.

Proof Sketch. As a first step, we show how to decompose the loss difference into the following quadratic terms:

$$\begin{aligned}\hat{L}^{kc}_t(c, \nu) - t\ell(c, \nu) &= \nu^\top\left(\sum_{s=1}^{t}\Big(\hat{x}_s\hat{x}_s^\top - \Sigma_x - \Sigma_n(c_s)\Big)\right)\nu \\ &\quad + \theta^{*\top}\left(\sum_{s=1}^{t}x_s x_s^\top - \Sigma_x\right)\theta^* \\ &\quad - 2\nu^\top\left(\sum_{s=1}^{t}x_s x_s^\top - \Sigma_x\right)\theta^* \\ &\quad - 2\nu^\top\left(\sum_{s=1}^{t}n_s x_s^\top\right)\theta^*.\end{aligned}$$

In particular, the first term already incorporates the covariance transfer from $\Sigma_n(c_s)$ to $\Sigma_n(c)$, leaving a well-behaved error term that depends on all the past payments, even when different costs were paid. The first three terms are quadratic forms w.r.t. the error between the empirical covariance of subgaussian vectors and their real covariance. The last term can also be written as such, noting that

$$\nu^\top n_s x_s^\top \theta^* = \begin{pmatrix}0\\ \nu\end{pmatrix}^\top\left(\begin{pmatrix}x_s\\ n_s\end{pmatrix}\begin{pmatrix}x_s\\ n_s\end{pmatrix}^\top - \begin{pmatrix}\Sigma_x & 0\\ 0 & \Sigma_n(c_s)\end{pmatrix}\right)\begin{pmatrix}\theta^*\\ 0\end{pmatrix}.$$

Using the inequality $x^\top A y \le \|x\|\|y\|\|A\|_{op}$, and given that the norms of ν, θ^* are bounded, we can uniformly bound the loss difference by upper bounding the operator norms of the covariance error terms. To bound these norms, we adapt martingale matrix concentration results (Tropp 2012) and apply them to the empirical covariance of subgaussian

Algorithm 1: Online Regression with Paid Features – Known Noise Covariances

Require: $\delta \in (0,1), K \in \mathbb{N}$
Initialize: $\hat{L}_0^{kc,R}(c,\nu) = 0$
for $t = 1, ..., T$ **do**
  for $k = 1, ..., K$ **do**
    Calculate $\hat{\nu}_{t-1}(k) \in \arg\min_{\nu:\|\nu\|\le S} \hat{L}_{t-1}^{kc,R}(k/K, \nu)$
  end for
  Set $k_{t-1} \in \arg\min_{k\in[K]} \hat{L}_{t-1}^{kc,R}(k/K, \hat{\nu}_{t-1}(k))$
  Pay $c_t = k_{t-1}/K$ and observe $\hat{x}_t = \hat{x}_t(c_t)$
  Predict $\hat{y}_t = \hat{x}_t^\top \hat{\nu}_{t-1}(k_{t-1})$ and observe y_t
  Update $\hat{L}_t^{kc,R}(k/K, \nu)$ for all $k \in [K]$
end for

vectors. In particular, if Z_t is a conditionally subgaussian vector martingale of conditional covariance Σ_t, it holds with high probability that

$$\left\| \sum_{s=1}^{t} Z_s Z_s^\top - \Sigma_s \right\|_{op} < \gamma_t/2$$

(see Lemma 6 in Section D for the precise statement and proof). This leads to the uniform concentration of the loss. One direct implication is the strict convexity of the regularized loss – noting that the loss is quadratic in ν and that $\hat{x}_t$ is $2R^2$-subgaussian, a regularization term of $\gamma_t\|\nu\|^2$ is enough to turn the quadratic loss term into strictly convex. □

Remark 3. *The concentration bound of Proposition 2 can be easily extended to the case where the regression output is noisy, namely, $y_t = x_t^\top \theta^* + \eta_t$ for conditionally zero-mean subgaussian noise η_t. This extra noise does not affect the algorithm or the concentration rates; see further discussion in Section B.2.*

A natural algorithmic approach is to minimize the (regularized) empirical loss – which we indeed do in Algorithm 1. Specifically, the algorithm minimizes this loss (w.r.t. the linear predictor ν) for every cost on a dense discretized grid $c \in \{k/K\}_{k\in[K]}$ and pays the cost

$$c_t \in \arg\min_{c\in\{k/K\}_{k\in[K]}} \min_{\nu:\|\nu\|\le S} \hat{L}_t^{kc,R}(c,\nu)$$

that achieves the minimal empirical loss on this grid. We note that the grid is only used to facilitate the optimal cost calculation; as long as it is sufficiently fine, it does not affect performance. Then, upon observing $\hat{x}_t(c_t)$, a prediction $\hat{y}_t(\hat{x}_t) = \hat{x}^\top \nu_t$ is calculated using the best empirical linear predictor for c_t:

$$\nu_t = \hat{\nu}_{t-1}(c_t) \in \arg\min_{\nu:\|\nu\|\le S} \hat{L}_t^{kc,R}(c_t,\nu).$$

This algorithm enjoys the following performance bound (see Section B for the proof):

Theorem 1. *For any $T \ge 1$, set $\delta = 1/T$ and $K = \lceil \lambda T \rceil$. Then, the regret of Algorithm 1 is bounded by*

$$\mathrm{Reg}(T) = \widetilde{\mathcal{O}}\Big(S^2(R^2 d + S^2)\sqrt{T} + \lambda\Big).$$

Proof Sketch. For any c, the uniform concentration in Proposition 2 guarantees that the empirical loss $\hat{L}_{t-1}^{kc}$ is similar to the expected one ℓ both for $\hat{\nu}_{t-1}(c)$ and for $\nu^*(c)$. Since $\hat{\nu}_{t-1}(c)$ outperforms $\nu^*(c)$ on the empirical loss $\hat{L}_{t-1}^{kc}$, the loss similarity implies that it cannot be much worse than $\nu^*(c)$ on the expected loss ℓ for any payment $c \in [0,1]$. The suboptimality of $\hat{\nu}_{t-1}(c)$ compared to $\nu^*(c)$ will be proportional to the maximal error between the average empirical loss $\frac{1}{t}\hat{L}_t^{kc}$ and the true loss ℓ: by Proposition 2, it is of order $\widetilde{\mathcal{O}}\big(S^2\beta_t/\sqrt{t}\big)$.

As for the choice of payment, the information sharing enables us to collect data on all costs simultaneously, without the need for explicit exploration. Thus, we can choose the best payment greedily on a payment grid, and the uniform concentration will again imply that the best empirical payment on the grid will perform similarly to the best cost on the grid (up to a similar error of $\widetilde{\mathcal{O}}\big(S^2\beta_t/\sqrt{t}\big)$). All that remains is to relate the optimal cost on the chosen grid to the best continuous payment: we do so leveraging the one-sided Lipschitzness of the loss, as stated in Claim 2. In particular, we choose a sufficiently fine grid to ensure that the discretization error is negligible. Thus, the instantaneous error compared to the optimal loss ℓ^* is of order $\widetilde{\mathcal{O}}\big(S^2\beta_t/\sqrt{t}\big)$, and accumulating it across all rounds leads to the desired regret bound. □

We now discuss two important aspects of our algorithm: its computational tractability and statistical efficiency (optimality of the regret bound).

Tractability of the loss minimization. One caveat in the covariance correction term $\Sigma_n(c) - \Sigma_n(c_s)$ is that it might render the loss nonconvex. To mitigate this, we add a regularization term $\gamma_t\|\nu\|^2$ to our loss. Then, we show in Proposition 2 that with high probability, calculating $\hat{\nu}_{t-1}(k)$ requires minimizing a strictly convex loss over a ball, which can be done efficiently. In particular, cases where the objective is nonconvex on some rounds are of low probability, so the learner can just choose arbitrary costs and predictions without affecting the performance. Nonetheless, it is important to note that the regularization is not strictly necessary for computational efficiency; the minimization of the unregularized empirical loss $\hat{L}_t^{kc}$ is in fact an instance of the trust region subproblem that can be efficiently solved even if the quadratic problem is nonconvex (see, e.g., Section 8.2.7 in Beck 2014). We also note that one could further increase the regularization term to make the objective $\Omega(R^2 d + S^2)$-strongly convex, while only deteriorating the regret by constant factors, thereby greatly accelerating the loss minimization step.

Statistical efficiency. At a glance, the regret rate that we obtained does not seem tight: the information sharing makes the problem similar to online linear regression, for which an $\mathcal{O}(d \ln T)$ regret is achievable (Vovk 1997). Somewhat surprisingly, we show that this is not the case in our setting: even in one-dimensional problems, an $\Omega(\sqrt{T})$ is unavoidable (see proof in Section B.1):

Theorem 2. *For any algorithm and $T \ge 1$, there exists a one-dimensional instance with a known noise variance profile*

Algorithm 2: Online Regression with Paid Features – Unknown Noise Covariances

Require: $\delta \in (0,1), K \in \mathbb{N}$
for $k = 1, ..., K$ **do**
  Pay $c_k = k/K$ and $\hat{y}_k = 0$; observe $\hat{x}_k = \hat{x}_k(c_k)$ and y_k
  Set $\hat{L}^{uc}_K(k/K, \nu) = \left(\hat{x}_k^\top \nu - y_k\right)^2 + \lambda k/K$
end for
for $t = K+1, ..., T$ **do**
  for $k = 1, ..., K$ **do**
    Calculate $\hat{\nu}_{t-1}(k) \in \arg\min_{\nu:\|\nu\|\le S} \hat{L}^{uc}_{t-1}(k/K, \nu)$
  end for
  choose k_{t-1} according to eq. (2)
  Pay $c_t = k_{t-1}/K$ and observe $\hat{x}_t = \hat{x}_t(c_t)$
  Predict $\hat{y}_t = \hat{x}_t^\top \hat{\nu}_{t-1}(k_{t-1})$ and observe y_t
  Update $\hat{L}^{uc}_t(k_{t-1}/K, \nu)$
end for

such that $\mathbb{E}[R(T)] \ge \frac{\sqrt{T}}{240}$.

The construction fixes the feature noise to be $\mathcal{N}(0,1)$ for all payments $c < 1/2$ and the noise to be deterministically zero when $c \ge 1/2$. We then create two instances with known $\theta^* = 1$ and $x_t \sim \mathcal{N}(0, 1 \pm \epsilon)$ and show that, depending on the instance, it is optimal to play either $c = 0$ or $c = 1/2$, thus reducing the problem to a two-armed bandit instance. From there, we use techniques from (Lattimore and Szepesvári 2020) to derive the lower bound. We hypothesize that the $\sqrt{T}$ rate results from the discontinuity in the noise variance, and that for sufficiently smooth noise, a rate of $\mathcal{O}(\ln T)$ might still be achievable, but leave this study for future work.

As a final remark, we note that the lower bound can be immediately generalized to d-dimensional features. In particular, we can choose a feature distribution that first samples a single coordinate (uniformly at random) to be non-zero, and then assigns its value according to the lower bound construction. This effectively creates d independent instances, each running for approximately T/d rounds. The regret of each instance will be lower bounded by $\Omega(\sqrt{T/d})$, so the overall regret from the interaction will be bounded by $\Omega(d\sqrt{T/d}) = \Omega(\sqrt{Td})$.

5 Extension to Unknown Noise Covariances

We now tackle the more challenging case in which the noise covariances are not known in advance. In this situation, we unfortunately cannot share information between different costs as we did in the previous section, and resort to estimating the loss at a cost c only based on samples collected with this specific cost. Formally, we define the loss

$$\hat{L}^{uc}_t(c, \nu) = \sum_{s=1}^{t} \Big(\left(\hat{x}_s^\top \nu - y_s\right)^2 + \lambda c \Big) \mathbb{1}\{c_s = c\},$$

which fully overlaps with $\hat{L}^{kc}_t(c, \nu)$ when all samples are collected with $c_s = c$. Leveraging this insight, we utilize Proposition 2 to derive concentration results on $\hat{L}^{uc}_t(c, \nu)$ (see proof in Section C).

Corollary 3. *Fix $c \in [0,1]$ and let $N_t(c) = \sum_{s=1}^{t} \mathbb{1}\{c_s = c\}$. Then, with probability at least $1 - 3\delta$, for all $t \ge 1$ and $\nu \in \mathbb{R}^d$ s.t. $\|\nu\| \le S$, it holds that*

$$\left| \hat{L}^{uc}_t(c, \nu) - N_t(c)\ell(c, \nu) \right| \le 9S^2 \sqrt{8 N_t(c) \beta_t^2 \ln \frac{3dt(t+1)}{\delta}},$$

where $\beta_t = \widetilde{\mathcal{O}}(R^2 d + S^2)$ is specified at Proposition 2.

We remark that without the information sharing term $\nu^\top(\Sigma_n(c) - \Sigma_n(c_s))\nu$, the loss $\hat{L}^{uc}_t(c, \nu)$ is always convex; therefore, the regularization is no longer needed and is omitted from the corollary.[3] A byproduct of separately estimating the loss for each cost is the need for exploration: as playing one cost does not yield sufficient information on all potential payments, we must test a diverse set of payments to identify the optimal cost c^*. We encourage our algorithm to explore by adding *optimism* to $\hat{L}^{uc}_t$ (Auer 2002), that is, replacing the empirical loss by the smallest plausible loss allowed by the confidence interval. Intuitively, the optimism introduces penalty terms that reduce the loss at payment levels that were not sufficiently observed, biasing toward their exploration.

The full algorithm is depicted in Algorithm 2. We first limit the learner to a fixed grid of potential payments $c \in \{k/K\}_{k \in [K]}$; exploration is costly, so only these costs will be explored. We play each of these costs once as a loss initialization step. Then, at each round, the algorithm chooses the cost index that minimizes the optimistic average loss:

$$k_{t-1} \in \arg\min_{k \in [K]} \left\{ \frac{\hat{L}^{uc}_{t-1}(k/K, \hat{\nu}_{t-1}(k))}{N_{t-1}(k/K)} - 9S^2 \sqrt{\frac{8\beta_t^2 \ln \frac{3dt(t+1)}{\delta}}{N_{t-1}(k/K)}} \right\}, \quad (2)$$

where

$$\hat{\nu}_{t-1}(k) \in \arg\min_{\nu:\|\nu\| \le S} \hat{L}^{uc}_{t-1}(k/K, \nu)$$

is a linear predictor that minimizes the empirical loss. As before, after paying $c_t = k_{t-1}/K$, the learner observes $\hat{x}_t(c_t)$ and uses $\hat{\nu}_{t-1}(k_{t-1})$ to predict $\hat{y}_t = \hat{x}_t(c_t)^\top \hat{\nu}_{t-1}(k_{t-1})$. Afterwards, upon observing y_t, the learner only updates the loss at c_t and continues to the next round.

The algorithm enjoys the following regret bound (see Section C for the proof):

Theorem 3. *For any $T \ge 1$, set $K = \left\lceil \frac{T^{1/3}\lambda^{2/3}}{(S^2(R^2 d + S^2))^{2/3}} \right\rceil$ and $\delta = \frac{1}{KT}$. Then, the regret of Algorithm 2 is bounded by*

$$\mathrm{Reg}(T) = \widetilde{\mathcal{O}}\Big(\left(S^2(R^2 d + S^2)\right)^{2/3} \lambda^{1/3} T^{2/3} \Big).$$

Proof Sketch. The proof again relies on the loss concentration (Corollary 3), now combined with bandit techniques. In particular, we show that with high probability, the instantaneous regret of playing a cost c_t compared to the best cost on the grid $\{k/K\}_{k \in [K]}$ is inversely proportional to

[3] The regularization could still be added to ensure strong convexity without affecting the results while requiring only minor algorithmic modifications.

the number of times it was previously played: of the order $\widetilde{\mathcal{O}}\Big(S^2\beta_{t-1}/\sqrt{N_{t-1}(c_t)}\Big)$. On the other hand, working with a grid of K costs leads to an instantaneous discretization error of λ/K (due to the Lipschitzness proved in Claim 2). Therefore, the cumulative regret is bounded (w.h.p.) by

$$\mathrm{Reg}(T) \lesssim \sum_{t=K+1}^{T}\left(\frac{S^2\beta_T}{\sqrt{N_{t-1}(c_t)}}+\frac{\lambda}{K}\right) \approx S^2\beta_T\sqrt{KT}+\frac{\lambda T}{K}.$$

The last relation uses standard bandit arguments: upon choosing a cost c_t, its corresponding count must increase, decreasing the denominator in future rounds in which this cost is played. Thus, the sum of the inverse counts cannot be too big. Fixing a grid size that optimally balances the two terms leads to the stated regret bound. □

Relation to Lipschitz Bandits. The discretization approach on the costs and local loss estimation resembles existing approaches in the Lipschitz bandit literature (also called $\mathcal{X}$-armed bandits, Bubeck et al. 2011); there, a regret bound of $\widetilde{\mathcal{O}}(T^{2/3})$ is obtained for one-dimensional problems. Indeed, although our problem is not Lipschitz, the one-sided Lipschitzness (Claim 2) might suffice for such a reduction. However, the resulting Lipschitz instance will be $d+1$-dimensional (representing both costs and predictions), thus leading to a regret bound of $\widetilde{\mathcal{O}}(T^{\frac{d+2}{d+3}})$. In contrast, we show how to separately optimize over the linear predictors, so that we can perform the discretization only in one dimension (somewhat similar to Tullii et al. 2024). This allows us to achieve an improved rate of $\widetilde{\mathcal{O}}(T^{2/3})$.

Next, we discuss the tightness of Theorem 3. Our discretization-based bandit approach yields a worse regret compared to the $\widetilde{\mathcal{O}}(\sqrt{T})$ rate with known noise covariances. This degradation is unavoidable, as we prove in the following one-dimensional lower bound:

Theorem 4. *For any algorithm and $T \geq 1$, there exists a one-dimensional instance with an unknown noise variance profile such that $\mathbb{E}[R(T)] \geq \frac{T^{2/3}}{256}$.*

The full proof can be found at Section C.1.

Proof Sketch. To derive this bound, we first show that for one dimensional problems with the parameters $\Sigma_x \triangleq \sigma_x^2 = 1$, $\theta^* = 1$ and $\lambda = 1/2$, the variance profile $\Sigma_n(c) \triangleq \sigma_n^2(c) = f(c) = \frac{1-c}{1+c}$ achieves a constant loss $\ell^*(c) = 1/2$ for all $c \in [0,1]$. Then, we carefully devise K noise variance profiles ('instances') that smoothly deviate from $f(c)$ only at a small interval ('modified interval') of size $\approx 1/K$, with no interval overlaps between instances. In particular, the minimal loss inside these intervals is slightly lower than $1/2$ (by $\approx 1/K$). We remark that although our problem is very different from Lipschitz bandits, our construction draws inspiration from characteristics of hard Lipschitz bandit instances, and we indeed obtain the same bound as Kleinberg (2004).

We then compare the behavior of any algorithm in these instances to that on a nominal instance with $\sigma_n^2(c) = f(c)$. By the pigeonhole principle, there must be an instance k^* such that its modified interval was sampled less than T/K times on average in the nominal instance. Since the nominal instance and k^* are identical outside the interval (no information on the instance is gained), we show that for $K \approx T^{1/3}$, the algorithm cannot behave too differently on k^*; using information-theoretic tools (Garivier, Ménard, and Stoltz 2019), we show that it chooses costs in the modified intervals no more than $3T/4$ times. Since the regret for not choosing a cost in the modified interval is roughly $1/K$ (loss of $1/2$ instead of $1/2 - 1/K$), this implies a lower bound of

$$\mathrm{Reg}(T) \gtrsim \left(T - \frac{3T}{4}\right)\frac{1}{K} = \Omega\Big(T^{2/3}\Big).$$

□

6 Summary and Future Work

In this work, we studied online linear regression problems with stochastic features that are corrupted by noise, and the learner can pay to reduce this noise. This setting is applicable, for example, in settings where features are measured in experiments, problems with privacy-protected information, and more. We focused on two variants of this problem: in the first, the learner has full knowledge of the noise covariance across all payment levels (e.g., privacy noise with known characteristics); in the second, the learner must estimate the effect of the payment on the quality of the observed features. In both cases, we devised learner algorithms that achieve order-optimal regret bounds as a function of the interaction length (up to logarithmic factors).

While the algorithms presented in this work achieve order-optimal horizon-dependence in their regret bounds, it is unclear whether these algorithms are optimal with respect to other problem parameters. In particular, all our lower bounds are derived on one-dimensional instances, and so the optimal dimension dependence remains to be determined. In addition, while both algorithms are polynomial, each time step requires solving a convex optimization problem over a grid of costs. It would be beneficial to derive more computationally efficient algorithms, including algorithms that build adaptive grids or avoid discretizing the payments.

Our work could also be extended beyond linear regression. Specifically, it would be interesting to extend our results to non-linear regression problems $y_t = f(x_t)$ (for possibly non-linear $f \in \mathcal{F}$), also potentially deviating from the quadratic loss. In this context, it is especially relevant to study *agnostic* settings: limiting the algorithm to a smaller class of predictors that does not include the mapping that minimizes the loss between $\hat{y}(x_t)$ and y_t. This might be necessary since the optimal predictors could be extremely complicated, as we discussed in Remark 1. Specifically, we limited our predictors to be linear, similarly to the mapping that created the original outputs y_t, but for more complex regression models, it might be beneficial to tailor the prediction class $\mathcal{P}$ to be different than the true regression model class $\mathcal{F}$.

Finally, in some situations, we need to pay for each feature separately and/or can obtain the same feature from multiple sources (for different payments and with different noise). It is then interesting to determine the optimal (combinatorial) payment scheme and devise algorithms that learn to optimally aggregate features gathered from different sources.

Acknowledgments

NM was supported by Israel Science Foundation research grant (ISF's No. 4118/25) and the Maimonides Fund's Future Scientists Center. NCB acknowledges the financial support from the MUR PRIN grant 2022EKNE5K (Learning in Markets and Society), the EU Horizon CL4-2021-HUMAN-01 research and innovation action under grant agreement 101070617, project ELSA, and the FAIR (Future Artificial Intelligence Research) project, funded by the NextGenerationEU program within the PNRR-PE-AI scheme.

References

Agarwal, A.; Harris, K.; Whitehouse, J.; and Wu, S. Z. 2023. Adaptive principal component regression with applications to panel data. *Advances in Neural Information Processing Systems*, 36: 77104–77118.

Agarwal, A.; Shah, D.; Shen, D.; and Song, D. 2019. On robustness of principal component regression. *Advances in Neural Information Processing Systems*, 32.

Agarwal, A.; and Singh, R. 2021. Causal Inference with Corrupted Data: Measurement Error, Missing Values, Discretization, and Differential Privacy. *arXiv preprint arXiv:2107.02780*.

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

Beck, A. 2014. *Introduction to nonlinear optimization: Theory, algorithms, and applications with MATLAB*. SIAM.

Belloni, A.; Rosenbaum, M.; and Tsybakov, A. B. 2017. Linear and conic programming estimators in high dimensional errors-in-variables models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 79(3): 939–956.

Bubeck, S.; Munos, R.; Stoltz, G.; and Szepesvári, C. 2011. X-Armed Bandits. *Journal of Machine Learning Research*, 12(5).

Cesa-Bianchi, N.; Shalev-Shwartz, S.; and Shamir, O. 2011. Online learning of noisy data. *IEEE Transactions on Information Theory*, 57(12): 7907–7931.

Chen, Y.; and Caramanis, C. 2013. Noisy and missing data regression: Distribution-oblivious support recovery. In *International Conference on Machine Learning*, 383–391. PMLR.

Datta, A.; and Zou, H. 2017. Cocolasso for high-dimensional error-in-variables regression. *Annals of Statistics*, 45(6): 2400–2426.

Duchi, J. C.; Jordan, M. I.; and Wainwright, M. J. 2013. Local privacy, data processing inequalities, and minimax rates. *arXiv preprint arXiv:1302.3203*.

Fallah, A.; Makhdoumi, A.; Malekian, A.; and Ozdaglar, A. 2024. Optimal and differentially private data acquisition: Central and local mechanisms. *Operations Research*, 72(3): 1105–1123.

Figueroa-Zúñiga, J. I.; Bayes, C. L.; Leiva, V.; and Liu, S. 2022. Robust beta regression modeling with errors-in-variables: a Bayesian approach and numerical applications. *Statistical Papers*, 63(3): 919–942.

Fukuchi, K.; Tran, Q. K.; and Sakuma, J. 2017. Differentially private empirical risk minimization with input perturbation. In *International Conference on Discovery Science*, 82–90. Springer.

Fuller, W. A. 2009. *Measurement error models*. John Wiley & Sons.

Garivier, A.; Ménard, P.; and Stoltz, G. 2019. Explore first, exploit next: The true shape of regret in bandit problems. *Mathematics of Operations Research*, 44(2): 377–399.

Ghosh, A.; and Roth, A. 2011. Selling privacy at auction. In *Proceedings of the 12th ACM conference on Electronic commerce*, 199–208.

Hsu, D.; Kakade, S. M.; and Zhang, T. 2012. A tail inequality for quadratic forms of subgaussian random vectors. *Electronic Communications in Probability*, 17: 1–6.

Hsu, J.; Gaboardi, M.; Haeberlen, A.; Khanna, S.; Narayan, A.; Pierce, B. C.; and Roth, A. 2014. Differential privacy: An economic method for choosing epsilon. In *2014 IEEE 27th Computer Security Foundations Symposium*, 398–410. IEEE.

Kasiviswanathan, S. P.; Lee, H. K.; Nissim, K.; Raskhodnikova, S.; and Smith, A. 2011. What can we learn privately? *SIAM Journal on Computing*, 40(3): 793–826.

Kaul, A.; and Koul, H. L. 2015. Weighted ℓ_1-penalized corrected quantile regression for high dimensional measurement error models. *Journal of Multivariate Analysis*, 140: 72–91.

Kim, J.-h.; Yun, S.-Y.; Jeong, M.; Nam, J.; Shin, J.; and Combes, R. 2023. Contextual linear bandits under noisy features: Towards bayesian oracles. In *International Conference on Artificial Intelligence and Statistics*, 1624–1645. PMLR.

Kleinberg, R. 2004. Nearly tight bounds for the continuum-armed bandit problem. *Advances in Neural Information Processing Systems*, 17.

Lattimore, T.; and Szepesvári, C. 2020. *Bandit algorithms*. Cambridge University Press.

Loh, P.-L.; and Wainwright, M. J. 2011. High-dimensional regression with noisy and missing data: Provable guarantees with non-convexity. *Advances in neural information processing systems*, 24.

Nissim, K.; Orlandi, C.; and Smorodinsky, R. 2012. Privacy-aware mechanism design. In *Proceedings of the 13th ACM conference on electronic commerce*, 774–789.

Nissim, K.; Vadhan, S.; and Xiao, D. 2014. Redrawing the boundaries on purchasing data from privacy-sensitive individuals. In *Proceedings of the 5th conference on Innovations in theoretical computer science*, 411–422.

Reilly, P. M.; and Patino-Leal, H. 1981. A Bayesian study of the error-in-variables model. *Technometrics*, 23(3): 221–231.

Rosenbaum, M.; and Tsybakov, A. B. 2010. Sparse Recovery under Matrix Uncertainty. *The Annals of Statistics*, 38(5): 2620–2651.

Rosenbaum, M.; and Tsybakov, A. B. 2013. Improved matrix uncertainty selector. In *From Probability to Statistics and Back: High-Dimensional Models and Processes–A Festschrift in Honor of Jon A. Wellner*, volume 9, 276–291. Institute of Mathematical Statistics.

Smith, A.; Thakurta, A.; and Upadhyay, J. 2017. Is interaction necessary for distributed private learning? In *2017 IEEE Symposium on Security and Privacy (SP)*, 58–77. IEEE.

Tropp, J. A. 2012. User-friendly tail bounds for sums of random matrices. *Foundations of computational mathematics*, 12: 389–434.

Tullii, M.; Gaucher, S.; Merlis, N.; and Perchet, V. 2024. Improved algorithms for contextual dynamic pricing. *Advances in Neural Information Processing Systems*, 37: 126088–126117.

Ungarala, S.; and Bakshi, B. R. 2000. A multiscale, Bayesian and error-in-variables approach for linear dynamic data rectification. *Computers & Chemical Engineering*, 24(2-7): 445–451.

Van Der Hoeven, D.; Pike-Burke, C.; Qiu, H.; and Cesa-Bianchi, N. 2023. Trading-off payments and accuracy in online classification with paid stochastic experts. In *International Conference on Machine Learning*, 34809–34830. PMLR.

Vovk, V. 1997. Competitive on-line linear regression. *Advances in Neural Information Processing Systems*, 10.

Wang, D.; and Xu, J. 2019. On sparse linear regression in the local differential privacy model. In *International Conference on Machine Learning*, 6628–6637. PMLR.

Wang, W.; Ying, L.; and Zhang, J. 2018. The value of privacy: Strategic data subjects, incentive mechanisms, and fundamental limits. *ACM Transactions on Economics and Computation (TEAC)*, 6(2): 1–26.

Yang, M.; Guo, T.; Zhu, T.; Tjuawinata, I.; Zhao, J.; and Lam, K.-Y. 2024. Local differential privacy and its applications: A comprehensive survey. *Computer Standards & Interfaces*, 89: 103827.

Zheng, K.; Cai, T.; Huang, W.; Li, Z.; and Wang, L. 2020. Locally differentially private (contextual) bandits learning. *Advances in Neural Information Processing Systems*, 33: 12300–12310.

A Properties of the Problem

Claim 1. *It holds that*

$$\ell(c,\nu) = \nu^\top \Sigma_{\hat{x}}(c)\nu - 2\nu^\top \Sigma_x \theta^* + {\theta^*}^\top \Sigma_x \theta^* + \lambda c.$$

Moreover, let $\bar{\nu}(c) = (\Sigma_{\hat{x}}(c))^{-1}\Sigma_x\theta^*$. *If* $\|\bar{\nu}(c)\| \le S$ *then* $\nu^*(c) = \bar{\nu}(c)$ *and*

$$\ell^*(c) = {\theta^*}^\top \Sigma_x \theta^* - {\theta^*}^T \Sigma_x \Sigma_{\hat{x}}(c)^{-1}\Sigma_x\theta^* + \lambda c.$$

Proof. By definition, it holds that

$$\begin{aligned}
\ell(c,\nu) &= \mathbb{E}_{x\sim D_x, n\sim D_n(c)}\left[\left((x+n)^\top \nu - x^\top\theta^*\right)^2\right] + \lambda c \\
&= \mathbb{E}_{x\sim D_x, n\sim D_n(c)}\left[\left(n^\top\nu + x^\top(\nu-\theta^*)\right)^2\right] + \lambda c \\
&\overset{(1)}{=} \mathbb{E}_{x\sim D_x, n\sim D_n(c)}\left[\nu^\top nn^\top\nu + (\nu-\theta^*)^\top xx^\top(\nu-\theta^*)\right] + \lambda c \\
&\overset{(2)}{=} \nu^\top\Sigma_n(c)\nu + (\nu-\theta^*)^\top\Sigma_x(\nu-\theta^*) + \lambda c \\
&= \nu^\top\Sigma_n(c)\nu + \nu^\top\Sigma_x\nu - 2\nu^\top\Sigma_x\theta^* + {\theta^*}^\top\Sigma_x\theta^* + \lambda c \\
&= \nu^\top\Sigma_{\hat{x}}(c)\nu - 2\nu^\top\Sigma_x\theta^* + {\theta^*}^\top\Sigma_x\theta^* + \lambda c.
\end{aligned}$$

Relation (1) holds since x and $n(c)$ are independent and $n(c)$ has zero mean, and relation (2) holds by the definition of the covariances. The function is strongly convex w.r.t. ν (due to the regularity of the covariances), so there exists a single optimal solution, obtainable by the first-order optimality condition:

$$2\Sigma_{\hat{x}}(c)\nu - 2\Sigma_x\theta^* = 0.$$

Reorganizing, we get that $\bar{\nu}(c) = (\Sigma_{\hat{x}}(c))^{-1}\Sigma_x\theta^*$ is the unique unconstrained global minimizer. In particular, if it satisfies the constraint $\|\bar{\nu}(c)\| \le S$, it is also the constrained minimizer, and thus $\nu^*(c) = (\Sigma_{\hat{x}}(c))^{-1}\Sigma_x\theta^*$. Substituting back to $\ell(c,\nu)$ directly leads to the stated optimal loss:

$$\begin{aligned}
\ell^*(c) &= \nu^*(c)^\top\Sigma_{\hat{x}}(c)\nu^*(c) - 2\nu^*(c)^\top\Sigma_x\theta^* + {\theta^*}^\top\Sigma_x\theta^* + \lambda c \\
&= {\theta^*}^\top\Sigma_x(\Sigma_{\hat{x}}(c))^{-1}\Sigma_{\hat{x}}(c)(\Sigma_{\hat{x}}(c))^{-1}\Sigma_x\theta^* - 2{\theta^*}^\top\Sigma_x(\Sigma_{\hat{x}}(c))^{-1}\Sigma_x\theta^* + {\theta^*}^\top\Sigma_x\theta^* + \lambda c \\
&= {\theta^*}^\top\Sigma_x\theta^* - {\theta^*}^\top\Sigma_x(\Sigma_{\hat{x}}(c))^{-1}\Sigma_x\theta^* + \lambda c.
\end{aligned}$$

Note that $\Sigma_{\hat{x}}(c)$ is invertible by the assumption $\Sigma_{\hat{x}}(1) \succ 0$ and the monotonicity of the cost, so both expressions for $\nu^*(c)$ and $\ell^*(c)$ are well-defined. □

Claim 3. *Assume* $d = 1$. *Then, for* $\bar{\nu}(c) = (\Sigma_{\hat{x}}(c))^{-1}\Sigma_x\theta^*$, *it holds that* $|\bar{\nu}(c)| \le S$ *for all* $c \in [0,1]$.

Proof. In dimension one, all parameters $\Sigma_x, \Sigma_n(c)$ and θ^* are scalars, and by the assumptions, $\Sigma_x, \Sigma_n(c) \ge 0$, $\Sigma_x + \Sigma_n(c) > 0$ and $|\theta^*| \le S$. Given this, the claim immediately holds from substitution:

$$|\bar{\nu}(c)| = \left|(\Sigma_{\hat{x}}(c))^{-1}\Sigma_x\theta^*\right| = \underbrace{\left|\frac{\Sigma_x}{\Sigma_x+\Sigma_n(c)}\right|}_{\le 1}\underbrace{|\theta^*|}_{\le S} \le S.$$

□

Claim 2. *The loss* $\ell^*(c)$ *is* λ*-one-sided Lipschitz: for any* $0 \le c_1 \le c_2 \le 1$, *it holds that* $\ell^*(c_2) \le \ell^*(c_1) + \lambda(c_2 - c_1)$.

Proof. By the noise covariance monotonicity assumption, we have that

$$\Sigma_{\hat{x}}(c_2) = \Sigma_x + \Sigma_n(c_2) \preceq \Sigma_x + \Sigma_n(c_1) = \Sigma_{\hat{x}}(c_1),$$

and thus, for any ν and $0 \le c_1 \le c_2 \le 1$, it holds by Claim 1 that

$$\begin{aligned}
\ell(c_1,\nu) &= \nu^\top\Sigma_{\hat{x}}(c_1)\nu - 2\nu^\top\Sigma_x\theta^* + {\theta^*}^\top\Sigma_x\theta^* + \lambda c_1 \\
&\ge \nu^\top\Sigma_{\hat{x}}(c_2)\nu - 2\nu^\top\Sigma_x\theta^* + {\theta^*}^\top\Sigma_x\theta^* + \lambda c_1 && (\Sigma_{\hat{x}}(c_1) \succeq \Sigma_{\hat{x}}(c_2)) \\
&= \nu^\top\Sigma_{\hat{x}}(c_2)\nu - 2\nu^\top\Sigma_x\theta^* + {\theta^*}^\top\Sigma_x\theta^* + \lambda c_2 + \lambda(c_1 - c_2) \\
&= \ell(c_2,\nu) + \lambda(c_1 - c_2).
\end{aligned}$$

Minimizing over ν of norm smaller than S in both sides of the inequality, we get

$$\ell^*(c_1) \ge \ell^*(c_2) + \lambda(c_1 - c_2),$$

and reorganizing this inequality concludes the proof. □

Lemma 1. *For any arbitrary history-dependent prediction rule $\hat{y}_t : \mathbb{R}^d \mapsto \mathbb{R}$ and any arbitrary sequence $\nu_1, \ldots, \nu_T \in \mathbb{R}^d$, both potentially depend on I^o_{t-1}, it holds that*

$$\mathrm{Reg}(T) = \mathbb{E}\left[\sum_{t=1}^{T}\Big((\hat{y}_t(\hat{x}_t) - y_t)^2 - \Big(\big(\hat{x}_t^\top \nu_t - y_t\big)^2\Big)\Big)\right] + \mathbb{E}\left[\sum_{t=1}^{T}(\ell(c_t, \nu_t) - \ell^*)\right].$$

In particular, if the algorithm only outputs linear predictions $\hat{y}_t(\hat{x}_t) = \hat{x}_t^\top \nu_t$, then

$$\mathrm{Reg}(T) = \mathbb{E}\left[\sum_{t=1}^{T}(\ell(c_t, \nu_t) - \ell^*)\right].$$

Proof. Recall that c_t and ν_t are assumed to be completely determined by I^o_{t-1} and that x_t, n_t are drawn respectively from $D_x, D_n(c_t)$ given I_{t-1} (and therefore, also given I^o_{t-1}). Then, for any t, we can write

$$\begin{aligned}
\mathbb{E}[\ell(c_t, \nu_t)] &= \mathbb{E}\left[\mathbb{E}_{x\sim D_x, n\sim D_n(c_t)}\left[\Big((x+n)^\top \nu_t - x^\top\theta^*\Big)^2 + \lambda c_t | c_t, \nu_t\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}_{x\sim D_x, n\sim D_n(c_t)}\left[\Big((x+n)^\top \nu_t - x^\top\theta^*\Big)^2 + \lambda c_t | I^o_{t-1}\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\Big((x_t+n_t)^\top \nu_t - x_t^\top\theta^*\Big)^2 + \lambda c_t | I^o_{t-1}\right]\right] \\
&= \mathbb{E}\Big[\big(\hat{x}_t^\top \nu_t - x_t^\top\theta^*\big)^2 + \lambda c_t\Big]. \qquad (3)
\end{aligned}$$

Starting from the regret definition and using this identity, we decompose it as follows.

$$\begin{aligned}
\mathrm{Reg}(T) &= \mathbb{E}\left[\sum_{t=1}^{T}\Big((\hat{y}_t(\hat{x}_t) - y_t)^2 + \lambda c_t\Big)\right] - T\ell^* \\
&= \mathbb{E}\left[\sum_{t=1}^{T}\Big((\hat{y}_t(\hat{x}_t) - y_t)^2 + \lambda c_t - \ell(c_t, \nu_t)\Big)\right] + \mathbb{E}\left[\sum_{t=1}^{T}\ell(c_t, \nu_t)\right] - T\ell^* \\
&= \mathbb{E}\left[\sum_{t=1}^{T}\Big((\hat{y}_t(\hat{x}_t) - y_t)^2 + \lambda c_t - \Big(\big(\hat{x}_t^\top \nu_t - y_t\big)^2 + \lambda c_t\Big)\Big)\right] + \mathbb{E}\left[\sum_{t=1}^{T}\ell(c_t, \nu_t)\right] - T\ell^* \qquad \text{(by eq. (3))} \\
&= \mathbb{E}\left[\sum_{t=1}^{T}\Big((\hat{y}_t(\hat{x}_t) - y_t)^2 - \Big(\big(\hat{x}_t^\top \nu_t - y_t\big)^2\Big)\Big)\right] + \mathbb{E}\left[\sum_{t=1}^{T}(\ell(c_t, \nu_t) - \ell^*)\right].
\end{aligned}$$

Specifically, if $\hat{y}_t(\hat{x}_t) = \hat{x}_t^\top \nu_t$, then the first term is canceled out and we get the desired expression. □

Claim 4. *Assume that D_x and $D_n(c)$ are zero-mean Gaussian independent distributions and assume that for $\bar{\nu}(c) = (\Sigma_x + \Sigma_n(c))^{-1}\Sigma_x\theta^*$, it holds that $\|\bar{\nu}(c)\| \le S$ for all $c \in [0,1]$. Then,*

$$\mathrm{Reg}(T) \ge \mathbb{E}\left[\sum_{t=1}^{T}(\ell^*(c_t) - \ell^*(c^*))\right].$$

Proof. Denote the history up to time t (up to c_t but not x_t or n_t) by I_{t-1}. We use the fact that the mean squared error is always minimized by the conditional expectation; that is, for any two random variables $Z \in \mathbb{R}^d, W \in \mathbb{R}$ and any function f, it holds that

$$\mathbb{E}\Big[\big(f(Z) - W\big)^2\Big] \ge \mathbb{E}\Big[\big(\mathbb{E}[W|Z] - W\big)^2\Big].$$

In particular, for $Z = \hat{x}_t$ and $W = y_t = x_t^\top\theta^*$, it holds that

$$\begin{aligned}
\mathbb{E}\Big[(\hat{y}_t(\hat{x}_t) - y_t)^2 | I_{t-1}\Big] &\ge \mathbb{E}\Big[\big(\mathbb{E}\big[x_t^\top\theta^* | \hat{x}_t, I_{t-1}\big] - x_t^\top\theta^*\big)^2 | I_{t-1}\Big] = \mathbb{E}\left[\Big(\big(\mathbb{E}[x_t|\hat{x}_t, I_{t-1}] - x_t\big)^\top\theta^*\Big)^2 | I_{t-1}\right] \\
&= {\theta^*}^\top \mathbb{E}\Big[\big(\mathbb{E}[x_t|\hat{x}_t, I_{t-1}] - x_t\big)\big(\mathbb{E}[x_t|\hat{x}_t, I_{t-1}] - x_t\big)^\top | I_{t-1}\Big]\theta^*.
\end{aligned}$$

Conditioned on the history, the vector $z_t = \begin{pmatrix} x_t \\ \hat{x}_t \end{pmatrix}$ is a Gaussian random vector of mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance $\begin{pmatrix} \Sigma_x & \Sigma_x \\ \Sigma_x & \Sigma_{\hat{x}}(c_t) \end{pmatrix}$ (due to the independence of x_t and n_t). Thus, by the classic results on the conditional expectation of Gaussian vectors, we have

$$\mathbb{E}[x_t|\hat{x}_t, I_{t-1}] = \Sigma_x(\Sigma_x + \Sigma_n(c_t))^{-1}\hat{x}_t.$$

Substituting back, we get

$$\begin{aligned}
\mathbb{E}\Big[(\hat{y}_t(\hat{x}_t) - y_t)^2|I_{t-1}\Big] &\geq \theta^{*\top}\mathbb{E}\Big[\Big(\Sigma_x(\Sigma_x + \Sigma_n(c_t))^{-1}\hat{x}_t - x_t\Big)\Big(\Sigma_x(\Sigma_x + \Sigma_n(c_t))^{-1}\hat{x}_t - x_t\Big)^\top|I_{t-1}\Big]\theta^* \\
&= \theta^{*\top}\Sigma_x(\Sigma_x + \Sigma_n(c_t))^{-1}\underbrace{\mathbb{E}\big[\hat{x}_t\hat{x}_t^\top|I_{t-1}\big]}_{=\Sigma_x+\Sigma_n(c_t)}(\Sigma_x + \Sigma_n(c_t))^{-1}\Sigma_x\theta^* - \theta^{*\top}\Sigma_x(\Sigma_x + \Sigma_n(c_t))^{-1}\underbrace{\mathbb{E}\big[\hat{x}_t x_t^\top|I_{t-1}\big]}_{=\Sigma_x}\theta^* \\
&\quad - \theta^{*\top}\underbrace{\mathbb{E}\Big[x_t\hat{x}_t\hat{|}I_{t-1}\Big]}_{=\Sigma_x}(\Sigma_x + \Sigma_n(c_t))^{-1}\Sigma_x\theta^* + \theta^{*\top}\underbrace{\mathbb{E}\big[x_t x_t^\top|I_{t-1}\big]}_{=\Sigma_x}\theta^* \\
&= \theta^{*\top}\Sigma_x\theta^* - \theta^{*\top}\Sigma_x(\Sigma_x + \Sigma_n(c_t))^{-1}\Sigma_x\theta^* \\
&= \ell^*(c_t) - \lambda c_t,
\end{aligned}$$

where the last equality is by Claim 1. Substituting into the regret, we get

$$\begin{aligned}
\mathrm{Reg}(T) &= \mathbb{E}\left[\sum_{t=1}^{T}\Big((\hat{y}_t(\hat{x}_t) - y_t)^2 + \lambda c_t\Big)\right] - T\ell^* \\
&= \mathbb{E}\left[\sum_{t=1}^{T}\Big(\mathbb{E}\Big[(\hat{y}_t(\hat{x}_t) - y_t)^2|I_{t-1}\Big] + \lambda c_t\Big)\right] - T\ell^* \\
&\geq \mathbb{E}\left[\sum_{t=1}^{T}((\ell^*(c_t) - \lambda c_t) + \lambda c_t)\right] - T\ell^* \\
&= \mathbb{E}\left[\sum_{t=1}^{T}(\ell^*(c_t) - \ell^*)\right].
\end{aligned}$$

□

Claim 5. *Assume that $x, n \in \mathbb{R}^d$ are R^2-subgaussian independent vectors s.t. $E[x] = \bar{x}$ and $\mathbb{E}[n] = 0$, where $\|\bar{x}\| \leq S$. Then,*

$$\max_{c\in[0,1],\nu:\|\nu\|\leq S}\left\{\mathbb{E}\left[\Big((x+n)^\top\nu - x^\top\theta^*\Big)^2\right] + \lambda c\right\} \leq 6S^2\big(R^2d + S^2\big) + \lambda.$$

Proof. For any $c \in [0,1]$ and $\nu \in \mathbb{R}^d$ s.t. $\|\nu\| \leq S$, it holds that

$$\begin{aligned}
\mathbb{E}\left[\Big((x+n)^\top\nu - x^\top\theta^*\Big)^2\right] + \lambda c &\leq 2\mathbb{E}\left[\Big((x+n)^\top\nu\Big)^2\right] + 2\mathbb{E}\Big[\big(x^\top\theta^*\big)^2\Big] + \lambda && ((a+b)^2 \leq 2a^2 + 2b^2) \\
&\leq 2\mathbb{E}\big[S^2\|x+n\|^2\big] + 2\mathbb{E}\big[S^2\|x\|^2\big] + \lambda && \text{(Cauchy Schwartz)}
\end{aligned}$$

We now use the fact that x is R^2 subgaussian and $x+n$ is $2R^2$ subgaussian, both of expectation $\bar{x}$. Thus, by Lemma 5, and using the identity $\mathbb{E}\Big[\|Y - \mathbb{E}[Y]\|^2\Big] = \mathbb{E}\Big[\|Y\|^2\Big] - \|\mathbb{E}[Y]\|^2$ (which holds for any random vector), we get

$$\begin{aligned}
\mathbb{E}\left[\Big((x+n)^\top\nu - x^\top\theta^*\Big)^2\right] + \lambda c &\leq 2S^2\big(\mathbb{E}\big[\|x+n-\bar{x}\|^2\big] + \|\bar{x}\|^2\big) + 2S^2\big(\mathbb{E}\big[\|x-\bar{x}\|^2\big] + \|\bar{x}\|^2\big) + \lambda \\
&\leq 2S^2\big(2R^2d + S^2\big) + 2S^2\big(R^2d + S^2\big) + \lambda && \text{(Lemma 5)} \\
&\leq 6S^2\big(R^2d + S^2\big) + \lambda.
\end{aligned}$$

□

B Proofs for Known Noise Covariances

Proposition 2. *With probability at least $1-3\delta$, for all $t \geq 1$, $c \in [0,1]$ and $\nu \in \mathbb{R}^d$ s.t. $\|\nu\| \leq S$, it holds that*

$$\left|\hat{L}_t^{kc}(c,\nu) - t\ell(c,\nu)\right| \leq 9S^2\sqrt{8t\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}$$

where

$$\beta_t = R^2\left(d + 2\sqrt{d\ln\frac{3t(t+1)}{\delta}} + 2\ln\frac{3t(t+1)}{\delta}\right) + S^2\left(1 + 2\sqrt{\frac{\ln\frac{3t(t+1)}{\delta}}{d}}\right) = \widetilde{\mathcal{O}}(R^2 d + S^2).$$

Furthermore, under the same event, the regularized loss

$$\hat{L}_t^{kc,R}(c,\nu) = \hat{L}_t^{kc}(c,\nu) + \gamma_t\|\nu\|^2 \tag{1}$$

for $\gamma_t = 2\sqrt{8t\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}$ is quadratic and strictly convex in ν for all $t \geq 1$.

Proof. The loss estimator can be written as follows:

$$\begin{aligned}\hat{L}_t^{kc}(c,\nu) &= \sum_{s=1}^{t}\left(\nu^\top\left(\hat{x}_s\hat{x}_s^\top + \Sigma_n(c) - \Sigma_n(c_s)\right)\nu - 2\nu^\top\hat{x}_s y_s + y_s^2 + \lambda c\right)\\ &= \nu^\top\left(\sum_{s=1}^{t}\left(\hat{x}_s\hat{x}_s^\top + \Sigma_n(c) - \Sigma_n(c_s)\right)\right)\nu - 2\nu^\top\left(\sum_{s=1}^{t}\hat{x}_s x_s^\top\right)\theta^* + {\theta^*}^\top\left(\sum_{s=1}^{t}x_s^\top x_s\right)\theta^* + t\lambda c. \end{aligned}\tag{4}$$

Therefore, by Claim 1, for any cost c and parameter ν, we have

$$\begin{aligned}&\hat{L}_t^{kc}(c,\nu) - t\ell(c,\nu)\\ &\quad= \nu^\top\left(\sum_{s=1}^{t}\left(\hat{x}_s\hat{x}_s^\top - \Sigma_x - \Sigma_n(c_s)\right)\right)\nu - 2\nu^\top\left(\sum_{s=1}^{t}\hat{x}_s x_s^\top - \Sigma_x\right)\theta^* + {\theta^*}^\top\left(\sum_{s=1}^{t}x_s x_s^\top - \Sigma_x\right)\theta^*\\ &\quad= \nu^\top\left(\sum_{s=1}^{t}\left(\hat{x}_s\hat{x}_s^\top - \Sigma_x - \Sigma_n(c_s)\right)\right)\nu - 2\nu^\top\left(\sum_{s=1}^{t}x_s x_s^\top - \Sigma_x\right)\theta^* - 2\nu^\top\left(\sum_{s=1}^{t}n_s x_s^\top\right)\theta^* + {\theta^*}^\top\left(\sum_{s=1}^{t}x_s x_s^\top - \Sigma_x\right)\theta^*.\end{aligned}$$

All but a single term form outer products. The remaining mixed term can also be represented as

$$\nu^\top\left(\sum_{s=1}^{t}n_s x_s^\top\right)\theta^* = \begin{pmatrix}0^\top & \nu^\top\end{pmatrix}\sum_{s=1}^{t}\left(\begin{pmatrix}x_s\\ n_s\end{pmatrix}\begin{pmatrix}x_s\\ n_s\end{pmatrix}^\top\right)\begin{pmatrix}\theta^*\\ 0\end{pmatrix} = \begin{pmatrix}0^\top & \nu^\top\end{pmatrix}\sum_{s=1}^{t}\left(\begin{pmatrix}x_s\\ n_s\end{pmatrix}\begin{pmatrix}x_s\\ n_s\end{pmatrix}^\top - \begin{pmatrix}\Sigma_x & 0\\ 0 & \Sigma_n(c_s)\end{pmatrix}\right)\begin{pmatrix}\theta^*\\ 0\end{pmatrix},$$

and using the inequality $x^\top Ay \leq \|x\|\|y\|\|A\|_{op}$ for any $x, y \in \mathbb{R}^d$ and $A \in \mathbb{R}^{d\times d}$, alongside the bounds $\|\nu\|, \|\theta^*\| \leq S$, we get

$$\begin{aligned}\left|\hat{L}_t^{kc}(c,\nu) - t\ell(c,\nu)\right| \leq{}& S^2\left\|\sum_{s=1}^{t}\left(\hat{x}_s\hat{x}_s^\top - \Sigma_x - \Sigma_n(c_s)\right)\right\|_{op} + 3S^2\left\|\sum_{s=1}^{t}\left(x_s x_s^\top - \Sigma_x\right)\right\|_{op}\\ &+ 2S^2\left\|\sum_{s=1}^{t}\left(\begin{pmatrix}x_s\\ n_s\end{pmatrix}\begin{pmatrix}x_s\\ n_s\end{pmatrix}^\top - \begin{pmatrix}\Sigma_x & 0\\ 0 & \Sigma_n(c_s)\end{pmatrix}\right)\right\|_{op}.\end{aligned}$$

Now, notice that x_t and n_t are R^2 conditionally subgaussian and conditionally independent; hence, $\hat{x}_t = x_t + n_t$ is $2R^2$ conditionally subgaussian and $z_t = \begin{pmatrix}x_t\\ n_t\end{pmatrix}$ is R^2 conditionally subgaussian (w.r.t. the filtration $\mathcal{F}_t = \sigma(I_t)$), and both their expectations are of norm smaller than S. We can therefore apply Lemma 6 on each of the three norms (noticing that the conditional covariance of each term is as needed for the lemma). While doing so, we remark that replacing $R^2 \to 2R^2$ or $d \to 2d$

can increase β_t by at most a factor of 2. Following this application of Lemma 6, we get that w.p. at least $1-3\delta$, for all $t \geq 1$:

$$\left\|\sum_{s=1}^{t}\left(\hat{x}_s\hat{x}_s^\top - \Sigma_x - \Sigma_n(c_s)\right)\right\|_{op} < 2\sqrt{8t\beta_t^2 \ln\frac{3dt(t+1)}{\delta}},$$
$$\left\|\sum_{s=1}^{t}\left(x_s x_s^\top - \Sigma_x\right)\right\|_{op} < \sqrt{8t\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}, \qquad \text{and}$$
$$\left\|\sum_{s=1}^{t}\left(\begin{pmatrix}x_t\\ n_t\end{pmatrix}\begin{pmatrix}x_t\\ n_t\end{pmatrix}^\top - \begin{pmatrix}\Sigma_x & 0\\ 0 & \Sigma_n(c)\end{pmatrix}\right)\right\|_{op} < 2\sqrt{8t\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}.$$

In particular, for all $t \geq 1$, $c \in [0,1]$ and $\nu \in \mathbb{R}^d$ s.t. $\|\nu\| \leq S$, it holds that

$$\begin{aligned}\left|\hat{L}_t^{kc}(c,\nu) - t\ell(c,\nu)\right| &\leq 2S^2\sqrt{8t\beta_t^2 \ln\frac{3dt(t+1)}{\delta}} + 3S^2\sqrt{8t\beta_t^2 \ln\frac{3dt(t+1)}{\delta}} + 4S^2\sqrt{8t\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}\\ &= 9S^2\sqrt{8t\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}.\end{aligned}$$

Finally, we move our focus to the regularized loss $\hat{L}_t^{kc,R}(c,\nu)$, which similarly to eq. (4), can be written as

$$\hat{L}_t^{kc,R}(c,\nu) = \hat{L}_t^{kc}(c,\nu) + \gamma_t\|\nu\|^2 = \nu^\top\sum_{s=1}^{t}\left(\hat{x}_s\hat{x}_s^\top + \Sigma_n(c) - \Sigma_n(c_s) + \gamma_t I\right)\nu - 2\nu^\top\sum_{s=1}^{t}(\hat{x}_s y_s) + \sum_{s=1}^{t} y_s^2 + t\lambda c$$

The loss is clearly quadratic in ν and is strictly convex iff the matrix in the quadratic form is positive definite. Indeed, under the aforementioned event, we have

$$\begin{aligned}\lambda_{\min}\left(\sum_{s=1}^{t}\left(\hat{x}_s\hat{x}_s^\top + \Sigma_n(c) - \Sigma_n(c_s) + \gamma_t I\right)\right) &= \gamma_t + \lambda_{\min}\left(\sum_{s=1}^{t}\left(\hat{x}_s\hat{x}_s^\top - \Sigma_n(c_s) - \Sigma_x\right) + t(\Sigma_x + \Sigma_n(c))\right)\\ &\geq \gamma_t + \lambda_{\min}\left(\sum_{s=1}^{t}\left(\hat{x}_s\hat{x}_s^\top - \Sigma_n(c_s) - \Sigma_x\right)\right) \qquad (\Sigma_x, \Sigma_n(c) \succeq 0)\\ &\geq \gamma_t - \left\|\sum_{s=1}^{t}\left(\hat{x}_s\hat{x}_s^\top - \Sigma_n(c_s) - \Sigma_x\right)\right\|_{op}\\ &> 0.\end{aligned}$$

□

Theorem 1. *For any $T \geq 1$, set $\delta = 1/T$ and $K = \lceil \lambda T \rceil$. Then, the regret of Algorithm 1 is bounded by*

$$\mathrm{Reg}(T) = \widetilde{\mathcal{O}}\Big(S^2(R^2 d + S^2)\sqrt{T} + \lambda\Big).$$

Proof. Define the good event $\mathcal{G}$ under which all the results of Proposition 2 hold; by the proposition, we know that the probability of $\mathcal{G}$ is at least $1 - 3/T$. In particular, when $\mathcal{G}$ holds, for all $t \geq 1$, $c \in [0,1]$ and ν s.t. $\|\nu\| \leq S$, we have that

$$\left|\hat{L}_t^{kc}(c,\nu) - t\ell(c,\nu)\right| \leq 9S^2\sqrt{8t\beta_t^2 \ln \frac{3dt(t+1)}{\delta}}. \tag{5}$$

We also define the regularized expected loss by $\ell_t^R(c,\nu) = \ell(c,\nu) + \frac{1}{t}\gamma_t\|\nu\|^2$ and its minimizer $\nu_t^R(c) \in \arg\min_{\nu:\|\nu\|\leq S} \ell_t^R(c,\nu)$. Similarly, define the optimal cost on the grid by $k_t^R \in \arg\min_{k\in[K]} \ell_t^R(k/K, \nu_t^R(k/K))$. Then, under G, for all $t \geq 1$, we have

$$\begin{aligned}
\ell\left(\frac{k_t}{K}, \hat{\nu}_t(k_t)\right) &\leq \ell\left(\frac{k_t}{K}, \hat{\nu}_t(k_t)\right) + \frac{1}{t}\gamma_t\|\hat{\nu}_t(k_t)\|^2 && \text{(The regularization is nonnegative)} \\
&\leq \frac{1}{t}\left(\hat{L}_t^{kc}\left(\frac{k_t}{K}, \hat{\nu}_t(k_t)\right) + \gamma_t\|\hat{\nu}_t(k_t)\|^2\right) + 9S^2\sqrt{\frac{8\beta_t^2 \ln \frac{3dt(t+1)}{\delta}}{t}} && \text{(By eq. (5))} \\
&\overset{(1)}{=} \frac{1}{t}\min_{k\in[K],\nu:\|\nu\|\leq S}\left\{\hat{L}_t^{kc}\left(\frac{k}{K}, \nu\right) + \gamma_t\|\nu\|^2\right\} + 9S^2\sqrt{\frac{8\beta_t^2 \ln \frac{3dt(t+1)}{\delta}}{t}} \\
&\leq \frac{1}{t}\left(\hat{L}_t^{kc}\left(\frac{k_t^R}{K}, \nu_t^R\left(\frac{k_t^R}{K}\right)\right) + \gamma_t\left\|\nu_t^R\left(\frac{k_t^R}{K}\right)\right\|^2\right) + 9S^2\sqrt{\frac{8\beta_t^2 \ln \frac{3dt(t+1)}{\delta}}{t}} \\
&\leq \left(\ell\left(\frac{k_t^R}{K}, \nu_t^R\left(\frac{k_t^R}{K}\right)\right) + \frac{\gamma_t}{t}\left\|\nu_t^R\left(\frac{k_t^R}{K}\right)\right\|^2\right) + 18S^2\sqrt{\frac{8\beta_t^2 \ln \frac{3dt(t+1)}{\delta}}{t}} && \text{(By eq. (5))} \\
&\overset{(2)}{=} \min_{k\in[K],\nu:\|\nu\|\leq S}\left\{\ell\left(\frac{k}{K}, \nu\right) + \frac{\gamma_t}{t}\|\nu\|^2\right\} + 18S^2\sqrt{\frac{8\beta_t^2 \ln \frac{3dt(t+1)}{\delta}}{t}} \\
&\leq \min_{k\in[K],\nu:\|\nu\|\leq S}\left\{\ell\left(\frac{k}{K}, \nu\right)\right\} + \frac{S^2\gamma_t}{t} + 18S^2\sqrt{\frac{8\beta_t^2 \ln \frac{3dt(t+1)}{\delta}}{t}} && \left(\left\|\nu_t^R\left(\tfrac{k_t^R}{K}\right)\right\| \leq S\right) \\
&\overset{(3)}{=} \min_{k\in[K]}\left\{\ell^*\left(\frac{k}{K}\right)\right\} + 20S^2\sqrt{\frac{8\beta_t^2 \ln \frac{3dt(t+1)}{\delta}}{t}},
\end{aligned}$$

where relation (1) is by the optimality of k_t and $\hat{\nu}_t$ w.r.t. the empirical regularized loss and (2) is by the optimality of k_t^R and $\nu_t^R\left(\frac{k_t^R}{K}\right)$ w.r.t. the true regularized loss. Relation (3) substitutes γ_t and ℓ^*.

We continue by analyzing the relation between the optimal loss and the loss over the grid:

$$\ell^* = \min_{c\in[0,1]} \ell^*(c) = \min_{k\in[K]} \min_{c\in\left[\frac{k-1}{K}, \frac{k}{K}\right)} \ell^*(c) \overset{(*)}{\geq} \min_{k\in[K]} \min_{c\in\left[\frac{k-1}{K}, \frac{k}{K}\right)} \left\{\ell^*\left(\frac{k}{K}\right) - \lambda\left(\frac{k}{K} - c\right)\right\} = \min_{k\in[K]} \ell^*\left(\frac{k}{K}\right) - \frac{\lambda}{K}, \tag{6}$$

where relation $(*)$ holds due to the one-sided Lipschitzness on loss (see Claim 2). Combining both inequalities, we get

$$\ell\left(\frac{k_t}{K}, \hat{\nu}_t(k_t)\right) \leq \ell^* + 20S^2\sqrt{\frac{8\beta_t^2 \ln \frac{3dt(t+1)}{\delta}}{t}} + \frac{\lambda}{K}.$$

We are now ready to bound the regret, using the regret decomposition of Lemma 1:

$$\begin{aligned}
\mathrm{Reg}(T) &= \mathbb{E}\left[\sum_{t=1}^{T}(\ell(c_t,\nu_t)-\ell^*)\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T}\left(\ell\left(\frac{k_{t-1}}{K},\hat{\nu}_{t-1}(k_{t-1})\right)-\ell^*\right)\right] \\
&\le \mathbb{E}\left[\sum_{t=2}^{T}\left(\ell\left(\frac{k_{t-1}}{K},\hat{\nu}_{t-1}(k_{t-1})\right)-\ell^*\right)\right] + \max_{c\in[0,1],\nu:\|\nu\|\le S}\ell(c,\nu) && \text{(separating } T=1\text{)} \\
&\le \mathbb{E}\left[\mathbb{1}\{\mathcal{G}\}\sum_{t=2}^{T}\left(\ell\left(\frac{k_{t-1}}{K},\hat{\nu}_{t-1}(k_{t-1})\right)-\ell^*\right)\right] + \underbrace{\Pr\{\bar{\mathcal{G}}\}T}_{\le 3}\max_{c\in[0,1],\nu:\|\nu\|\le S}\ell(c,\nu) + \max_{c\in[0,1],\nu:\|\nu\|\le S}\ell(c,\nu) \\
&\le \sum_{t=1}^{T-1}\left(20S^2\sqrt{\frac{8\beta_t^2\ln\frac{3dt(t+1)}{\delta}}{t}}+\frac{\lambda}{K}\right)+4\max_{c\in[0,1],\nu:\|\nu\|\le S}\ell(c,\nu) \\
&\le 40S^2\sqrt{8T\beta_T^2\ln(3dT^2(T+1))}+1+4\max_{c\in[0,1],\nu:\|\nu\|\le S}\ell(c,\nu) && (K\ge\lambda T,\delta=1/T) \\
&\le 40S^2\sqrt{8T\beta_T^2\ln(3dT^2(T+1))}+1+24S^2\big(R^2d+S^2\big)+4\lambda && \text{(Claim 5)} \\
&= \widetilde{\mathcal{O}}\Big(S^2(R^2d+S^2)\sqrt{T}+\lambda\Big).
\end{aligned}$$

□

B.1 Lower Bounds for Known Noise Covariance

We consider the following two one-dimensional instances. In the first instance p^-, it holds that $x_t \sim \mathcal{N}(0, 1-\epsilon)$, while in the second instance p^+, we have $x_t \sim \mathcal{N}(0, 1+\epsilon)$, where $\epsilon \in (0, 1/2]$ will be determined later. Aside for this difference, both instances are completely identical; in both, $\theta^* = 1$ and $\lambda = 1$, and for any given cost c, the feature noise is distributed as $\mathcal{N}(0, \sigma_n^2(c))$ for $\sigma_n^2(c) = \mathbb{1}\{c < 1/2\}$. In other words, the noise has unit variance for all costs smaller than $1/2$, and from this payment onward, the noise variance is zero – there is no noise. For this choice of θ^*, we get $y_t = x_t$, and therefore, we can assume w.l.o.g. that at the end of each round, we observe x_t and n_t (so I_t and I_t^o contain the same information). By Claim 1 and Claim 3, for feature covariance σ_x^2 and noise covariance $\sigma_n^2(c)$ (and recalling that $\sigma_z^2 = \sigma_x^2 + \sigma_n^2$), we get an optimal loss

$$\ell^*(c) = \sigma_x^2 - \frac{\sigma_x^4}{\sigma_x^2+\sigma_n^2(c)} + \lambda c = \frac{\sigma_x^2\sigma_n^2(c)}{\sigma_x^2+\sigma_n^2(c)} + \lambda c.$$

Notably, for our choice of noise, the loss is strictly increasing inside the intervals $c \in [0, 1/2)$ and $c \in [1/2, 1]$ since in these intervals, the noise variance remains the same while the payment increases. Thus, it is always optimal to either play $c = 0$ or $c = 1/2$. Now, looking at these two specific payments, we have

$$\begin{aligned}
\ell_+^*(0) &= \frac{1+\epsilon}{1+\epsilon+1} = \frac{1+\epsilon}{2+\epsilon} > 1/2, \qquad \text{and,} \qquad \ell_+^*(1/2) = 0+\frac{1}{2} = \frac{1}{2}, \\
\ell_-^*(0) &= \frac{1-\epsilon}{1-\epsilon+1} = \frac{1-\epsilon}{2-\epsilon} < 1/2, \qquad \text{and,} \qquad \ell_-^*(1/2) = 0+\frac{1}{2} = \frac{1}{2}.
\end{aligned}$$

Therefore, in instance p^+, an optimal algorithm will play $c_+^* = 1/2$, and in instance p^-, an optimal algorithm will play $c_-^* = 0$. The suboptimality of playing the wrong choice is

$$\begin{aligned}
\Delta_+ &= \ell_+^*(0) - \ell_+^*(1/2) = \frac{1+\epsilon}{2+\epsilon} - \frac{1}{2} = \frac{\epsilon}{2(2+\epsilon)} \ge \frac{\epsilon}{5}, \\
\Delta_- &= \ell_-^*(1/2) - \ell_-^*(0) = \frac{1}{2} - \frac{1-\epsilon}{2-\epsilon} = \frac{\epsilon}{2(2-\epsilon)} \ge \frac{\epsilon}{4} \ge \frac{\epsilon}{5},
\end{aligned}$$

where in both cases, the inequalities use the fact that $\epsilon \in (0, 1/2]$.

Using these properties allows us to prove a lower bound for the regret:

Theorem 2. *For any algorithm and $T \ge 1$, there exists a one-dimensional instance with a known noise variance profile such that $\mathbb{E}[R(T)] \ge \frac{\sqrt{T}}{240}$.*

Proof. We adapt the technique of Theorem 15.2 in (Lattimore and Szepesvári 2020). Let $\mathbb{P}_{p^+}$ and $\mathbb{P}_{p^-}$ be the distribution of the entire process of the algorithm when applied to instances p^+ and p^-, respectively. Since $\theta^* = 1$, we effectively observe x_t, n_t at the end of each round, so the observed history I_t^o and the full one I_t contain the same information. We also remark that $\hat{y}_t$ is completely determined by the algorithm and does not affect the observation. Indeed, all information gained between I_t and I_{t+1} depends only on c_{t+1} and any internal stochasticity of the algorithm.

By the chain rule of KL divergence, we have (see. e.g., (9) in (Garivier, Ménard, and Stoltz 2019)

$$\mathrm{KL}\Big(\mathbb{P}_{p^+}^{I_T}, \mathbb{P}_{p^-}^{I_T}\Big) = \sum_{t=1}^{T} \mathrm{KL}\Big(\mathbb{P}_{p^+}^{\hat{x}_t,\hat{y}_t,y_t,x_t,n_t,c_{t+1}|I_{t-1}}, \mathbb{P}_{p^-}^{\hat{x}_t,\hat{y}_t,y_t,x_t,n_t,c_{t+1}|I_{t-1}}\Big).$$

Each term can be greatly simplified to

$$\begin{aligned}
&\mathrm{KL}\Big(\mathbb{P}_{p^+}^{\hat{x}_t,\hat{y}_t,y_t,x_t,n_t,c_{t+1}|I_{t-1}}, \mathbb{P}_{p^-}^{\hat{x}_t,\hat{y}_t,y_t,x_t,n_t,c_{t+1}|I_{t-1}}\Big) \\
&\quad = \mathrm{KL}\Big(\mathbb{P}_{p^+}^{x_t,n_t|I_{t-1}}, \mathbb{P}_{p^-}^{x_t,n_t|I_{t-1}}\Big) + \underbrace{\mathrm{KL}\Big(\mathbb{P}_{p^+}^{\hat{x}_t,\hat{y}_t,y_t,c_{t+1}|I_{t-1},x_t,n_t}, \mathbb{P}_{p^-}^{\hat{x}_t,\hat{y}_t,y_t,c_{t+1}|I_{t-1},x_t,n_t}\Big)}_{=0} \\
&\quad \overset{(1)}{=} \mathrm{KL}\Big(\mathbb{P}_{p^+}^{x_t,n_t|I_{t-1}}, \mathbb{P}_{p^-}^{x_t,n_t|I_{t-1}}\Big) \\
&\quad \overset{(2)}{=} \mathrm{KL}\Big(\mathbb{P}_{p^+}^{x_t|I_{t-1}}, \mathbb{P}_{p^-}^{x_t|I_{t-1}}\Big) \\
&\quad \overset{(3)}{=} \mathrm{KL}(\mathcal{N}(0,1+\epsilon), \mathcal{N}(0,1-\epsilon)).
\end{aligned}$$

In this derivation, relation (1) holds since $\hat{x}_t, y_t$ are deterministic given x_t, n_t and that $\hat{y}_t$ and c_{t+1} are determined identically by the algorithm in both instances – given the same history (and the same x_t, n_t), they have the same distribution and the KL divergence equals zero. Relation (2) holds since n_t and x_t are independent given I_{t-1} and n_t has the same distribution in both instances (as I_{t-1} already encapsulates the choice of c_t). Finally, (3) substitutes the conditional feature distribution. The resulting KL divergence can be directly bounded by

$$\mathrm{KL}(\mathcal{N}(0,1+\epsilon), \mathcal{N}(0,1-\epsilon)) = \frac{1}{2}\left(\frac{1+\epsilon}{1-\epsilon} - \ln\frac{1+\epsilon}{1-\epsilon} - 1\right) \overset{(*)}{\leq} \frac{1}{4}\left(\frac{1+\epsilon}{1-\epsilon} - 1\right)^2 = \frac{1}{4}\left(\frac{2\epsilon}{1-\epsilon}\right)^2 \leq 4\epsilon^2,$$

where $(*)$ use the inequality $\ln(x) \geq x - 1 - (x-1)^2/2$, which holds for all $x \geq 1$, and the last inequality is since $\epsilon \leq 1/2$. Thus, $\mathrm{KL}\Big(\mathbb{P}_{p^+}^{I_T}, \mathbb{P}_{p^-}^{I_T}\Big) \leq 4T\epsilon^2$.

We now use this result to lower-bound the regret. In particular, define the random variable $N_T = \sum_{t=1}^{T} \mathbb{1}\{c_t \in [1/2, 1]\}$ (which is completely determined by the information in I_T) and fix $\epsilon = \frac{1}{2\sqrt{T}} \in (0, 1/2]$. Then, by the Bretagnolle–Huber inequality (Theorem 14.2 of Lattimore and Szepesvári 2020), it holds that

$$\max\big\{\mathbb{P}_{p^-}(N_T > T/2), \mathbb{P}_{p^+}(N_T \leq T/2)\big\} \geq \frac{1}{4}\exp\Big(-\mathrm{KL}\Big(\mathbb{P}_{p^+}^{I_T}, \mathbb{P}_{p^-}^{I_T}\Big)\Big) \geq \frac{1}{4}\exp\left(-4T\left(\frac{1}{2\sqrt{T}}\right)^2\right) \geq \frac{1}{12}. \tag{7}$$

Since all distributions are zero-mean Gaussian, these probabilities can be related to the regret in both instances by Claim 4 (and Claim 3), namely,

$$\begin{aligned}
\mathrm{Reg}_{p^-}(T) &\geq \mathbb{E}\left[\sum_{t=1}^{T}\big(\ell_-^*(c_t) - \ell_-^*(0)\big)\right] \\
&\geq \mathbb{E}\left[\sum_{t=1}^{T}\big(\ell_-^*(c_t) - \ell_-^*(0)\big)\mathbb{1}\{c_t \in [1/2,1]\}\right] \\
&\geq \mathbb{E}\left[\sum_{t=1}^{T}\big(\ell_-^*(1/2) - \ell_-^*(0)\big)\mathbb{1}\{c_t \in [1/2,1]\}\right] && \text{(The loss increases in } [1/2,1]) \\
&= \Delta_- \mathbb{E}\left[\sum_{t=1}^{T}\mathbb{1}\{c_t \in [1/2,1]\}\right] \\
&\geq \Delta_- \cdot \frac{T}{2}\mathbb{P}_{p^-}(N_T > T/2).
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
\mathrm{Reg}_{p^+}(T) &\geq \mathbb{E}\left[\sum_{t=1}^{T}\big(\ell_+^*(c_t) - \ell_+^*(1/2)\big)\right] \\
&\geq \mathbb{E}\left[\sum_{t=1}^{T}\big(\ell_+^*(c_t) - \ell_+^*(1/2)\big)\mathbb{1}\{c_t \in [0,1/2)\}\right] \\
&\geq \mathbb{E}\left[\sum_{t=1}^{T}\big(\ell_+^*(0) - \ell_+^*(1/2)\big)\mathbb{1}\{c_t \in [0,1/2)\}\right] && \text{(The loss increases in } [0,1/2)) \\
&= \Delta_+ \mathbb{E}\left[\underbrace{\sum_{t=1}^{T}\mathbb{1}\{c_t \in [0,1/2)\}}_{=T-N_T}\right] \\
&\geq \Delta_+ \cdot \frac{T}{2}\mathbb{P}_{p^-}(T - N_T \geq T/2) \\
&= \Delta_+ \cdot \frac{T}{2}\mathbb{P}_{p^-}(N_T \leq T/2).
\end{aligned}$$

Combining both with Equation (7) and lower bounding both gaps by $\epsilon/5$, we get

$$\max\big\{\mathrm{Reg}_{p^-}(T), \mathrm{Reg}_{p^+}(T)\big\} \geq \frac{\epsilon}{5} \cdot \frac{T}{2} \cdot \max\big\{\mathbb{P}_{p^-}(N_T > T/2), \mathbb{P}_{p^+}(N_T \leq T/2)\big\} \geq \frac{\epsilon T}{10} \cdot \frac{1}{12} = \frac{\sqrt{T}}{240}.$$

□

B.2 Regression with Paid Feature and Additive Noise

In this appendix, we shortly discuss how the results change if the regression output is noisy, that is, when observing $\widetilde{y}_t = x_t^\top \theta^* + \eta_t = y_t + \eta_t$ for some zero-mean noise η_t that is R^2-subgaussian, conditionally independent of all other quantities at round t and is of variance σ_η^2. While we describe the modification only for the case where the noise covariances are known, similar conclusions can be drawn for the other problem variants studied in the paper.

Due to the independence and the fact the the noise is of expectation zero, the instantaneous expected loss can be written as

$$\begin{aligned}
\widetilde{\ell}(c,\nu) &= \mathbb{E}_{x\sim D_x, n\sim D_n(c),\eta}\left[\Big((x+n)^\top\nu - x^\top\theta^* - \eta\Big)^2\right] + \lambda c \\
&= \mathbb{E}_{x\sim D_x, n\sim D_n(c)}\left[\Big((x+n)^\top\nu - x^\top\theta^*\Big)^2\right] + \mathbb{E}_\eta\big[\eta^2\big] + \lambda c && \text{(independence of } \eta) \\
&= \ell(c,\nu) + \sigma_\eta^2.
\end{aligned}$$

That is, the loss is only shifted by σ_η^2, so that the optimal values $\nu^*(c)$ and c^* remain unchanged, and the optimal loss is also biased by the same quantity.

To show that our algorithmic approach still applies, we need to show that the loss estimation, as described in Proposition 2, behaves similarly. The new loss estimator that uses the noisy regression outputs can be written as

$$\begin{aligned}
\widetilde{L}_t^{kc}(c,\nu) &= \sum_{s=1}^{t}\Big(\big(\hat{x}_s^\top\nu - \widetilde{y}_s\big)^2 + \nu^\top(\Sigma_n(c) - \Sigma_n(c_s))\nu + \lambda c\Big) \\
&= \sum_{s=1}^{t}\Big(\big(\hat{x}_s^\top\nu - y_s - \eta_s\big)^2 + \nu^\top(\Sigma_n(c) - \Sigma_n(c_s))\nu + \lambda c\Big) \\
&= \sum_{s=1}^{t}\Big(\big(\hat{x}_s^\top\nu - y_s\big)^2 + \nu^\top(\Sigma_n(c) - \Sigma_n(c_s))\nu + \lambda c\Big) - 2\sum_{s=1}^{t}\eta_s\big(\hat{x}_s^\top\nu - y_s\big) + \sum_{s=1}^{t}\eta_s^2 \\
&= \hat{L}_t^{kc}(c,\nu) - 2\sum_{s=1}^{t}\eta_s x_s^\top(\nu - \theta^*) - 2\sum_{s=1}^{t}\eta_s n_s^\top\nu + \sum_{s=1}^{t}\eta_s^2.
\end{aligned}$$

In Proposition 2, we already showed that $\hat{L}_t^{kc}(c,\nu)$ converges to the expected loss $t\ell(c,\nu)$. Moreover, the last term does not depend on c or ν, not affecting the minimizers of the loss. Thus, to maintain the same results, one only needs to prove that the two middle terms converge to zero at a comparable rate as the empirical loss to the expected loss. This naturally follows from the fact that the noise is conditionally zero-mean subgaussian given x_t, n_t and the high-probability boundedness of x_t and n_t. Alternatively, one can readily extend the analysis technique used in Proposition 2 and write, for example,

$$\sum_{s=1}^{t} \eta_s x_s^\top (\nu - \theta^*) = \begin{pmatrix} 0^\top & 1 \end{pmatrix} \sum_{s=1}^{t} \left(\begin{pmatrix} x_s \\ \eta_s \end{pmatrix} \begin{pmatrix} x_s \\ \eta_s \end{pmatrix}^\top \right) \begin{pmatrix} \nu - \theta^* \\ 0 \end{pmatrix} = \begin{pmatrix} 0^\top & 1 \end{pmatrix} \sum_{s=1}^{t} \left(\begin{pmatrix} x_s \\ \eta_s \end{pmatrix} \begin{pmatrix} x_s \\ \eta_s \end{pmatrix}^\top - \begin{pmatrix} \Sigma_x & 0 \\ 0 & \sigma_\eta^2 \end{pmatrix} \right) \begin{pmatrix} \nu - \theta^* \\ 0 \end{pmatrix}$$

$$\Rightarrow \left| \sum_{s=1}^{t} \eta_s x_s^\top (\nu - \theta^*) \right| \leq 2S \left\| \sum_{s=1}^{t} \left(\begin{pmatrix} x_s \\ \eta_s \end{pmatrix} \begin{pmatrix} x_s \\ \eta_s \end{pmatrix}^\top - \begin{pmatrix} \Sigma_x & 0 \\ 0 & \sigma_\eta^2 \end{pmatrix} \right) \right\|_{op},$$

where the operator norm can be bounded by Lemma 6, leading to similar rates.

To summarize, one can easily extend the concentration bounds in Proposition 2 to the case of noisy outputs, with only mild modifications to the constants. Since this is the concentration property that we use for both regret proofs (either directly or via Corollary 3), the same constant change will propagate to the optimistic index and regret bounds, leading to very minor algorithmic and bound changes.

C Proofs for Unknown Noise Covariances

Corollary 3. *Fix $c \in [0,1]$ and let $N_t(c) = \sum_{s=1}^{t} \mathbb{1}\{c_s = c\}$. Then, with probability at least $1-3\delta$, for all $t \geq 1$ and $\nu \in \mathbb{R}^d$ s.t. $\|\nu\| \leq S$, it holds that*

$$\left|\hat{L}_t^{uc}(c,\nu) - N_t(c)\ell(c,\nu)\right| \leq 9S^2\sqrt{8N_t(c)\beta_t^2 \ln\frac{3dt(t+1)}{\delta}},$$

where $\beta_t = \widetilde{\mathcal{O}}(R^2 d + S^2)$ is specified at Proposition 2.

Proof. Before the interaction starts, we sample T i.i.d. samples $\{z_t\}_{t\in[T]}$ and $\{\eta_t\}_{t\in[T]}$ from the distributions D_x and $D_n(c)$. Then, without loss of generality, we assume that upon choosing $c_t = c$ for the i^{th} time, the feature vector is taken as $x_t = z_i$ and the noise vector is given by $n_t = \eta_i$, so that the regression output is $y_t = x_t^\top \theta^* = z_i^\top \theta^* \triangleq \xi_i$. Then, defining the loss

$$\widetilde{L}_i(c,\nu) = \sum_{s=1}^{i}\bigg(\Big((z_s+\eta_s)^\top \nu - \xi_i\Big)^2 + \lambda c\bigg),$$

we have that $\hat{L}_t^{uc}(c,\nu) = \widetilde{L}_{N_t(c)}(c,\nu)$.

Now, by Proposition 2, we have that for all $i \geq 1$ and $\nu \in \mathbb{R}^d$ s.t. $\|\nu\| \leq S$, it holds w.p. $1-3\delta$ that

$$\left|\widetilde{L}_i(c,\nu) - i\ell(c,\nu)\right| \leq 9S^2\sqrt{8i\beta_i^2 \ln\frac{3di(i+1)}{\delta}}.$$

Since the event holds simultaneously for all $i \geq 1$, it also holds for $i = N_t(c)$ for all $t \geq 1$. In particular, w.p. $1-3\delta$, for all $t \geq 1$ and $\nu \in \mathbb{R}^d$ s.t. $\|\nu\| \leq S$,

$$\left|\hat{L}_t^{uc}(c,\nu) - N_t(c)\ell(c,\nu)\right| \leq 9S^2\sqrt{8N_t(c)\beta_{N_t(c)}^2 \ln\frac{3dN_t(c)(N_t(c)+1)}{\delta}} \leq 9S^2\sqrt{8N_t(c)\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}.$$

where we used the fact that $N_t(c) \leq t$ and the monotonicity of β_t in t. □

Theorem 3. *For any $T \geq 1$, set $K = \left\lceil \frac{T^{1/3}\lambda^{2/3}}{(S^2(R^2d+S^2))^{2/3}} \right\rceil$ and $\delta = \frac{1}{KT}$. Then, the regret of Algorithm 2 is bounded by*

$$\mathrm{Reg}(T) = \widetilde{\mathcal{O}}\Big(\big(S^2(R^2d+S^2)\big)^{2/3}\lambda^{1/3}T^{2/3}\Big).$$

Proof. Define the good event $\mathcal{G}_k$ under which all the results of Corollary 3 hold for $c = k/K$ and let $\mathcal{G} = \bigcap_{k\in[K]}\mathcal{G}_k$. By the corollary, for $\delta = \frac{1}{KT}$, $\Pr\{\mathcal{G}\} \geq 1 - 3/T$, and under this event, for all $t \geq 1$, $c \in \{k/K\}_{k\in[K]}$ and ν s.t. $\|\nu\| \leq S$,

$$\left|\frac{1}{N_t(c)}\hat{L}_t^{uc}(c,\nu) - \ell(c,\nu)\right| \leq 9S^2\sqrt{\frac{8\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}{N_t(c)}}. \tag{8}$$

As a first step, we relate the empirical loss at the selected cost and linear regressor to the real loss when $\mathbb{G}$ holds.

$$\begin{aligned}
\hat{L}_t^{uc}\left(\frac{k_t}{K},\hat{\nu}_t(k_t)\right) &= \left(\hat{L}_t^{uc}\left(\frac{k_t}{K},\hat{\nu}_t(k_t)\right) - 9S^2\sqrt{8N_t\left(\frac{k_t}{K}\right)\beta_t^2\ln\frac{3dt(t+1)}{\delta}}\right) + 9S^2\sqrt{8N_t\left(\frac{k_t}{K}\right)\beta_t^2\ln\frac{3dt(t+1)}{\delta}} \\
&= N_t\left(\frac{k_t}{K}\right)\min_{k\in[K],\nu:\|\nu\|\leq S}\left\{\frac{\hat{L}_t^{uc}(k/K,\nu)}{N_t(k/K)} - 9S^2\sqrt{\frac{8\beta_t^2\ln\frac{3dt(t+1)}{\delta}}{N_t(k/K)}}\right\} + 9S^2\sqrt{8N_t\left(\frac{k_t}{K}\right)\beta_t^2\ln\frac{3dt(t+1)}{\delta}}.
\end{aligned}$$

In particular, denoting $k^* \in \arg\min_{k\in[K]}\ell^*(k/K)$, we can bound

$$\begin{aligned}
\hat{L}_t^{uc}\left(\frac{k_t}{K},\hat{\nu}_t(k_t)\right) &\leq N_t\left(\frac{k_t}{K}\right)\left(\frac{\hat{L}_t^{uc}(k^*/K,\nu^*(k^*/K))}{N_t(k^*/K)} - 9S^2\sqrt{\frac{8\beta_t^2\ln\frac{3dt(t+1)}{\delta}}{N_t(k^*/K)}}\right) + 9S^2\sqrt{8N_t\left(\frac{k_t}{K}\right)\beta_t^2\ln\frac{3dt(t+1)}{\delta}} \\
&\leq N_t\left(\frac{k_t}{K}\right)\ell\left(\frac{k^*}{K},\nu^*\left(\frac{k^*}{K}\right)\right) + 9S^2\sqrt{8N_t\left(\frac{k_t}{K}\right)\beta_t^2\ln\frac{3dt(t+1)}{\delta}} \qquad \text{(By eq. (8))} \\
&= N_t\left(\frac{k_t}{K}\right)\ell^*\left(\frac{k^*}{K}\right) + 9S^2\sqrt{8N_t\left(\frac{k_t}{K}\right)\beta_t^2\ln\frac{3dt(t+1)}{\delta}}. \qquad (9)
\end{aligned}$$

Combined with eq. (8), we can leverage this inequality to relate the expected loss at the played cost/prediction to the optimal expected loss, namely

$$\begin{aligned}
\ell\left(\frac{k_t}{K}, \hat{\nu}_t(k_t)\right) &\leq \frac{\hat{L}_t^{uc}(k_t/K, \hat{\nu}_t(k_t))}{N_t(k/K)} + 9S^2\sqrt{\frac{8\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}{N_t(k_t/K)}} && \text{(By eq. (8))} \\
&\leq \ell^*\left(\frac{k^*}{K}\right) + 18S^2\sqrt{\frac{8\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}{N_t(k_t/K)}} && \text{(By eq. (9))} \\
&\leq \ell^* + \frac{\lambda}{K} + 18S^2\sqrt{\frac{8\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}{N_t(k_t/K)}},
\end{aligned}$$

where the last transition is by eq. (6).

We are finally ready to bound the regret, using the regret decomposition of Lemma 1:

$$\begin{aligned}
\mathrm{Reg}(T) &= \mathbb{E}\left[\sum_{t=1}^{T}(\ell(c_t, \nu_t) - \ell^*)\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T}\left(\ell\left(\frac{k_{t-1}}{K}, \hat{\nu}_{t-1}(k_{t-1})\right) - \ell^*\right)\right] \\
&\leq \mathbb{E}\left[\sum_{t=K+1}^{T}\left(\ell\left(\frac{k_{t-1}}{K}, \hat{\nu}_{t-1}(k_{t-1})\right) - \ell^*\right)\right] + K\max_{c\in[0,1],\nu:\|\nu\|\leq S}\ell(c,\nu) && \text{(separating } T \leq K\text{)} \\
&\leq \mathbb{E}\left[\mathbb{1}\{\mathcal{G}\}\sum_{t=K+1}^{T}\left(\ell\left(\frac{k_{t-1}}{K}, \hat{\nu}_{t-1}(k_{t-1})\right) - \ell^*\right)\right] + \underbrace{\Pr\{\bar{\mathcal{G}}\}T}_{\leq 3}\max_{c\in[0,1],\nu:\|\nu\|\leq S}\ell(c,\nu) + K\max_{c\in[0,1],\nu:\|\nu\|\leq S}\ell(c,\nu) \\
&\leq \sum_{t=K+1}^{T}\left(18S^2\sqrt{\frac{8\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}{N_{t-1}(k_{t-1}/K)}} + \frac{\lambda}{K}\right) + 4K\max_{c\in[0,1],\nu:\|\nu\|\leq S}\ell(c,\nu).
\end{aligned}$$

The sum over the first term can be bounded using standard bandit arguments – noticing that the counts start from 1 due to the initial sampling stage and increase by one every time we play $k_t = k$, we have

$$\begin{aligned}
\sum_{t=K+1}^{T}\frac{1}{\sqrt{N_{t-1}(k_{t-1}/K)}} = \sum_{k\in[K]}\sum_{t=K+1}^{T}\frac{\mathbb{1}\{k_{t-1}=k\}}{\sqrt{N_{t-1}(k/K)}} &\leq \sum_{k\in[K]}\sum_{i=1}^{N_T(k/K)}\frac{1}{\sqrt{i}} \leq 2\sum_{k\in[K]}\sqrt{N_T(k/K)} \\
&\overset{(*)}{\leq} 2\sqrt{K}\sqrt{\sum_{k\in[K]}N_T(k/K)} = 2\sqrt{KT},
\end{aligned}$$

where inequality $(*)$ is by Cauchy-Schwarz. Substituting back while noting that β_t increases with t yields

$$\begin{aligned}
\mathrm{Reg}(T) &\leq 36S^2\sqrt{8KT\beta_T^2 \ln\frac{3dT(T+1)}{\delta}} + \frac{\lambda T}{K} + 4K\max_{c\in[0,1],\nu:\|\nu\|\leq S}\ell(c,\nu) \\
&\leq 36S^2\sqrt{8KT\beta_T^2 \ln(3dKT^2(T+1))} + \frac{\lambda T}{K} + 4K\left(6S^2\left(R^2d + S^2\right) + \lambda\right) && \text{(Claim 5)} \\
&= \widetilde{\mathcal{O}}\left(\left(S^2(R^2d+S^2)\right)^{2/3}\lambda^{1/3}T^{2/3}\right),
\end{aligned}$$

where the last inequality is due to the specific choice of K and the fact that $\beta_T = \widetilde{\mathcal{O}}\left(R^2d + S^2\right)$. □

C.1 Lower Bounds for Unknown Noise Covariance

We consider the following one-dimensional example. Let $x_t \sim \mathcal{N}(0,1)$, $\theta^* = 1$ and $\lambda = \frac{1}{2}$. For any given cost c, the feature noise is distributed as $\mathcal{N}(0, \sigma_n^2(c))$; we will briefly specify $\sigma_n^2(c)$. Notably, $y_t = x_t$, and therefore, we can assume w.l.o.g. that at the end of each round, we observe x_t and n_t. By Claim 1 and Claim 3, we have

$$\ell^*(c) = 1 - \frac{1}{1+\sigma_n^2(c)} + \frac{1}{2}c = \frac{\sigma_n^2(c)}{1+\sigma_n^2(c)} + \frac{1}{2}c.$$

Finally, we denote $f(c) = \frac{1-c}{1+c}$. Notice that if $\sigma_n^2(c) = f(c)$, it holds that

$$\ell^*(c) = \frac{\frac{1-c}{1+c}}{1+\frac{1-c}{1+c}} + \frac{1}{2}c = \frac{1-c}{2} + \frac{1}{2}c = \frac{1}{2}. \tag{10}$$

Let $K \in \mathbb{N}$ and consider the following instances:

1. The baseline instance p is such that $\sigma_n^2(c) = f(c)$, and so $\ell^*(c) = \frac{1}{2}$.
2. For any $k \in [K]$, define $c_k = \frac{1}{2} + \frac{k-1}{4K}$ (with $c_{K+1} = \frac{3}{4}$) . We then define the instance p_k with the noise variance

$$\sigma_n^2(c;k) = \begin{cases} f(c_k) + 2\frac{c-c_k}{c_{k+1}-c_k}(f(c_{k+1}) - f(c_k)) & c \in \left[c_k, \frac{c_k+c_{k+1}}{2}\right) \\ f(c_{k+1}) & c \in \left[\frac{c_k+c_{k+1}}{2}, c_{k+1}\right) \\ f(c) & \text{else} \end{cases}$$

Since $f(c)$ is 1-Lipschitz and decreasing, with $f(c) \in \left[\frac{1}{8}, \frac{1}{2}\right]$ for all $c \in \left[\frac{1}{2}, \frac{3}{4}\right]$, then $\sigma_n^2(c;k)$ is both decreasing and 1-Lipschitz.

This choice enjoys the following properties (which we formally prove after proving the lower bound theorem):

Claim 6. *1. For any c, k, the KL divergence between the two noise distributions is bounded by*

$$\mathrm{KL}\big(\mathcal{N}(0,\sigma_n^2(c)), \mathcal{N}(0,\sigma_n^2(c;k))\big) \le 16(c_{k+1}-c_k)^2 \mathbb{1}\{c \in [c_k, c_{k+1})]\} = \frac{1}{K^2}\mathbb{1}\{c \in [c_k, c_{k+1})]\}.$$

2. The optimal loss in instance p_k is bounded by

$$\min_{c\in[0,1]} \ell^*(c;k) \le \frac{1}{2} - \frac{c_{k+1}-c_k}{4} = \frac{1}{2} - \frac{1}{16K}.$$

Using these properties allows us to prove a lower bound for the regret:

Theorem 4. *For any algorithm and $T \ge 1$, there exists a one-dimensional instance with an unknown noise variance profile such that $\mathbb{E}[R(T)] \ge \frac{T^{2/3}}{256}$.*

Proof. We adapt the technique of Theorem 6 in (Garivier, Ménard, and Stoltz 2019). Fix some $K \in \mathbb{N}$ that will be determined later. Let $\mathbb{P}_p$ and $\mathbb{P}_{p_k}$ be the distribution of the entire process of the algorithm when applied to instances p and p_k, respectively. As in Theorem 2, notice that the choice $\theta^* = 1$ renders x_t, n_t completely identifiable at the end of each round from $y_t, \hat{x}_t$, so that I_t^o and I_t contain the same information. We also remark that $\hat{y}_t$ is completely determined by the algorithm and has no effect on the observation; therefore, all the new information between I_t and I_{t+1} only depends on c_{t+1}.

By the chain rule of KL divergence, for any $k \in [K]$, we have (see. e.g., eq. (9) in Garivier, Ménard, and Stoltz 2019)

$$\mathrm{KL}\big(\mathbb{P}_p^{I_T}, \mathbb{P}_{p_k}^{I_T}\big) = \sum_{t=1}^{T} \mathrm{KL}\Big(\mathbb{P}_p^{\hat{x}_t,\hat{y}_t,y_t,x_t,n_t,c_{t+1}|I_{t-1}}, \mathbb{P}_{p_k}^{\hat{x}_t,\hat{y}_t,y_t,x_t,n_t,c_{t+1}|I_{t-1}}\Big).$$

Each term can be greatly simplified to

$$\begin{aligned}
&\mathrm{KL}\Big(\mathbb{P}_p^{\hat{x}_t,\hat{y}_t,y_t,x_t,n_t,c_{t+1}|I_{t-1}}, \mathbb{P}_{p_k}^{\hat{x}_t,\hat{y}_t,y_t,x_t,n_t,c_{t+1}|I_{t-1}}\Big) \\
&\quad= \mathrm{KL}\Big(\mathbb{P}_p^{x_t,n_t|I_{t-1}}, \mathbb{P}_{p_k}^{x_t,n_t|I_{t-1}}\Big) + \underbrace{\mathrm{KL}\Big(\mathbb{P}_p^{\hat{x}_t,\hat{y}_t,y_t,c_{t+1}|I_{t-1}x_t,n_t}, \mathbb{P}_{p_k}^{\hat{x}_t,\hat{y}_t,y_t,c_{t+1}|I_{t-1}x_t,n_t}\Big)}_{=0} \\
&\quad\overset{(1)}{=} \mathrm{KL}\Big(\mathbb{P}_p^{x_t,n_t|I_{t-1}}, \mathbb{P}_{p_k}^{x_t,n_t|I_{t-1}}\Big) \\
&\quad\overset{(2)}{=} \mathrm{KL}\Big(\mathbb{P}_p^{n_t|I_{t-1}}, \mathbb{P}_{p_k}^{n_t|I_{t-1}}\Big) \\
&\quad\overset{(3)}{=} \mathbb{E}_p\big[\mathrm{KL}\big(\mathcal{N}(0,\sigma_n^2(c_t)), \mathcal{N}(0,\sigma_n^2(c_t;k))\big)\big] \\
&\quad\overset{(4)}{\le} \frac{1}{K^2}\mathbb{E}_p[\mathbb{1}\{c_t \in [c_k, c_{k+1})]\}].
\end{aligned}$$

In this derivation, relation (1) holds since given the same history and x_t, y_t, the variables $\hat{x}_t, y_t$ are deterministic, so their KL is zero. Moreover, given the same quantities, $\hat{y}$ and c_{t+1} are identically determined by the algorithm in both instances – have the same conditional distribution – and the KL divergence is equal to zero. Relation (2) holds since n_t and x_t are independent given I_{t-1} and x_t has the same distribution in both instances. (3) substitutes the conditional noise distribution and (4) is by Claim 6.

Denoting $N_T(k) = \sum_{t=1}^{T} \mathbb{1}\{c_t \in [c_k, c_{k+1})]\}$, we can thus bound

$$\mathrm{KL}\big(\mathbb{P}_p^{I_T}, \mathbb{P}_{p_k}^{I_T}\big) \leq \frac{1}{K^2}\mathbb{E}_p[N_T(k)].$$

On the other hand, by the data processing inequality (e.g., eq. (8) in Garivier, Ménard, and Stoltz 2019) it holds that

$$\mathrm{KL}\big(\mathbb{P}_p^{I_T}, \mathbb{P}_{p_k}^{I_T}\big) \geq \mathrm{kl}\left(\frac{\mathbb{E}_p[N_T(k)]}{T}, \frac{\mathbb{E}_{p_k}[N_T(k)]}{T}\right),$$

where $\mathrm{kl}(p,q) = p\ln\frac{p}{q} + (1-p)\ln\frac{1-p}{1-q}$. Combining both inequalities and using Pinsker's inequality, we get

$$\frac{1}{K^2}\mathbb{E}_p[N_T(k)] \geq 2\left(\frac{\mathbb{E}_p[N_T(k)]}{T} - \frac{\mathbb{E}_{p_k}[N_T(k)]}{T}\right)^2,$$

and reorganizing leads to the bound

$$\mathbb{E}_{p_k}[N_T(k)] \leq \frac{T}{K}\sqrt{\frac{\mathbb{E}_p[N_T(k)]}{2}} + \mathbb{E}_p[N_T(k)].$$

We are now ready to choose an instance k. Since the intervals $[c_k, c_{k+1})$ are disjoint and $\sum_{k\in[K]} N_T(k) \leq T$, by the pigeonhole principle, there exists k^* such that $\mathbb{E}_p[N_T(k^*)] \leq \frac{T}{K}$. For this instance, choosing $K = \lceil 2T^{1/3} \rceil$ it holds that

$$\mathbb{E}_{p_{k^*}}[N_T(k^*)] \leq \frac{1}{\sqrt{2}}\left(\frac{T}{K}\right)^{3/2} + \frac{T}{K} \leq \frac{3T}{4}.$$

Moreover, by Claim 6, we know that

$$\min_{c\in[0,1]} \ell^*(c;k) \leq \frac{1}{2} - \frac{1}{16K} \leq \frac{1}{2} - \frac{T^{-1/3}}{64},$$

while outside the interval $[c_k, c_{k+1})$, the loss is $\ell^*(c;k) = \frac{1}{2}$. Therefore, using our zero-mean Gaussian choice for x_t, n_t, by Claim 4 (and Claim 3), the regret under instance k^* is lower bounded by

$$\begin{aligned}
\mathbb{E}_{p_{k^*}}[R(T)] &\geq \mathbb{E}_{p_{k^*}}\left[\sum_{t=1}^{T} \ell^*(c_t;k^*) - \min_{c\in[0,1]} \ell^*(c;k^*)\right] \\
&\geq \mathbb{E}_{p_{k^*}}\left[\sum_{t=1}^{T} \mathbb{1}\{c_t \notin [c_{k^*}, c_{k^*+1})]\}\left(\ell^*(c_t;k^*) - \min_{c\in[0,1]} \ell^*(c;k^*)\right)\right] \\
&\geq \mathbb{E}_{p_{k^*}}\left[\sum_{t=1}^{T} \mathbb{1}\{c_t \notin [c_{k^*}, c_{k^*+1})]\}\frac{T^{-1/3}}{64}\right] \\
&= \frac{T^{-1/3}}{64}(T - \mathbb{E}_{p_{k^*}}[N_T(k^*)]) \\
&\geq \frac{T^{2/3}}{256}.
\end{aligned}$$

□

Claim 6. *1. For any c, k, the KL divergence between the two noise distributions is bounded by*

$$\mathrm{KL}\big(\mathcal{N}(0,\sigma_n^2(c)),\mathcal{N}(0,\sigma_n^2(c;k))\big) \le 16(c_{k+1}-c_k)^2\mathbb{1}\{c\in[c_k,c_{k+1})]\} = \frac{1}{K^2}\mathbb{1}\{c\in[c_k,c_{k+1})]\}.$$

2. The optimal loss in instance p_k is bounded by

$$\min_{c\in[0,1]}\ell^*(c;k) \le \frac{1}{2}-\frac{c_{k+1}-c_k}{4} = \frac{1}{2}-\frac{1}{16K}.$$

Proof. We start by stating a few properties of the noise variances under the two instances.

- Outside the interval $[c_k, c_{k+1}]$ it holds that $\sigma_n^2(c)=\sigma_n^2(c;k)$; in particular,
$$\mathrm{KL}\big(\mathcal{N}(0,\sigma_n^2(c)),\mathcal{N}(0,\sigma_n^2(c;k))\big)=0.$$
- We can rewrite the variance $\sigma_n^2(c)=f(c)=\frac{2}{1+c}-1$. Specifically, the variance slope $\frac{d\sigma_n^2(c)}{dc}=-\frac{2}{(1+c)^2}$ is increasing (becoming less negative) as c increases. Moreover, for $c\in\left[c_k,\frac{c_k+c_{k+1}}{2}\right)$, it holds that
$$\begin{aligned}\frac{d\sigma_n^2(c;k)}{dc} &= \frac{2}{c_{k+1}-c_k}(f(c_{k+1})-f(c_k)) = \frac{2}{c_{k+1}-c_k}\left(\frac{2}{1+c_{k+1}}-\frac{2}{1+c_k}\right) = -\frac{4}{(1+c_{k+1})(1+c_k)}\\ &\le -\frac{2}{(1+c_k)^2},\end{aligned}$$
where the last inequality is since $c_k, c_{k+1}\in[0,1]$. In other words, we have that $\sigma_n^2(c_k)=\sigma_n^2(c_k;k)$ and $\frac{d\sigma_n^2(c;k)}{dc}\le\frac{d\sigma_n^2(c_k)}{dc}\le\frac{d\sigma_n^2(c)}{dc}$ for all $c\in\left[c_k,\frac{c_k+c_{k+1}}{2}\right)$, and so inside this interval, the difference $\sigma_n^2(c)-\sigma_n^2(c;k)$ is non-negative and maximized at $c=\frac{c_k+c_{k+1}}{2}$. This cost is also the maximizer of the difference in $c\in[c_k,c_{k+1})$, since inside the interval $c\in\left[\frac{c_k+c_{k+1}}{2},c_{k+1}\right)$, $\sigma_n^2(c)$ decreases while $\sigma_n^2(c;k)$ remains constant. We can similarly conclude that $\sigma_n^2(c)-\sigma_n^2(c;k)\ge 0$ for $c\in\left[\frac{c_k+c_{k+1}}{2},c_{k+1}\right)$, since $\sigma_n^2(c)\ge\sigma_n^2(c_{k+1})=\sigma_n^2(c;k)$; therefore, the difference is non-negative across the interval $[c_k,c_{k+1})$. To summarize, we showed that for any $c\in[c_k,c_{k+1})$, it holds that
$$0\le\sigma_n^2(c)-\sigma_n^2(c;k)\le\sigma_n^2\left(\frac{c_k+c_{k+1}}{2}\right)-\sigma_n^2\left(\frac{c_k+c_{k+1}}{2};k\right)=\sigma_n^2\left(\frac{c_k+c_{k+1}}{2}\right)-\sigma_n^2(c_{k+1}).$$
We now explicitly bound the difference
$$\sigma_n^2\left(\frac{c_k+c_{k+1}}{2}\right)-\sigma_n^2(c_{k+1})=\frac{2}{1+\frac{c_k+c_{k+1}}{2}}-\frac{2}{1+c_{k+1}}=\frac{c_{k+1}-c_k}{\left(1+\frac{c_k+c_{k+1}}{2}\right)(1+c_{k+1})}\le c_{k+1}-c_k$$
To conclude this property, we have for all $c\in[c_k,c_{k+1})$ that
$$0\le\sigma_n^2(c)-\sigma_n^2(c;k)\le c_{k+1}-c_k. \tag{11}$$
- As stated near its definition, since $f(c)$ is decreasing in c, it can be easily verified that $\sigma_n^2(c;k)$ is also decreasing, and so for any $c\in[c_k,c_{k+1})$, it holds that
$$\sigma_n^2(c;k)\ge\sigma_n^2(c_{k+1};k)=f(c_{k+1})\ge f\left(\frac{3}{4}\right)\ge\frac{1}{8}.$$

Using these calculations, we are now ready to bound the KL divergence between the two distributions:

$$\begin{aligned}\mathrm{KL}\big(\mathcal{N}(0,\sigma_n^2(c)),\mathcal{N}(0,\sigma_n^2(c;k))\big) &= \frac{1}{2}\left(\frac{\sigma_n^2(c)}{\sigma_n^2(c;k)}-\ln\frac{\sigma_n^2(c)}{\sigma_n^2(c;k)}-1\right)\mathbb{1}\{c\in[c_k,c_{k+1})]\}\\ &\le \frac{\left(\frac{\sigma_n^2(c)}{\sigma_n^2(c;k)}-1\right)^2}{4}\mathbb{1}\{c\in[c_k,c_{k+1})]\} \qquad (\ln x\ge x-1-\tfrac{(x-1)^2}{2}\text{ for }x\ge 1)\\ &= \frac{\left(\sigma_n^2(c)-\sigma_n^2(c;k)\right)^2}{4\sigma_n^4(c;k)}\mathbb{1}\{c\in[c_k,c_{k+1})]\}\\ &\le 16(c_{k+1}-c_k)^2\mathbb{1}\{c\in[c_k,c_{k+1})]\},\end{aligned}$$

where the last inequality is by eq. (11) and the lower bound $\sigma_n^2(c;k) \geq \frac{1}{8}$. This concludes the first part of the claim.

For the second part, we have

$$\begin{aligned}
\min_{c\in[0,1]} \ell^*(c;k) &\leq \ell^*\left(\frac{c_k + c_{k+1}}{2}; k\right) \\
&= \frac{\sigma_n^2\left(\frac{c_k+c_{k+1}}{2}; k\right)}{1+\sigma_n^2\left(\frac{c_k+c_{k+1}}{2}; k\right)} + \frac{1}{2}\frac{c_k + c_{k+1}}{2} \\
&= \frac{f(c_{k+1})}{1+f(c_{k+1})} + \frac{c_k + c_{k+1}}{4} \\
&= \frac{1}{2} - \frac{c_{k+1}}{2} + \frac{c_k + c_{k+1}}{4} && \text{(by eq. (10), } \tfrac{f(c)}{1+f(c)} + \tfrac{c}{2} = \tfrac{1}{2}\text{)} \\
&= \frac{1}{2} - \frac{c_{k+1} - c_k}{4}.
\end{aligned}$$

□

D Useful Concentration Results

We start by stating two existing concentration results and then prove a lemma that combines both of them.

Lemma 4 (Matrix Azuma, Tropp 2012, Theorem 7.1)**.** *Let $M_1, M_2, \ldots, M_t \in \mathbb{R}^{d\times d}$ a sequence of self-adjoint matrices adapted to a filtration $\mathcal{F}_t$ such that $\mathbb{E}[M_t|\mathcal{F}_{t-1}] = 0$, and assume that there exist a fixed sequence of self-adjoint matrices $A_1, \ldots A_t$ such that $M_s^2 \preceq A_s^2$ a.s. for all s. Finally, denote $\sigma^2 = \left\|\sum_{s=1}^t A_s^2\right\|_{op}$. Then, for all $t \geq 1$ and $\delta > 0$,*

$$\Pr\left\{\lambda_{\max}\left(\sum_{s=1}^t M_s\right) \geq \sqrt{8\sigma^2 \ln\frac{d}{\delta}}\right\} \leq \delta.$$

Lemma 5 (Concentration of Subgaussian Norm, Hsu, Kakade, and Zhang 2012, Theorem 2.1 with $A = I$ and Remark 2.2)**.** *Let $X \in \mathbb{R}^d$ be a random R^2-subgaussian vector of mean $\mathbb{E}[X] = \mu$ such that*

$$\forall \alpha \in \mathbb{R}^d, \qquad \mathbb{E}\Big[e^{\alpha^\top (X-\mu)}\Big] \leq e^{\|\alpha\|^2 R^2/2}.$$

Then, for all $\delta > 0$,

$$\Pr\left\{\|X\|^2 \geq R^2\left(d + 2\sqrt{d\ln\frac{1}{\delta}} + 2\ln\frac{1}{\delta}\right) + \|\mu\|^2\left(1 + 2\sqrt{\frac{\ln\frac{1}{\delta}}{d}}\right)\right\} \leq \delta.$$

and $\mathbb{E}\Big[\|X_t - \mu\|^2\Big] \leq R^2 d$.

Lemma 6. *Let $X_1, \ldots, X_t \in \mathbb{R}^d$ be a sequence of R^2-subgaussians vectors adapted to the filtration $\mathcal{F}_t$ such that $\mathbb{E}[X_t|\mathcal{F}_{t-1}] = \mu_t$, $\mathbb{E}\big[X_t X_t^\top|\mathcal{F}_{t-1}\big] = \Sigma_t$ and*

$$\forall \alpha \in \mathbb{R}^d, \qquad \mathbb{E}\Big[e^{\alpha^\top (X_t-\mu_t)}|\mathcal{F}_{t-1}\Big] \leq e^{\|\alpha\|^2 R^2/2}.$$

Further assume that $\|\mu_t\| \leq S$ a.s. for all $t \geq 1$. Then, with probability at least $1-\delta$, for all $t \geq 1$, it holds that

$$\left\|\sum_{s=1}^t X_s X_s^\top - \Sigma_t\right\|_{op} < \sqrt{8t\beta_t^2 \ln\frac{3dt(t+1)}{\delta}}.$$

where $\beta_t = R^2\left(d + 2\sqrt{d\ln\frac{3t(t+1)}{\delta}} + 2\ln\frac{3t(t+1)}{\delta}\right) + S^2\left(1 + 2\sqrt{\frac{\ln\frac{3t(t+1)}{\delta}}{d}}\right)$.

Proof. Throughout the proof, we only work with symmetric matrices, for which $\|A\|_{op} = \max(\lambda_{\max}(A), -\lambda_{\min}(A))$.

We consider the matrix $M_t = X_t X_t^\top - \Sigma_t$ and aim to bound its maximal and minimal eigenvalues with high probability.

- To bound the minimal eigenvalue, notice that $M_t \succeq -\Sigma_t$ a.s., so we focus on bounding the maximal singular value of Σ_t. Specifically, we have

$$\lambda_{\max}(\Sigma_t) = \lambda_{\max}\big(\mathbb{E}\big[X_t X_t^\top|\mathcal{F}_{t-1}\big]\big) = \max_{u\in\mathbb{R}^d:\|u\|\leq 1} u^\top \mathbb{E}\big[X_t X_t^\top|\mathcal{F}_{t-1}\big]u \leq \mathbb{E}\left[\max_{u\in\mathbb{R}^d:\|u\|\leq 1} u^\top X_t X_t^\top u|\mathcal{F}_{t-1}\right]$$
$$= \mathbb{E}\Big[\|X_t\|^2|\mathcal{F}_{t-1}\Big].$$

  In addition, by Lemma 5, we have

$$R^2 d \geq \mathbb{E}\Big[\|X_t - \mu_t\|^2|\mathcal{F}_{t-1}\Big] = \mathbb{E}\Big[\|X_t\|^2|\mathcal{F}_{t-1}\Big] - 2\mathbb{E}[X_t|\mathcal{F}_{t-1}]^\top\mu_t + \|\mu_t\|^2 = \mathbb{E}\Big[\|X_t\|^2|\mathcal{F}_{t-1}\Big] - \|\mu_t\|^2,$$

  and therefore,

$$\lambda_{\max}(\Sigma) \leq R^2 d + \|\mu_t\|^2 \leq R^2 d + S^2,$$

  or $\lambda_{\min}(M_t) \geq -\big(R^2 d + S^2\big) \geq -\beta_t$ a.s. for all $t \geq 1$.
- To bound the maximal eigenvalue, define the event

$$E_n = \Big\{\forall t \geq 1 : \|X_t\|^2 \leq \beta_t\Big\}.$$

By Lemma 5 and the union bound, it holds that

$$\begin{aligned}
\Pr\{E_n\} &= 1 - \Pr\Big\{\exists t \geq 1 : \|X_t\|^2 > \beta_t\Big\} \\
&\geq 1 - \sum_{t=1}^{\infty} \Pr\Big\{\|X_t\|^2 > \beta_t\Big\} \\
&\geq 1 - \sum_{t=1}^{\infty} \Pr\left\{\|X_t\|^2 > R^2\left(d + 2\sqrt{d\ln\frac{3t(t+1)}{\delta}} + 2\ln\frac{3t(t+1)}{\delta}\right) + \|\mu_t\|^2\left(1 + 2\sqrt{\frac{\ln\frac{3t(t+1)}{\delta}}{d}}\right)\right\} \\
&\qquad\qquad (\|\mu_t\| \leq S) \\
&\geq 1 - \sum_{t=1}^{\infty} \frac{\delta}{3t(t+1)} \\
&= 1 - \frac{\delta}{3}.
\end{aligned}$$

Under E_n, since Σ_t is positive semi-definite, we have that

$$\lambda_{\max}(M_t) \leq \lambda_{\max}\big(X_t X_t^\top\big) = \max_{u\in\mathbb{R}^d:\|u\|\leq 1} u^\top X_t X_t^\top u = \|X_t\|^2 \leq \beta_t.$$

To summarize, under E_n, the operator norm of M_t is $\|M_t\|_{op} \leq \beta$; therefore, we have that a.s., $\big\|M_t^2 \mathbb{1}\{E_n\}\big\|_{op} \leq \beta^2$ (or $(M_t\mathbb{1}\{E_n\})^2 \preceq \beta_t^2 I$). We now apply Lemma 4 on the sequences $M_t^+ = M_t\mathbb{1}\{E_n\}$ and $M_t^- = -M_t\mathbb{1}\{E_n\}$.

Denoting $A_t = \beta_t I$, we saw that $(M_t^+)^2 \preceq A_t^2$ and $(M_t^-)^2 \preceq A_t^2$, where

$$\left\|\sum_{s=1}^{t} A_t^2\right\|_{op} = \left\|\sum_{s=1}^{t} \beta_t^2 I\right\|_{op} = t\beta_t^2.$$

Thus, by Lemma 4, the following two events hold w.p. at least $1 - \frac{\delta}{3t(t+1)}$:

$$\begin{aligned}
&\lambda_{\max}\left(\sum_{s=1}^{t} M_s\mathbb{1}\{E_n\}\right) = \lambda_{\max}\left(\sum_{s=1}^{t} M_s^+\right) < \sqrt{8t\beta_t^2\ln\frac{3dt(t+1)}{\delta}}, \quad \text{and,} \\
&\lambda_{\min}\left(\sum_{s=1}^{t} M_s\mathbb{1}\{E_n\}\right) = -\lambda_{\max}\left(\sum_{s=1}^{t} M_s^-\right) > -\sqrt{8t\beta_t^2\ln\frac{3dt(t+1)}{\delta}}.
\end{aligned}$$

Taking the union bound over both events and all $t \geq 1$ while recalling that $\sum_{t\geq 1}\frac{1}{t(t+1)} = 1$, we get that w.p. at least $1 - \frac{2\delta}{3}$,

$$\forall t \geq 1, \qquad \left\|\sum_{s=1}^{t} M_s\mathbb{1}\{E_n\}\right\|_{op} < \sqrt{8t\beta_t^2\ln\frac{3dt(t+1)}{\delta}}.$$

Now using the fact that $\Pr\{E_n\} \geq 1 - \frac{\delta}{3}$, we have

$$\begin{aligned}
&\Pr\left\{\exists t : \left\|\sum_{s=1}^{t}\big(X_s X_s^\top - \Sigma_s\big)\right\|_{op} \geq \sqrt{8t\beta_t^2\ln\frac{3dt(t+1)}{\delta}}\right\} \\
&\leq \Pr\left\{E_n, \exists t : \left\|\sum_{s=1}^{t}\big(X_s X_s^\top - \Sigma_s\big)\mathbb{1}\{E_n\}\right\|_{op} \geq \sqrt{8t\beta_t^2\ln\frac{3dt(t+1)}{\delta}}\right\} + \Pr\big\{\bar{E}_n\big\} \\
&\leq \Pr\left\{\exists t : \left\|\sum_{s=1}^{t} M_s\mathbb{1}\{E_n\}\right\|_{op} \geq \sqrt{8t\beta_t^2\ln\frac{3dt(t+1)}{\delta}}\right\} + \frac{\delta}{3} \\
&\leq \frac{2\delta}{3} + \frac{\delta}{3} \\
&= \delta.
\end{aligned}$$

□