Proceedings of Machine Learning Research vol 178:1–27, 2022 35th Annual Conference on Learning Theory

Faster online calibration without randomization: interval forecasts and the power of two choices

Chirag Gupta CHIRAGG@CMU.EDU
Carnegie Mellon University

Aaditya Ramdas ARAMDAS@CMU.EDU
Carnegie Mellon University

Editors: Po-Ling Loh and Maxim Raginsky

Abstract

We study the problem of making calibrated probabilistic forecasts for a binary sequence generated by an adversarial nature. Following the seminal paper of Foster and Vohra (1998), nature is often modeled as an adaptive adversary who sees all activity of the forecaster except the randomization that the forecaster may deploy. A number of papers have proposed randomized forecasting strategies that achieve an ϵ-calibration error rate of $O(1/\sqrt{T})$, which we prove is tight in general. On the other hand, it is well known that it is not possible to be calibrated without randomization, or if nature also sees the forecaster's randomization; in both cases the calibration error could be $\Omega(1)$. Inspired by the equally seminal works on the power of two choices and imprecise probability theory, we study a small variant of the standard online calibration problem. The adversary gives the forecaster the option of making *two nearby probabilistic forecasts*, or equivalently an interval forecast of small width, and the endpoint closest to the revealed outcome is used to judge calibration. This power of two choices, or imprecise forecast, accords the forecaster with significant power—we show that a faster ϵ-calibration rate of $O(1/T)$ can be achieved even without deploying any randomization.

1. Introduction

A number of machine learning and statistics applications rely on probabilistic predictions. In economics, the influential discrete choice framework uses probabilistic modeling at its core (McFadden, 1974). Spiegelhalter (1986) argued that when predictive models are used in medicine for detecting disease, categorizing patient risk, and clinical trials, it is imperative that they provide accurate probabilities, in order to appropriately guide downstream decisions. Weather forecasters (and their audiences) would like to know the probability of precipitation on a given day (Brier, 1950).

We study the problem of producing probabilistic forecasts for binary events, that are calibrated without any assumptions on the data-generating process. Informally, a forecaster is calibrated if, on all the days that the forecaster produces a forecast p_t that is approximately equal to $p \in [0, 1]$, the empirical average of the observations $y_t \in \{0, 1\}$ is also approximately equal to p, and this is true for every $p \in [0, 1]$ that is frequently close to a forecast (Dawid, 1982). We formalize this next.

1.1. Calibration games and ϵ-calibration

Calibration-Game-I (classical)
(nature is an adaptive adversary)

At time $t = 1, 2, \ldots,$

- Forecaster plays $u_t \in \Delta([0,1])$.
- Nature plays $v_t \in \Delta(\{0,1\})$.
- Forecaster predicts $p_t \sim u_t$.
- Nature reveals $y_t \sim v_t$.

Calibration-Game-II (POTC)
(nature is an adaptive adversary, forecaster has two nearby choices)

Fix $\epsilon > 0$. At time $t = 1, 2, \ldots,$

- Forecaster plays $p_{t0}, p_{t1} \in [0,1]$, such that $p_{t0} \leqslant p_{t1}$ and $|p_{t1} - p_{t0}| \leqslant 2\epsilon$.
- Nature reveals $y_t \in \{0,1\}$.
- If $y_t = 1$, set $p_t = p_{t1}$; else set $p_t = p_{t0}$.

Calibration-Game-I models the problem as a game between a forecaster and nature. The forecaster produces a *randomized* forecast $u_t \in \Delta([0,1])$, which is a distribution over the space of forecasts $[0,1]$. $\Delta(S)$ denotes the set of probability distributions over the set S (in every case, S is a standard set like $[0,1]$ with a canonical σ-algebra). Nature observes u_t and responds with a Bernoulli distribution for the outcome $v_t \in \Delta(\{0,1\}) = [0,1]$. We abuse notation slightly and use v_t to denote both the Bernoulli distribution and its parameter in $[0,1]$. Then forecaster and nature draw their actual actions, the forecast $p_t \sim u_t$ and the outcome $y_t \sim v_t$, simultaneously. At time $T > 1$, the prior activities $(u_t, v_t, p_t, y_t)_{t=1}^{T-1}$ are known to both players. The goal of the forecaster is to appear calibrated, defined shortly. Nature wishes to prevent the forecaster from appearing calibrated. Such a nature is typically referred to as an adaptive adversary.

Even before defining calibration formally, we can see that randomization is essential for the forecaster to demonstrate any semblance of being calibrated. If the forecaster is forced to put all his mass on a single p_t at each time (or equivalently if nature is an adaptive *offline* adversary), nature can play $v_t = y_t = \mathbf{1}\{p_t \leqslant 0.5\}$ to render the forecaster highly miscalibrated (Oakes, 1985; Dawid, 1985).

In anticipation of a forthcoming definition of ϵ-calibration error (equation (1)), we note that it will suffice for forecasters to only make discrete forecasts. Let $\epsilon > 0$ be a discretization or tolerance level, which is a small constant such as 0.1 or 0.01 depending on the application. For technical simplicity, we assume that $\epsilon = 1/2m$ for some integer $m \geqslant 2$. Consider the ϵ-cover of $[0,1]$ given by the m intervals $I_1 = [0, 1/m), I_2 = [1/m, 2/m), \ldots, I_m = [1-1/m, 1]$. At time t, the forecaster makes a forecast on the '2ϵ-grid' of the mid-points of these intervals:

$$p_t \in \{M_1 := 1/2m = \epsilon, M_2 := 3/2m = 3\epsilon, \ldots, M_m := 1 - 1/2m = 1 - \epsilon\}.$$

Denote the total number of times the forecast is $p_t = M_i$ until time $T \geqslant 1$ as

$$N_i^T := |\{t \leqslant T : p_t = M_i\}|,$$

and the observed average of the y_t's when $p_t = M_i$ as

$$p_i^T := \begin{cases} \frac{1}{N_i^T} \sum_{t \leqslant T : p_t = M_i} y_t & \text{if } N_i^T > 0, \\ M_i & \text{otherwise.} \end{cases}$$

Following Foster (1999), the (ℓ_1-)calibration error at time T, CE_T, is defined as the weighted sum of the prediction errors for each possible forecast:

$$\mathrm{CE}_T := \sum_{i=1}^{m} \frac{N_i^T}{T} \cdot \left| M_i - p_i^T \right|, \text{ or equivalently } \sum_{i=1}^{m} \left| \frac{1}{T} \sum_{t=1}^{T} \mathbf{1}\{p_t = M_i\}(M_i - y_t) \right|.$$

Finally, we define the ϵ-calibration error (ϵ-CE_T) as the calibration error with a slack of ϵ:

$$\epsilon\text{-}\mathrm{CE}_T := \max(\mathrm{CE}_T - \epsilon, 0). \tag{1}$$

In Calibration-Game-I, the forecaster and nature are allowed to randomize, thus ϵ-CE_T is a random quantity. A commonly studied object is its expected value. A forecaster is said to be ϵ-calibrated if, for *any* strategy of nature, the forecaster satisfies

$$\lim_{T\to\infty} \mathbb{E}\left[\epsilon\text{-}\mathrm{CE}_T\right] = 0, \text{ or equivalently } \mathbb{E}\left[\epsilon\text{-}\mathrm{CE}_T\right] = \underbrace{o_T(1)}_{f(T)}. \tag{2}$$

We are interested in the worst case value of $\mathbb{E}\left[\epsilon\text{-}\mathrm{CE}_T\right]$ against an adversarial nature, denoted as $f(T)$, henceforth called the ϵ-calibration rate or simply calibration rate. We show results about the asymptotic dependence of $f(T)$ as $T \to \infty$, holding ϵ as a fixed problem parameter on which f may depend arbitrarily.

1.2. Related work and our contributions

A number of papers have proposed ϵ-calibrated forecasting algorithms which guarantee $f(T) = O(1/\sqrt{T})$—the first was the seminal work of Foster and Vohra (1998), followed by a number of alternative proofs and generalizations of their result (Foster (1999), Fudenberg and Levine (1999), Vovk et al. (2005), Mannor and Stoltz (2010, Section 4.1), Abernethy et al. (2011, Theorem 22), Perchet (2015, Section 4.2)).

In Theorem 7, we show that the $O(1/\sqrt{T})$ rate achieved by these algorithms is tight. There is a strategy for nature that ensures $f(T) = \Omega(1/\sqrt{T})$. Our proof uses a non-constructive lower bound for Blackwell approachability games (Mannor and Perchet, 2013).

Qiao and Valiant (2021) recently showed that the worst-case calibration error without the ϵ-slack, $\mathbb{E}\left[\mathrm{CE}_T\right]$, is $\Omega(T^{-0.472})$. In contrast, we treat ϵ as a small constant fixed ahead of time, and consider lower bounds on $\mathbb{E}\left[\epsilon\text{-}\mathrm{CE}_T\right]$. Neither goal subsumes the other, so our lower bound complements theirs. In particular, observe that $\mathbb{E}\left[\mathrm{CE}_T\right] = \Omega(T^{-1/2})$ can be forced by nature by playing a non-adaptive Bernoulli strategy, drawing independently each $y_t \sim \text{Bernoulli}(p)$ for some p. This strategy seems insufficient for deriving a useful lower bound on $\mathbb{E}\left[\epsilon\text{-}\mathrm{CE}_T\right]$. These comparisons are further discussed in Section 4.6.

Foster (1999) showed that calibration is a Blackwell approachability instance (see Section 4.1), and while the rate $f(T) = \Omega(1/\sqrt{T})$ has not been formally established earlier (to the best of our knowledge), it is the rate one expects from a general Blackwell approachability instance (Cesa-Bianchi and Lugosi, 2006, Remark 7.7). Instead, the community has looked to establish positive results for alternative notions: calibration with more stringent tests than ϵ-calibration (Perchet, 2015; Rakhlin et al., 2011), calibration where the output space takes more than two values (Mannor and Stoltz, 2010), calibration with checking rules (Lehrer, 2001; Sandroni et al., 2003; Vovk et al., 2005), weak calibration (Kakade and

Foster, 2004), and smooth calibration (Foster and Hart, 2018). In particular, while no deterministic forecaster playing Calibration-Game-I can be ϵ-calibrated, there exist deterministic forecasters who are weakly/smoothly calibrated (Foster and Hart, 2018).

In our work, we take a slightly different approach from these papers. We retain the classical definition of ϵ-calibration but change the calibration game. In Calibration-Game-II, also called the Power-Of-Two-Choices (POTC) game, the forecaster reveals two forecasts $p_{t0}, p_{t1} \in [0,1]$, such that $p_{t0} \leqslant p_{t1}$ and $|p_{t1} - p_{t0}| \leqslant 2\epsilon$. Since the earlier binning scheme used a 2ϵ-grid, that is $1/m = 2\epsilon$, this effectively allows the forecaster to choose a full bin as their forecast (rather than its midpoint), or equivalently to choose two consecutive bin midpoints. Thus there is no randomization, and nature knows the two forecasts. If nature chooses to play $y_t = 0$, $p_t = p_{t0}$ is used to judge the calibration of the forecaster, and if nature chooses to play $y_t = 1$, $p_t = p_{t1}$ is used. (One could say that the forecast closer to reality is used for measuring calibration, or that the forecaster decides which one of the two forecasts to use; these are all equivalent.) Obviously, without the restriction of p_{t0}, p_{t1} being 2ϵ-close, the problem is trivial: the forecaster would predict $p_{t0} = 0$ and $p_{t1} = 1$ in each round, and achieve zero error in every round. Requiring $|p_{t1} - p_{t0}| \leqslant 2\epsilon$ makes the problem interesting. The POTC setup may appear surprising to some and we devote Section 2 to motivating it.

The summary of our main result (Theorem 2) is as follows. In the POTC game, the forecaster can ensure—deterministically—that

$$\epsilon\text{-CE}_T = O(1/T). \tag{3}$$

Compared to (2), there is no expectation operator anymore since the forecaster is deterministic and nature being fully adaptively adversarial does not benefit from randomizing.

Our forecaster is a variant of Foster (1999). While Foster's forecaster randomizes over two nearby forecasts (making it *almost deterministic* in the sense of Foster and Hart (2021)), our forecaster predicts both these values and is judged with respect to the better one (and is actually, not almost, deterministic).

Remark 1 (Generalization from binary to bounded outputs) *The POTC game can be modified for bounded, instead of binary, outputs. That is, nature can play $v_t \in [0,1]$ and calibration would be judged with respect to the average of the v_t's on the instances when $p_t = M_i$. Note that this is not the same as nature playing $y_t \sim$ Bernoulli(v_t), since the calibration loss (left-hand-side of (3)) is not linear in y_t. With bounded outputs, the same $O(1/T)$ calibration rate can be achieved by a minor modification to our proposed forecasting strategy; see Appendix C for more details. A similar remark holds for Calibration-Game-I and the corresponding lower bound of $\Omega(1/\sqrt{T})$. This latter fact is evident without further details since the lower bound can only increase if nature is given more flexibility.*

Organization. Section 2 provides further context and motivation for the POTC game. Section 3 presents our algorithm for the POTC game and proves the fast calibration rate of $O(1/T)$ for it (Theorem 2). Section 4 reviews the well-known equivalence between calibration and Blackwell's approachability theorem (Blackwell, 1956), using which we prove the slow calibration rate of $\Omega(1/\sqrt{T})$ for Calibration-Game-I (Theorem 7). Most proofs are presented alongside the results. Section 5 concludes with a discussion.

2. Motivation for the POTC calibration game

Calibration-Game-II or the POTC game is motivated by two rich fields of literature: imprecise probability and the power of two choices.

2.1. A practical perspective via imprecise probability

The reader may wonder what the practical usefulness of the POTC game is. Why would we judge the forecaster in such a manner? The answer is that our earlier problem was phrased in a fashion that makes the connection to the power of two choices transparent. But one can also re-cast the problem in the language of *imprecise probability*. In this area, one is typically not restricted to work with single, unique probability measures, but instead the axioms of probability are relaxed, and added flexibility is provided in order to work with *upper* and *lower* probability measures (Walley and Fine, 1982).

In the context of our problem, instead of saying that the probability of rain is 0.3, a forecaster is allowed to say $0.3 \pm \epsilon$. One may just say that the forecaster is slightly uncertain and does not wish to commit to a point forecast, and indeed we may not force a forecaster to announce a point forecast against their will. From a Bayesian or game-theoretic perspective, we may say that the forecaster allows bets against their forecast, represented as a contract which pays off y_t, but the forecaster's prices for buying and selling such a contract are slightly different. From a practical perspective, this type of *interval* forecast arguably has almost the same utility and interpretability to a layman as the corresponding point forecast. The use of upper and lower forecasts (translated to prices or betting odds) is standard in game-theoretic probability (Shafer and Vovk, 2019). Separately, the recent work of de Cooman and De Bock (2022) establishes that *randomness is inherently imprecise* in a formal sense, and provides a different justification for the use of interval forecasts for binary sequences.

Remarkably, this small and seemingly insignificant change in reporting leads to a huge change in our ability to achieve calibration. (This gain can be rather puzzling: we were binning/gridding anyway, so why not report a full bin rather than its midpoint? How could that possibly improve our calibration error?!) Of course, we must figure out how to judge the quality of such an interval forecast: we must swap out $\mathbf{1}\{p_t = M_i\}(M_i - y_t)$ in CE_T for a generalized notion of error that dictates how far y_t was from the forecasted interval $A_t := [p_{t0}, p_{t1}]$. To do this, we use the distance from a point to a convex set: we replace M_i with the projection of y_t onto A_t, denoted $\mathrm{proj}(y_t, A_t)$. This is exactly what our POTC version does, just expressed differently.

When we generalize the definition of calibration, it is notationally simpler to restrict the forecasted interval endpoints to be the same m gridpoints, meaning that $A_t = [M_i, M_{i+1}]$ for some $i \leqslant m-1$ (rather than $A_t = I_i$, the intervals whose midpoints are M_i). In this case, we call a method that produces interval forecasts $(I_t)_{t \geqslant 1}$ as being ϵ-calibrated if:

$$\max\left(\sum_{i=1}^{m}\left|\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}\{\mathrm{proj}(y_t, A_t) = M_i\}\mathrm{dist}(y_t, A_t)\right| - \epsilon, 0\right) = o(1), \tag{4}$$

where $\mathrm{dist}(y, A) := |y - \mathrm{proj}(y, A)|$ is the distance of y to interval A. (When the interval is a single point, we recover the original definition of calibration, but in this case we know that randomization is necessary for ϵ-calibration.)

Imprecise probability has also made an intriguing appearance in the simpler setting commonly considered in machine learning, of achieving calibration in offline binary classification in the presence of i.i.d. data. This is a problem where theoretical progress has been made on designing *distribution-free* algorithms that have calibration guarantees by just assuming that the covariate-label pairs of data are i.i.d., while also performing well on real data (Gupta et al., 2020; Gupta and Ramdas, 2021, 2022). Venn predictors are a class of distribution-free algorithms that produce imprecise probability forecasts (Vovk et al., 2003; Vovk and Petej, 2014). On observing the covariates of a new point, Venn predictors output a particular interval of probabilities $[p_0, p_1]$ for the unknown binary label. A strong, but slightly odd, calibration property holds: the authors prove that p_Y (a random and unknown prediction, since Y is unknown and random) achieves *exact* calibration in finite samples. One can, in some sense, view our work as extending the use of such imprecise interval forecasts to the online calibration setting with adversarial data.

2.2. The varied applications of the power of two choices

The power of two choices (POTC) refers to a remarkable result by Azar et al. (1994) for the problem of load balancing. Suppose n balls are placed independently and uniformly at random into n bins. It can be shown that with high probability, the maximum number of balls in a bin (the maximum *load*) will be $\widetilde{\Theta}(\log n)$. Consider a different setup where the balls are placed sequentially, and for each ball, two bin indices are drawn uniformly at random and offered to a *load-balancer* who gets to decide which of the two bins to place the ball in. The load-balancer attempts to reduce the maximum load by following a natural strategy: at each step, place the ball in the bin with lesser load. It turns out that with this strategy the maximum load drops exponentially to $\widetilde{\Theta}(\log \log n)$.

The POTC result has led a number of applications. In a network where one of many servers can fulfil a request, it is exponentially better to choose two servers (instead of one) at random and allocate the server with fewer existing requests (Azar et al., 1994). Using two hashes instead of one significantly reduces the load of a single hash bucket (Broder and Mitzenmacher, 2001). In circuit routing, selecting one of two possible circuits provably leads to decongestion (Cole et al., 1998). When allocating a task to one of many resources where an intensive query needs to be made about the resource capacity, querying two resources is often better than querying all resources, or querying a single resource (Azar et al., 1994). Recently, Dwivedi et al. (2019) used the POTC to develop an online thinning algorithm that produces *low-discrepancy sequences* on hypercubes, with applications to quasi Monte Carlo integration. For further applications and a survey of mathematical techniques, we refer the reader to the thesis of Mitzenmacher (1996), or the survey by Mitzenmacher et al. (2001).

In this paper, we find yet another intriguing phenomenon involving the POTC, this time in the context of calibration. We modify the classical setup of calibration (Calibration-Game-I) to the POTC setup (Calibration-Game-II), by offering the forecaster two nearby choices. We show that this change accords the forecaster with significant power, enabling faster calibration, even without randomization.

POTC-Cal: Algorithm for forecaster in Calibration-Game-II
(for notation, see Section 3 below)

- At time $t = 1$, play $(p_{10}, p_{11}) = (M_1, M_1)$. Thus $p_1 = M_1$.
- At time $t + 1$ ($t \geqslant 1$), if

  condition A: there exists an $i \in [m]$ such that $d_i^t \leqslant 0$ and $e_i^t \leqslant 0$,

  is satisfied, play (M_i, M_i) for any i that verifies condition A (that is, $p_{t+1} = M_i$). Otherwise,

  condition B: there exists an $i \in [m-1]$ such that $e_i^t > 0$ and $d_{i+1}^t > 0$,

  must be satisfied (see Lemma 14). Play (M_i, M_{i+1}) for any index i that verifies condition B (that is, $p_{t+1} = M_i$ if $y_{t+1} = 0$ and $p_{t+1} = M_{i+1}$ if $y_{t+1} = 1$).

3. Main results: algorithm and analysis

Consider the POTC game (Calibration-Game-II). Recall that the forecaster's probabilities correspond to the mid-points of the intervals $I_1 = [0, 1/m), \ldots, I_m = [1 - 1/m, 1]$, given by $M_1 = 1/2m, \ldots, M_m = 1 - 1/2m$. The forecaster can play either $(p_{t0}, p_{t1}) = (M_i, M_i)$ or $(p_{t0}, p_{t1}) = (M_i, M_{i+1})$ for some i. We can also say that the forecaster predicts one of the two intervals $\{M_i\}$ or $[M_i, M_{i+1}]$ respectively.

We introduce some notation to describe the algorithm. For $i \in [m] := \{1, 2, \ldots, m\}$ and $t \geqslant 1$, define:

$$\begin{aligned}
\text{(left endpoint of interval } i\text{)} \quad & l_i = (i-1)/m, \\
\text{(right endpoint of interval } i\text{)} \quad & r_i = i/m, \\
\text{(frequency of interval } i\text{)} \quad & N_i^t = \left|\{\mathbf{1}\{p_s = M_i\} : s \leqslant t\}\right|, \\
\text{(observed average when } M_i \text{ was forecasted)} \quad & p_i^t = \begin{cases} \sum_{s=1}^t y_s \mathbf{1}\{p_s = M_i\}/N_i^t & \text{if } N_i^t > 0 \\ M_i & \text{if } N_i^t = 0, \end{cases} \\
\text{(deficit)} \quad & d_i^t = l_i - p_i^t, \\
\text{(excess)} \quad & e_i^t = p_i^t - r_i.
\end{aligned}$$

The terminology 'deficit' alludes to the fact that if p_i^t is smaller than desired (to the left of l_i), then $d_i^t > 0$ (p_i^t is 'in deficit'). 'Excess' has the opposite interpretation.

3.1. Forecasting algorithm

The algorithm, presented on top of this page, is a variant of the one proposed by Foster (1999). Foster's forecaster isolates two relevant M_i's and randomizes over them; we use the same M_i's to form the reported interval. At time $t + 1$, if there is a forecast M_i that is already 'good' in the sense that $p_i^t \in [l_i, r_i]$, the forecaster predicts M_i. Otherwise, the

forecaster finds two consecutive values (M_i, M_{i+1}) such that p_i^t is in excess and p_{i+1}^t is in deficit (such an i exists by Lemma 14, Appendix A). The forecaster plays (M_i, M_{i+1}). If nature reveals $y_{t+1} = 0$, then $p_{t+1} = M_i$, and the excess of p_i^t decreases. If nature reveals $y_{t+1} = 1$, then $p_{t+1} = M_{i+1}$, and the deficit of p_{i+1}^t decreases.

3.2. Analysis of POTC-Cal

We now present our main result along with a short proof.

Theorem 2 *POTC-Cal satisfies, at any time $T \geqslant 1$, for any strategy of nature,*

$$\epsilon\text{-}CE_T \leqslant m/T. \tag{5}$$

Proof Consider any $t \geqslant 1$. We write each of the m terms in the calibration error at time t, CE_t, as follows:

$$\left| \frac{1}{t} \sum_{s=1}^{t} \mathbf{1}\{p_s = M_i\}(M_i - y_s) \right| = \frac{N_i^t \left| M_i - p_i^t \right|}{t} = \frac{N_i^t(\epsilon + \max(d_i^t, e_i^t))}{t}.$$

Define $E_t^{(i)} := N_i^t \max(d_i^t, e_i^t)$, and observe that

$$\epsilon\text{-CE}_T = \max\left(\sum_{i=1}^{m} \frac{N_i^T \epsilon + E_T^{(i)}}{T} - \epsilon, 0 \right) = \max\left(\sum_{i=1}^{m} \frac{E_T^{(i)}}{T}, 0 \right).$$

We will show that for every $i \in [m]$, $E_T^{(i)} \leqslant 1$, proving the theorem.

Consider some specific $i \in [m]$. If action i is never played, then $N_i^T = 0$, and $E_T^{(i)} = 0$. Suppose an action i has $N_i^T > 0$. For each $1 \leqslant t < T$, if $a_{t+1} \neq i$, then $E_{t+1}^{(i)} = E_t^{(i)}$. If $a_{t+1} = i$, then by Lemmas 3 and 4 (stated and proved below), $E_{t+1}^{(i)} \leqslant \max(E_t^{(i)}, 1)$. In other words, at all t, the value of $E_{t+1}^{(i)}$ either stays bounded by 1, or decreases compared to the previous value $E_t^{(i)}$. A trivial inductive argument thus implies $E_T^{(i)} \leqslant 1$. For completeness, we verify the base case: since $p_1 = M_1$, $E_1^{(i \neq 1)} = 0$ and $E_1^{(1)} \leqslant 1$ (as $d_1^1 \leqslant 0$ and $e_1^1 \leqslant 1$). ■

Lemma 3 *Suppose condition A was satisfied at time $t+1$ and the forecast was $p_{t+1} = M_i$. Then, $N_i^{t+1} \max(d_i^{t+1}, e_i^{t+1}) \leqslant 1$.*

Proof Since $p_{t+1} = M_i$, $N_i^{t+1} = N_i^t + 1$, and $N_i^{t+1} p_i^{t+1} = N_i^t p_i^t + y_{t+1}$. Then,

$$\begin{aligned} \left| d_i^{t+1} - d_i^t \right| = \left| e_i^{t+1} - e_i^t \right| = \left| p_i^{t+1} - p_i^t \right| &= \left| \frac{N_i^t p_i^t + y_{t+1}}{N_i^t + 1} - \frac{N_i^t p_i^t + p_i^t}{N_i^t + 1} \right| \\ &= \left| \frac{y_{t+1} - p_i^t}{N_i^t + 1} \right| \leqslant \frac{1}{N_i^t + 1} = \frac{1}{N_i^{t+1}}. \end{aligned} \tag{6}$$

Since by condition A, $\max(d_i^t, e_i^t) \leqslant 0$, we obtain $\max(d_i^{t+1}, e_i^{t+1}) \leqslant 1/N_i^{t+1}$. ■

Lemma 4 *Suppose condition A was not satisfied at time $t+1$ and the forecast was $p_{t+1} = M_i$, following condition B. Then $N_i^{t+1} \max(d_i^{t+1}, e_i^{t+1}) \leqslant \max(N_i^t \max(d_i^t, e_i^t), 1)$.*

Proof Suppose $y_{t+1} = 0$. Since we are playing as per condition B, $e_i^t > 0$. Since $d_i^t + e_i^t = l_i - r_i = -1/m$, we have that $d_i^t < 0$. Thus,

$$y_{t+1} = 0 \implies e_i^t > 0 \text{ and } d_i^t < 0.$$

Similarly, it can be verified that $y_{t+1} = 1 \implies e_i^t < 0$ and $d_i^t > 0$. Below we assume without loss of generality that $y_{t+1} = 0$. (A similar argument goes through for the case $y_{t+1} = 1$.)

We derive how $N_i^t \max(d_i^t, e_i^t)$ changes when going from t to $t+1$. There are two cases: $e_i^{t+1} \geqslant d_i^{t+1}$ or $e_i^{t+1} < d_i^{t+1}$. If $e_i^{t+1} \geqslant d_i^{t+1}$, then

$$\begin{aligned} N_i^{t+1} \max(d_i^{t+1}, e_i^{t+1}) = N_i^{t+1} e_i^{t+1} &= N_i^{t+1} p_i^{t+1} - N_i^{t+1} r_i \\ &= N_i^t p_i^t - N_i^{t+1} r_i \quad (\text{since } y_{t+1} = 0) \\ &= N_i^t e_i^t - r_i \\ &= N_i^t \max(d_i^t, e_i^t) - r_i \leqslant N_i^t \max(d_i^t, e_i^t). \end{aligned}$$

On the other hand if $e_i^{t+1} < d_i^{t+1}$, then,

$$\begin{aligned} N_i^{t+1} \max(d_i^{t+1}, e_i^{t+1}) &= N_i^{t+1} d_i^{t+1} \\ &\leqslant N_i^{t+1} (d_i^t + \left| d_i^{t+1} - d_i^t \right|) \\ &\overset{(*)}{<} N_i^{t+1} (0 + 1/N_i^{t+1}) = 1. \end{aligned}$$

Inequality $(*)$ holds since $d_i^t < 0$ and $\left| d_i^{t+1} - d_i^t \right| \leqslant 1/N_i^{t+1}$ (see set of equations (6)). ■

4. $\Omega(1/\sqrt{T})$ lower bound for the classical calibration game

Calibration-Game-I can be viewed as a repeated game with vector-valued payoffs/rewards. Such games were studied by Blackwell (1956), and are now commonly referred to as Blackwell approachability games. We review the reduction from calibration to Blackwell approachability and use it to prove the lower bound. Throughout this section, we denote the action space of the forecaster as $\mathcal{X} = \{M_1, M_2, \ldots, M_m\}$ and that of nature as $\mathcal{Y} = \{0, 1\}$. The random plays of the forecaster lie in $\Delta(\mathcal{X})$ which is a probability simplex in m dimensions. We embed $\Delta(\mathcal{X})$ in $\mathbb{R}^m$ to simplify discussion.

4.1. Calibration as an instance of Blackwell approachability

The fact that calibration can be modelled as a Blackwell approachability instance is well-known (since Foster (1999); Hart and Mas-Colell (2000)). Suppose the actions of the forecaster and nature give a reward $r : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^m$ defined as follows: the i-th component of the reward vector $a = r(p \in \mathcal{X}, y \in \mathcal{Y}) \in \mathbb{R}^m$ is given by

$$a_i = \mathbf{1}\{p = M_i\} \cdot (M_i - y). \tag{7}$$

Let $\bar{a}^T := \sum_{i=1}^T r(p_t, y_t)/T$ be the average reward vector given component-wise by $\bar{a}_i^T = \sum_{t=1}^T \mathbf{1}\{p_t = M_i\}(M_i - y_t)/T$. Let B_ϵ be the ℓ_1-ball with radius ϵ, and dist the ℓ_1-distance function. Note that

$$\text{dist}(\bar{a}^T, B_\epsilon) = \max\left(\sum_{i=1}^m \left|\frac{1}{T}\sum_{t=1}^T \mathbf{1}\{p_t = M_i\}(M_i - y_t)\right| - \epsilon, 0\right) = \epsilon\text{-CE}_T.$$

Thus, the ϵ-calibration condition (2) is equivalent to $\lim_{T\to\infty} \mathbb{E}\left[\text{dist}(\bar{a}^T, B_\epsilon)\right] = 0$. If this condition is satisfied, we say that $\bar{a}^T$ *approaches* B_ϵ in the limit. Blackwell (1956) established necessary and sufficient conditions for approachability.

Theorem 5 (Corollary to Theorem 3 by Blackwell (1956)) *Assume the same setup as Calibration-Game-I, but the players receive a vector-valued reward r, as defined component-wise in (7). The forecaster can ensure that $\lim_{T\to\infty} \mathbb{E}\left[dist(\bar{a}^T, B_\epsilon)\right] = 0$ if and only if for every $v \in \Delta(\mathcal{Y}) = [0,1]$, there exists $u \in \Delta(\mathcal{X})$ such that the expected reward belongs to B_ϵ:*

$$\forall v \in \Delta(\mathcal{Y}),\ \exists u \in \Delta(\mathcal{X}) : \mathbb{E}_{p\sim u, y\sim v}\left[r(p,y)\right] = \sum_{i=1}^m u_i((1-v)\cdot r(M_i,0) + v\cdot r(M_i,1)) \in B_\epsilon. \quad (8)$$

The hypothetical situation considered in the theorem is akin to a one-shot game but with the order of the players reversed: nature plays v first and the forecaster responds with u. If the forecaster can *respond* to every play by nature and ensure that the expected reward lies in B_ϵ, then the forecaster can ensure that $\bar{a}^T$ approaches B_ϵ in the sequential game (where nature goes second each time). Abernethy et al. (2011) call this *response-satisfiability*; in their words, Theorem 5 is interepreted as response-satisfiability $\iff$ approachability.

Proposition 6 *The forecaster can exhibit response satisfiability (8). Thus the forecaster playing Calibration-Game-I can ensure $\lim_{T\to\infty} \mathbb{E}\left[dist(\bar{a}^T, B_\epsilon)\right] = 0$ and be ϵ-calibrated (2).*

The proof of this well-known result is in Appendix A.3. A second question is of the rate at which the expected reward vector approaches the desired set. As reviewed in the introduction, a number of papers have shown that the rate of approachability for the ϵ-calibration game is $O(1/\sqrt{T})$. We show that this rate cannot be improved.

4.2. $\Omega(1/\sqrt{T})$ lower bound for the ϵ-calibration error rate

Theorem 7 *A forecaster playing Calibration-Game-I against an adversarial nature cannot achieve ϵ-calibration at a rate faster than $O(1/\sqrt{T})$. That is, for every strategy of the forecaster, there is a strategy of nature that ensures*

$$\mathbb{E}\left[\epsilon\text{-}CE_T\right] = \Omega(1/\sqrt{T}). \quad (9)$$

Mannor and Perchet (2013) analyzed the convergence rate in approachability games and characterized conditions, which if satisfied by a target set $\mathcal{C}$, entail that nature can ensure $\mathbb{E}\left[\text{dist}(\bar{a}^T, \mathcal{C})\right] = \Omega(1/\sqrt{T})$. In particular, if these conditions are satisfied B_ϵ, Theorem 7 follows immediately since $\text{dist}(\bar{a}^T, \mathcal{C}) = \epsilon\text{-CE}_T$.

Theorem 8 (Theorem 6.ii by Mannor and Perchet (2013)) *Let $\mathcal{C}$ be a closed convex set that is (i) minimal approachable, and (ii) mixed approachable. Then $\mathcal{C}$ cannot be approached at a rate faster than $O(1/\sqrt{T})$, or in other words, $\mathbb{E}\left[dist(\bar{a}^T, \mathcal{C})\right] = \Omega(1/\sqrt{T})$, where $\bar{a}^T$ is the average reward vector.*

In what follows, we introduce the conditions (i) minimal approachability and (ii) mixed approachability in the context of our calibration game, and show that they are satisfied by B_ϵ (Lemma 10 and Lemma 13 respectively).

4.3. Minimal approachability

For a point $u \in \mathbb{R}^m$ and a convex set $K \subseteq \mathbb{R}^m$, define the distance of u from K as $d_K(u) = \inf_{u' \in K} \|u - u'\|_2$. For any $\lambda > 0$, a convex set $K' \subseteq K$ is said to be a λ-shrinkage of K if $\{u : d_{K'}(u) \leqslant \lambda\} \subseteq K$. In the following definition of minimal approachability, we implicitly assume that the set of action sets of the players and the corresponding rewards have been fixed, and the goal is to characterize which convex sets are approachable and which are not.

Definition 9 *A set K is* minimal *approachable if K is approachable, but no λ-shrinkage of K is approachable.*

We now show the first condition required by Theorem 8.

Lemma 10 *The set B_ϵ is minimal approachable.*

Proof Proposition 6 shows that B_ϵ is approachable so it remains to show that the minimality condition holds. Let K be a λ-shrinkage of B_ϵ for some $\lambda > 0$. We first argue that $K \subseteq B_{\epsilon-\lambda}$. Suppose this were not the case, that is, there exists $u \in K$ such that $u \notin B_{\epsilon-\lambda}$. By definition of the ℓ_1-ball $B_{\epsilon-\lambda}$, this means that $\|u\|_1 > \epsilon - \lambda$. We will show that such a u cannot belong to any λ-shrinkage of B_ϵ, in particular it cannot belong to K, leading to a contradiction.

Consider the point $u' = u + (\lambda/\|u\|_2)u$. Note that $d_K(u') \leqslant \|u - u'\|_2 = \lambda$. Since K is a λ-shrinkage of B_ϵ, this implies that $u' \in B_\epsilon$, or $\|u'\|_1 \leqslant \epsilon$. On the other hand, we have,

$$\begin{aligned}
(\epsilon \geqslant)\ \|u'\|_1 &= \|u\|_1(1 + \lambda/\|u\|_2) \\
&\geqslant \|u\|_1(1 + \lambda/\|u\|_1) && \text{(for any vector } v \in \mathbb{R}^m,\ \|v\|_1 \geqslant \|v\|_2) \\
&= \|u\|_1 + \lambda > (\epsilon - \lambda) + \lambda = \epsilon,
\end{aligned}$$

which is a contradiction. Thus $K \subseteq B_{\epsilon-\lambda}$, as claimed.

It follows that K is approachable only if $B_{\epsilon-\lambda}$ is approachable. We now show that for every $\lambda > 0$, $B_{\epsilon-\lambda}$ is not approachable. As in the proof of Proposition 6, the i-th component of the reward vector is given by $u_i(M_i - v)$. Suppose $v = 1/m$. Then for every M_i, $|M_i - v| \geqslant 1/2m = \epsilon$, by definition of m. Thus $|u_i(M_i - v)| \geqslant u_i\epsilon$, and $\sum_{i=1}^m |u_i(M_i - v)| \geqslant \sum u_i\epsilon = \epsilon$. Equivalently, $\mathbb{E}\left[r(p, y)\right] \notin B_{\epsilon-\lambda}$. By Theorem 5, $B_{\epsilon-\lambda}$ is not approachable. ■

In order to describe the second condition required by Theorem 8 and show that it holds for B_ϵ, we need additional technical setup. The following subsection serves this purpose.

4.4. Reducing approachability to scalar-valued games

The vector-valued approachability game induces a number of scalar-valued min-max games, one for each direction in $\mathbb{R}^m$. The value of these scalar games is closely connected to the question of approachability.

Consider the approachability of B_ϵ with respect to individual directions, represented by arbitrary vectors $q \in \mathbb{R}^m$ (for intuition, one may equivalently think of q being direction vectors, those with ℓ_2-norm equal to one, but this restriction is technically unnecessary; we stick to $q \in \mathbb{R}^m$). Let $c \in B_\epsilon$ be such that q belongs to the *normal cone* of B_ϵ at c, that is, $\langle c, q\rangle = \sup_{c' \in B_\epsilon} \langle c', q\rangle = \epsilon\|q\|_\infty$. We call such a pair (c, q) as *admissible*. Consider the following one-shot min-max game defined for every admissible (c, q):

$$\begin{aligned}
\mathrm{Val}(c, q) &= \min_{u \in \Delta(\mathcal{X})} \max_{v \in [0,1]} \langle \mathbb{E}_{p\sim u, y\sim v}\left[r(p, y)\right] - c, q\rangle \\
&= \min_{u \in \Delta(\mathcal{X})} \max_{v \in [0,1]} \left(\sum_{i=1}^{m} u_i q_i((1-v)M_i + v(M_i - 1)) - \langle c, q\rangle \right) \\
&= \min_{u \in \Delta(\mathcal{X})} \max_{v \in [0,1]} \left(\sum_{i=1}^{m} u_i q_i(M_i - v) - \epsilon\|q\|_\infty \right).
\end{aligned}$$

To appreciate the relationship between the $\mathrm{Val}(c, q)$ games and the B_ϵ-approachability game, consider the following. Suppose the forecaster can guarantee *one-shot approachability*, that is, there exists a fixed $u^\star \in \Delta(\mathcal{X})$ such that for every $v \in [0, 1]$, $\mathbb{E}_{p\sim u^\star, y\sim v}\left[r(p, y)\right] \in B_\epsilon$. By definition of the normal cone, for every admissible (c, q), and any $c' \in B_\epsilon$, it holds that $\langle c' - c, q\rangle \leqslant 0$. In particular, $\mathbb{E}_{p\sim u^\star, y\sim v}\left[r(p, y)\right] \in B_\epsilon$ is such a c'. It follows that for every admissible (c, q), Val(c, q) $\leqslant 0$.

This observation does not hold in the reverse direction: even if $\mathrm{Val}(c, q) \leqslant 0$ for every admissible (c, q), one-shot approachability need not hold. The intuition is that the optimal u for different (c, q) can be different, and it is unclear how to merge them to achieve one-shot forecasting for the approachability game. In a remarkable result, Blackwell (1956) showed that the result does hold in the reverse direction for the *repeated* approachability game (as opposed to the one-shot approachability game).

Theorem 11 (Theorem 1 by Blackwell (1956)) *A convex set K is approachable if and only if for every admissible (c, q), $\mathrm{Val}(c, q) \leqslant 0$.*

This condition has also been termed as halfspace-satisfiability by Abernethy et al. (2011). The result was stated in the language of convex cones by Mannor and Perchet (2013). Notice that for our problem, the min-max game does not depend on c, once we replace $\langle c, q\rangle$ with $\epsilon\|q\|_\infty$. We thus simplify notation and index our games only by $q \in \mathbb{R}^m$:

$$\mathrm{Val}(q) := \min_{u \in \Delta(\mathcal{X})} \max_{v \in [0,1]} \left(\sum_{i=1}^{m} u_i q_i(M_i - v) - \epsilon\|q\|_\infty \right). \tag{10}$$

We know that B_ϵ is approachable (Proposition 6) and hence by Blackwell's result, halfspace-satisfiability must hold. That is, for every $q \in \mathbb{R}^m$, $\mathrm{Val}(q) \leqslant 0$. For completeness, we verify this in Proposition 15 (Appendix A).

Having defined the $\mathrm{Val}(q)$ games, we are now ready to define mixed approachability.

4.5. Pure* game and mixed approachability

In order to achieve a small value in the $\mathrm{Val}(q)$ game, the forecaster may play a randomized strategy, that is, $u^\star \neq e_i$, where e_i is one of the canonical basis vectors of $\mathbb{R}^m$. On the other hand, since nature goes second, she has no incentive to randomize: there will exist an optimal strategy $v^\star \in \{0,1\}$. In the following 'pure' game, the forecaster is also not allowed to randomize over his actions.

$$\mathrm{Val}^{\mathrm{p}}(q) := \min_{p \in \mathcal{X}} \max_{y \in \{0,1\}} \langle r(p,y), q \rangle - \epsilon \|q\|_\infty. \tag{11}$$

The superscript 'p' in $\mathrm{Val}^{\mathrm{p}}(\cdot)$ refers to the min-max game being over pure actions $p \in \mathcal{X}$ and $y \in \{0,1\}$. Let us refer to this as the pure game, and the game (10) as the mixed game.

Suppose the approaching player (forecaster in our case) can ensure halfspace-satisfiability using only pure actions: $\forall q,\ \mathrm{Val}^{\mathrm{p}}(q) \leqslant 0$. Mannor and Perchet (2013) showed that if this is true then then the approaching player can ensure approachability at a fast rate of $O(1/n)$. However, it is possible to achieve the fast rate even if the above condition is not true. Characterizing a situation where the fast rate is unachievable requires considering another game, whose value lies between the pure and mixed games. Define

$$\mathcal{X}^\star = \{p \in \mathcal{X} : p \in \mathrm{support}(u^\star), \text{ where } u^\star \text{ is some optimal mixed strategy for the forecaster}\};$$
$$\mathcal{Y}^\star = \{y \in \{0,1\} : y \in \mathrm{support}(v^\star), \text{ where } v^\star \text{ is some optimal mixed strategy for nature}\}.$$

Then define the pure* game and its value as follows,

$$\mathrm{Val}^\star(q) := \min_{p \in \mathcal{X}^\star} \max_{y \in \mathcal{Y}^\star} \langle r(p,y), q \rangle - \epsilon \|q\|_\infty. \tag{12}$$

It can be shown that for any q, $\mathrm{Val}(q) \leqslant \mathrm{Val}^\star(q) \leqslant \mathrm{Val}^{\mathrm{p}}(q)$ (Mannor and Perchet, 2013). We now define mixed approachability.

Definition 12 *An approachable set is said to be mixed approachable if there exists a $q \in \mathbb{R}^m$ such that while $Val(q) = 0$, $Val^\star(q) > 0$.*

The following lemma witnesses a q that shows that the mixed approachability condition required by Theorem 8 is satisfied. The witnessed q in the proof is identified based on case (d) in the proof of Proposition 15.

Lemma 13 *There exists a $q \in \mathbb{R}^m$ such that $Val^\star(q) > Val(q) = 0$. Thus B_ϵ is mixed approachable.*

The proof of Lemma 13 is deferred to Appendix A.4 due to space constraints. We briefly note that the witnessed q is motivated by case (d) in the proof of Proposition 15, wherein the optimal strategy for the forecaster is to randomize.

4.6. Relationship to previous lower bounds for calibration

Qiao and Valiant (2021) recently constructed a strategy of nature that ensures that the calibration error of any forecaster playing Calibration-Game-I satisfies $\mathbb{E}\left[\mathrm{CE}_T\right] = \Omega(T^{-0.472})$. This bound is interesting on its own and neither weaker nor stronger than the bound we

show in Theorem 7. By studying $\mathbb{E}[\epsilon\text{-CE}_T]$, we allow the forecaster a slack of ϵ in his calibration error, which is standard in several earlier cited works (see Section 1.2), and may be sufficient in practice given that the forecasts are themselves on a 2ϵ-grid.

Qiao and Valiant (2021) also noted that $\mathbb{E}[\text{CE}_T] = \Omega(T^{-0.5})$ can be forced using a *Bernoulli strategy*: at each time step, nature plays $y_t \sim \text{Bernoulli}(p)$ for some fixed $p \in [0,1]$ unknown to the forecaster. However, in Appendix B, we provide initial (but not conclusive) evidence that the Bernoulli strategy seems insufficient to guarantee $\mathbb{E}[\epsilon\text{-CE}_T] = \Omega(T^{-0.5})$. We construct an ϵ-calibrated forecaster that satisfies $\mathbb{E}[\text{CE}_T - \epsilon] \leqslant O(\text{poly}(\log T)/T)$ for the Bernoulli strategy ($\text{poly}(\log T)$ corresponds to polynomial terms in $\log(T)$). We conjecture that the stronger statement $\mathbb{E}[\epsilon\text{-CE}_T] = \mathbb{E}[\max(\text{CE}_T - \epsilon, 0)] \leqslant O(\text{poly}(\log T)/T)$ also holds, which would mean that the Bernoulli strategy is insufficient to derive a $\Omega(1/\sqrt{T})$ bound on the ϵ-calibration rate (as shown by Theorem 7).

5. Summary

This paper connects three rich areas of the literature in a natural way: online calibration, the power of two choices, and imprecise probability. In summary, we show that by allowing the forecaster to output a deterministic short interval of probabilities (of length at most 2ϵ), we can achieve a faster rate of $O(1/T)$ for ϵ-calibration against a fully adaptive adversarial reality who presents the binary outcome after observing the interval forecast. This should be compared to the $\Theta(1/\sqrt{T})$ rate achievable with randomized point forecasts (the upper bound is a seminal result by Foster and Vohra (1998), the lower bound is ours), or the $\Theta(1)$ for deterministic point forecasts.

Arguably, such narrow intervals are as practically interpretable as point forecasts, and since some sort of binning anyway underlies most calibration algorithms, it also feels natural to allow the forecaster to express their uncertainty in this fashion, especially since it avoids randomization and improves calibration. Thus, we view our work as a theoretical contribution with clear practical implications.

Several open questions remain, since we open a rather new line of investigation. We mention two: (a) lower bounds for our setting are unknown, and (b) we don't know if models providing more than two choices could possibly improve the rate further. We suspect that $1/T$ is the optimal rate since it corresponds to constant cumulative calibration error (without normalization by T), which is incurred at $t = 1$ itself and seems unavoidable. Finally, it would be interesting to (c) figure out multidimensional analogs of our paper. We leave these for future work.

We also note some POTC-style results in online learning. Neu and Zhivotovskiy (2020) show that for expert-based classification, providing the learner the choice to abstain from making a prediction improves the regret from $\Omega(1/\sqrt{T})$ to $O(1/T)$, similar to what we obtain in Theorem 2. In zero-order convex optimization or bandit convex optimization, allowing the learner two function evaluations enables the unknown gradient to be estimated using finite difference, leading to significantly improved rates (Agarwal et al. (2010) and follow-up work). For example, in the strongly convex case the rate improves from $\Omega(\sqrt{T})$ to $O(\log T)$. Such improvements also hold for non-smooth functions (Shamir, 2017). It would be interesting to consider POTC setups for multi-armed bandits (two arm-pulls instead of one) or expert-based online learning (choosing two experts instead of randomizing or choosing one expert).

Acknowledgements

CG would like to thank the Bloomberg Data Science Ph.D. Fellowship for their generous support. We are grateful to Yusha Liu, Mingda Qiao, and the anonymous COLT reviewers, for feedback that helped improve a previous version of the paper.

References

Jacob Abernethy, Peter L Bartlett, and Elad Hazan. Blackwell approachability and no-regret learning are equivalent. In *Conference on Learning Theory*, 2011.

Alekh Agarwal, Ofer Dekel, and Lin Xiao. Optimal algorithms for online convex optimization with multi-point bandit feedback. In *Conference on Learning Theory*, 2010.

Yossi Azar, Andrei Z Broder, Anna R Karlin, and Eli Upfal. Balanced allocations. In *Symposium on Theory of Computing*, 1994.

David Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8, 1956.

Glenn W Brier. Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78(1):1–3, 1950.

Andrei Broder and Michael Mitzenmacher. Using multiple hash functions to improve IP lookups. In *IEEE International Conference on Computer Communications*, 2001.

Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.

Richard Cole, Bruce M Maggs, Friedhelm Meyer auf der Heide, Michael Mitzenmacher, Andréa W Richa, Klaus Schröder, Ramesh K Sitaraman, and Berthold Vöcking. Randomized protocols for low-congestion circuit routing in multistage interconnection networks. In *Symposium on Theory of Computing*, 1998.

A Philip Dawid. The well-calibrated Bayesian. *Journal of the American Statistical Association*, 77(379):605–610, 1982.

A Philip Dawid. Comment: The impossibility of inductive inference. *Journal of the American Statistical Association*, 80(390):340–341, 1985.

Gert de Cooman and Jasper De Bock. Randomness is inherently imprecise. *International Journal of Approximate Reasoning*, 141:28–68, 2022.

Raaz Dwivedi, Ohad N Feldheim, Ori Gurel-Gurevich, and Aaditya Ramdas. The power of online thinning in reducing discrepancy. *Probability Theory and Related Fields*, 174(1): 103–131, 2019.

Dean P Foster. A proof of calibration via Blackwell's approachability theorem. *Games and Economic Behavior*, 29(1-2):73–78, 1999.

Dean P Foster and Sergiu Hart. Smooth calibration, leaky forecasts, finite recall, and nash dynamics. *Games and Economic Behavior*, 109:271–293, 2018.

Dean P Foster and Sergiu Hart. Forecast hedging and calibration. *Journal of Political Economy*, 129(12):3447–3490, 2021.

Dean P Foster and Rakesh V Vohra. Asymptotic calibration. *Biometrika*, 85(2):379–390, 1998.

Drew Fudenberg and David K Levine. An easier way to calibrate. *Games and economic behavior*, 29(1-2):131–137, 1999.

Chirag Gupta and Aaditya Ramdas. Distribution-free calibration guarantees for histogram binning without sample splitting. In *International Conference on Machine Learning*, 2021.

Chirag Gupta and Aaditya Ramdas. Top-label calibration and multiclass-to-binary reductions. In *International Conference on Learning Representations*, 2022.

Chirag Gupta, Aleksandr Podkopaev, and Aaditya Ramdas. Distribution-free binary classification: prediction sets, confidence intervals and calibration. In *Advances in Neural Information Processing Systems*, 2020.

Sergiu Hart and Andreu Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

Steven R Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 49(2): 1055–1080, 2021.

Sham M Kakade and Dean P Foster. Deterministic calibration and nash equilibrium. In *Conference on Learning Theory*, 2004.

Ehud Lehrer. Any inspection is manipulable. *Econometrica*, 69(5):1333–1347, 2001.

Shie Mannor and Vianney Perchet. Approachability, fast and slow. In *Conference on Learning Theory*, 2013.

Shie Mannor and Gilles Stoltz. A geometric proof of calibration. *Mathematics of Operations Research*, 35(4):721–727, 2010.

Daniel McFadden. Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, pages 105–142, 1974.

Michael Mitzenmacher. *The power of two choices in randomized load balancing*. PhD thesis, University of California, Berkeley, 1996.

Michael Mitzenmacher, Andréa W Richa, and Ramesh Sitaraman. The power of two random choices: A survey of techniques and results. *Combinatorial Optimization*, 9:255–304, 2001.

Gergely Neu and Nikita Zhivotovskiy. Fast rates for online prediction with abstention. In *Conference on Learning Theory*, 2020.

David Oakes. Self-calibrating priors do not exist. *Journal of the American Statistical Association*, 80(390):339–339, 1985.

Vianney Perchet. Exponential weight approachability, applications to calibration and regret minimization. *Dynamic Games and Applications*, 5(1):136–153, 2015.

Mingda Qiao and Gregory Valiant. Stronger calibration lower bounds via sidestepping. In *Symposium on Theory of Computing*, 2021.

Alexander Rakhlin, Karthik Sridharan, and Ambuj Tewari. Online learning: Beyond regret. In *Conference on Learning Theory*, 2011.

Alvaro Sandroni, Rann Smorodinsky, and Rakesh V Vohra. Calibration with many checking rules. *Mathematics of operations Research*, 28(1):141–153, 2003.

Glenn Shafer and Vladimir Vovk. *Game-Theoretic Foundations for Probability and Finance*. John Wiley & Sons, 2019.

Ohad Shamir. An optimal algorithm for bandit and zero-order convex optimization with two-point feedback. *The Journal of Machine Learning Research*, 18(1):1703–1713, 2017.

David J Spiegelhalter. Probabilistic prediction in patient management and clinical trials. *Statistics in Medicine*, 5(5):421–433, 1986.

Vladimir Vovk and Ivan Petej. Venn-abers predictors. In *Conference on Uncertainty in Artificial Intelligence*, 2014.

Vladimir Vovk, Glenn Shafer, and Ilia Nouretdinov. Self-calibrating probability forecasting. In *Advances in Neural Information Processing Systems*, 2003.

Vladimir Vovk, Akimichi Takemura, and Glenn Shafer. Defensive forecasting. In *International Workshop on Artificial Intelligence and Statistics*, 2005.

Peter Walley and Terrence L Fine. Towards a frequentist theory of upper and lower probability. *The Annals of Statistics*, 10(3):741–761, 1982.

Appendix A. Supplementary results and deferred proofs

A.1. Lemma 14 with proof

Lemma 14 *In POTC-Cal, for any $t \geqslant 1$, if condition A is not satisfied, then condition B must be satisfied.*

Proof Note that for all t, $d_1^t \leqslant 0$ and $e_m^t \leqslant 0$, since $l_1 = 0$ and $r_m = 1$ (there cannot be a deficit for interval 1 or an excess for interval m). If condition A does not hold for $i = m$, $d_m^t > 0$. Since $d_1^t \leqslant 0$ and $d_m^t > 0$, there exists an $i \in [m-1]$ such that, $d_i^t \leqslant 0$ and $d_{i+1}^t > 0$. If condition A does not hold, then $d_i^t \leqslant 0$ implies $e_i^t > 0$. Thus we have that $d_{i+1}^t > 0$ and $e_i^t > 0$; in other words, condition B holds at the exhibited i. ■

A.2. Proposition 15 with proof

Proposition 15 *The forecaster in Calibration-Game-I can ensure halfspace-satisfiability. That is, for every $q \in \mathbb{R}^m$, $Val(q) \leqslant 0$.*

Proof Our construction is directly inspired by the proof of calibration by Foster (1999). We perform a case analysis for different values of q:

(a) Suppose any $q_i = 0$. Then, playing $u_i = 1$ and $u_{j\neq i} = 0$ gives the objective value of $-\epsilon\|q\|_\infty \leqslant 0$ irrespective of the value of v.

(b) Suppose $q_1 > 0$. Then, playing $u_1 = 1$ and $u_{j>1} = 0$ gives the objective value as $q_1(M_1 - v) - \epsilon\|q\|_\infty \leqslant q_1\epsilon - \epsilon\|q\|_\infty \leqslant 0$ (note that $M_1 = \epsilon$ and $v \geqslant 0$).

(c) Suppose $q_m < 0$. Then, playing $u_m = 1$ and $u_{j<m} = 0$ gives the objective value as $q_m(M_m - v) - \epsilon\|q\|_\infty \leqslant |q_m|\,\epsilon - \epsilon\|q\|_\infty \leqslant 0$ (note that $M_m = 1 - \epsilon$ and $v \leqslant 1$).

(d) Suppose that neither of cases (a), (b), or (c) hold. Namely, $q_1 < 0$, $q_m > 0$ and $q_i \neq 0$ for any i. Let $j \in [m-1]$ be the smallest index such that $q_j < 0$ and $q_{j+1} > 0$. Then consider u given by

$$u_j = \frac{|q_{j+1}|}{|q_j| + |q_{j+1}|},\ u_{j+1} = \frac{|q_j|}{|q_j| + |q_{j+1}|},\ u_{i\notin\{j,j+1\}} = 0.$$

Observe two facts. First, $\sum_{i_1}^m u_i q_i v = (u_j q_j + u_{j+1}q_{j+1})v = 0$ since the value inside the brackets is itself equal to 0 (any play v of nature is thus rendered ineffective). Second,

$$\begin{aligned}
\sum_{i_1}^m u_i q_i M_i &= u_j q_j M_j + u_{j+1}q_{j+1}M_{j+1} \\
&= (u_j q_j + u_{j+1}q_{j+1})M_j + u_{j+1}q_{j+1}(M_{j+1} - M_j) \\
&= 0 + 2u_{j+1}q_{j+1}\epsilon \\
&= \frac{2\,|q_j|\,|q_{j+1}|\,\epsilon}{|q_j| + |q_{j+1}|} \\
&\leqslant \frac{\|q\|_\infty\,|q_j|\,\epsilon}{|q_j| + |q_{j+1}|} + \frac{\|q\|_\infty\,|q_{j+1}|\,\epsilon}{|q_j| + |q_{j+1}|} \\
&= \|q\|_\infty\epsilon.
\end{aligned}$$

Thus the overall objective value is at most 0.

The cases considered for q are exhaustive, and in each case we verified that the forecaster can guarantee that the objective value is at most 0. ■

A.3. Proof of Proposition 6

The i-th component of the expected reward vector (8) is $u_i[(1-v)M_i + v(M_i - 1)] = u_i(M_i - v)$. Suppose $v \in I_j$. Consider playing $u \in \Delta(\mathcal{X})$ given by $u_j = 1$, $u_{i\neq j} = 0$. Then $|\mathbb{E}_{p\sim u, y\sim v}[r(p,y)]| = |\sum_{i=1}^m u_i(M_i - v)| = |M_j - v|$. Since $v \in I_j$, M_j is the mid-point of I_j, and the radius of each interval is ϵ, $|M_j - v| \leqslant \epsilon$. Thus $\mathbb{E}_{u,v}[r(p,y)] \in B_\epsilon$.

A.4. Proof of Proposition 13

Set $q_{1:m-1} = -1$ (i.e. $q_i = -1$ for all $i \in [m-1]$) and $q_m = 1$. Let us compute $\mathrm{Val}(q)$. The game for nature reduces to maximizing $(\sum_{i=1}^{m-1} u_i - u_m)v$ which can be done by playing $v = \mathbf{1}\{u_m \leqslant \sum_{i=1}^{m-1} u_i\}$. We perform case work to identify the optimal play for the forecaster.

- If $\sum_{i=1}^{m-1} u_i \leqslant u_m$, $v(u_m - \sum_{i=1}^{m-1} u_i) = 0$. The objective for the forecaster reduces to:

$$\min_{u \in \Delta(\mathcal{X}), \sum_{i=1}^{m-1} u_i \leqslant u_m} u_m M_m - \sum_{i=1}^{m-1} u_i M_i - \epsilon.$$

  Note that $0 < M_1 < M_2 < \ldots < M_m$. Thus the forecaster would set the minimum possible value of u_m, which under the constraints is equal to 0.5. For the second term, the largest coefficient of a u_i in the summation is M_{m-1}. Thus in order to minimize, the forecaster would set the maximum possible value of u_{m-1}, which under the constraints is $1 - u_m = 0.5$. We conclude that the minimum occurs at $u_m = u_{m-1} = 0.5$, $u_{i \notin \{m-1,m\}} = 0$. The objective value is equal to $0.5(M_m - M_{m-1}) - \epsilon = 0$.

- If $\sum_{i=1}^{m-1} u_i \geqslant u_m$, $v = 1$ is an optimal play for nature. The objective for the forecaster reduces to:

$$\min_{u \in \Delta(\mathcal{X}), \sum_{i=1}^{m-1} u_i \geqslant u_m} u_m(M_m - 1) - \sum_{i=1}^{m-1} u_i(M_i - 1) - \epsilon.$$

  As in the other case, this game is solved by observing that since $M_1 < M_2 < \ldots < M_m < 1$, the forecaster would want to put the maximum possible value on u_m which is 0.5 under the constraints. Among $u_{1:m-1}$, the multiplier of $(M_{m-1} - 1)$ hurts the least. Thus the forecaster sets $u_{m-1} = 0.5$ and $u_{i \notin \{m-1,m\}}$. The objective value at the minimum is equal to 0.

In each case, the forecaster's optimal play is $u_m = u_{m-1} = 0.5$, $u_{i \notin \{m-1,m\}} = 0$. This essentially makes nature's action irrelevant; nature's optimal response is any $v \in [0,1]$. Thus we have shown that $\mathrm{Val}(q) = 0$, $\mathcal{X}^\star = \{M_{m-1}, M_m\}$, and $\mathcal{Y}^\star = \{0,1\}$. Let us now compute

$$\mathrm{Val}^\star(q) = \min_{p \in \mathcal{X}^\star} \max_{y \in \mathcal{Y}^\star} \left(\mathbf{1}\{p = M_m\}(M_m - y) - \mathbf{1}\{p = M_{m-1}\}(M_{m-1} - y) - \epsilon\right).$$

If the forecast is $p = M_m$, nature can respond $y = 0$ to achieve an overall objective of $1 - 2\epsilon > 0$. (We have assumed $m \geqslant 2$ so that $\epsilon < 0.5$.) If the forecast is $p = M_{m=1}$, nature can respond $y = 1$ to achieve an overall objective of $2\epsilon > 0$. Thus, $\mathrm{Val}^\star(q) > 0$.

Appendix B. Bernoulli strategy seems insufficient to prove an $\Omega(1/\sqrt{T})$ lower bound on the ϵ-calibration rate

This section is in the setup of Calibration-Game-I. Nature plays the 'Bernoulli strategy': for a fixed $p \in [0,1]$ (unknown to the forecaster), nature plays $y_t \sim \mathrm{Bernoulli}(p)$ each time. We describe a strategy for the forecaster that satisfies $\mathbb{E}\left[\epsilon\text{-CE}_T\right] = O(1/\sqrt{T})$ against a

PI-F99: pre-initialized version of the ϵ-calibrated strategy by Foster (1999)

Fix $T_0 \in \mathbb{N}$. Set $T_k := 2^k T_0$, K_k as in (13), $T^{(0)} := 0$, and $T^{(k)} := \sum_{j=1}^{k-1} T_j = (2^k - 1)T_0$.

For each time $t = 1, 2, \ldots$

- Identify smallest $k \in \mathbb{N}$ such that $t \leqslant T^{(k)}$.
- Play PI-F99(T_{k-1}) based on observations from $T^{(k-1)} + 1$ until t:
  - if $t \leqslant T^{(k-1)} + mK_k$ (initialization phase):
    identify j such that $t - T^{(k-1)} \in ((j-1)K_k + jK_k]$ and predict M_j.
  - if $t > T^{(k-1)} + mK_k$:
    follow Foster's strategy based on observations from $T^{(k-1)} + 1$ until t.

general strategy of nature; however, if nature is playing the Bernoulli strategy (instead of an arbitrary strategy), then our proposed strategy satisfies $\mathbb{E}\left[\mathrm{CE}_T - \epsilon\right] \leqslant O(\mathrm{poly}(\log T)/T)$. Our result stops short of proving the following stronger statement against the Bernoulli strategy: $\mathbb{E}\left[\epsilon\text{-}\mathrm{CE}_T\right] = \mathbb{E}\left[\max(\mathrm{CE}_T - \epsilon, 0)\right] \leqslant O(\mathrm{poly}(\log T)/T)$. We conjecture that this stronger statement holds as well, meaning that the Bernoulli strategy is insufficient to derive a $\Omega(1/\sqrt{T})$ bound on the ϵ-calibration rate as shown in Theorem 7.

B.1. The strategy

The strategy we propose is a pre-initialized version of the ϵ-calibration strategy of Foster (1999)—we call it PI-F99 (for pre-initialized-Foster-99). PI-F99 relies on a few constants: a large enough *doubling* horizon $T_0 \in \mathbb{N}$, *doubled* versions of T_0, namely $T_k =: 2^k T_0$ for $k \in \mathbb{N}$, and an initialization parameter

$$K_k := \left\lceil (0.85 \log T_k/\epsilon)^2 (\log\log(T_k/2) + 0.72 \log(5.2 m T_k^2)) \right\rceil \tag{13}$$

defined for each k. K_k is the sufficient number of samples required to estimate the bias of m Bernoulli random variables simultaneously and uniformly across time to a certain degree of reliability; further details become clear when analyzing. The constants in the definition of K_k are not crucial (looser constants still lead to the same asymptotic dependence on T), but we identify them nevertheless in order to be precise.

PI-F99 is a concatenation of certain sub-strategies PI-F99(T_k) for $k \in \mathbb{N}$, each of which are strategies for Calibration-Game-I assuming the game only goes on until time $t = T_k$. The forecaster playing PI-F99(T_k) first forecasts $p_t = M_i$, K_k times each, for each $i \in [m]$ (thus until time $t \leqslant mK_k$). This is the *initialization* phase. Then, for $t \in \{mK_k + 1, \ldots, T_k\}$, the forecaster follows Foster's strategy initialized with *current empirical frequencies* for each bin, based on what has been observed so far in the initialization phase.

The actual strategy of the forecaster corresponds to a concatenation of PI-F99(T_0), PI-F99(T_1), PI-F99(T_2), and so on. This is a version of the doubling trick (Cesa-Bianchi and Lugosi, 2006). The forecaster first plays PI-F99(T_0) from $t = 1$ to $t = T_0$, then plays

PI-F99(T_1) from $t = T_0 + 1$ to $t = T_0 + T_1$, then PI-F99(T_2) and so on. To be clear, when switching from PI-F99(T_{k-1}) to PI-F99(T_k), the forecaster completely ignores the forecasts and observations so far, and restarts. The overall strategy is described in the box on top of the previous page.

B.2. Analysis

Ignoring terms in m and ϵ which are constants in T_k, $mK_k = O(\log^2 T_k) \ll \sqrt{T_k}$ (for sufficiently large T_0 and all $k \geqslant 0$). Assuming a worst case error of 1 for each time until $T \leqslant mK_k$, we can show that the ϵ-calibration error of PI-F99(T_k) at any time $T \leqslant T_k$ satisfies the following:

$$\begin{aligned}
&\mathbb{E}\left[\max\left(\sum_{i=1}^{m}\left|\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}\{p_t = M_i\}(M_i - y_t)\right| - \epsilon, 0\right)\right] \\
\leqslant\ &\mathbb{E}\left[\max\left(\sum_{i=1}^{m}\left|\frac{1}{T}\sum_{t=1}^{mK_k}\mathbf{1}\{p_t = M_i\}(M_i - y_t)\right| + \sum_{i=1}^{m}\left|\frac{1}{T}\sum_{t=mK_k+1}^{T}\mathbf{1}\{p_t = M_i\}(M_i - y_t)\right| - \epsilon, 0\right)\right] \\
\leqslant\ &\frac{mK_k}{T} + \mathbb{E}\left[\max\left(\sum_{i=1}^{m}\left|\frac{1}{T}\sum_{t=mK_k+1}^{T}\mathbf{1}\{p_t = M_i\}(M_i - y_t)\right| - \epsilon, 0\right)\right] \\
\leqslant\ &\frac{O(\sqrt{T_k})}{T} + \mathbb{E}\left[\max\left(\sum_{i=1}^{m}\left|\frac{1}{T}\sum_{t=mK_k+1}^{T}\mathbf{1}\{p_t = M_i\}(M_i - y_t)\right| - \epsilon, 0\right)\right].
\end{aligned}$$

To bound the second term, note that Foster's strategy has a calibration rate of $O(1/\sqrt{T})$ starting with any initialization. Thus,

$$\mathbb{E}\left[\max\left(\sum_{i=1}^{m}\left|\frac{1}{T}\sum_{t=mK_k+1}^{T}\mathbf{1}\{p_t = M_i\}(M_i - y_t)\right| - \epsilon, 0\right)\right] \leqslant \frac{O(\sqrt{T - mK_k})}{T} \leqslant \frac{C\sqrt{T_k}}{T}, \tag{14}$$

for some constant C independent of T_0 and k. Thus for PI-F99(T_k) at any time $T \leqslant T_k$, we have shown that $\mathbb{E}\left[\epsilon\text{-CE}_T\right] \leqslant O(1/\sqrt{T})$. From this, it follows that the overall strategy PI-F99 satisfies $f(T) = O(1/\sqrt{T})$ asymptotically. This is shown in Proposition 19, later in this subsection. Next, we perform the analysis for the Bernoulli strategy.

Proposition 16 *PI-F99 satisfies $\mathbb{E}\left[CE_T\right] \leqslant \epsilon + O(poly(\log T)/T)$ if nature follows the Bernoulli strategy.*

Proof Following the notation of Section 3, let N_i^T denote the number of times the mid-point M_i is forecasted until time T. We have,

$$\mathbb{E}\left[\sum_{i=1}^{m}\left|\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}\{p_t = M_i\}(M_i - y_t)\right|\right] = \underbrace{\frac{\sum_{i\in[m]:|p-M_i|>\epsilon}\mathbb{E}\left[N_i^T\left|M_i - \sum_{t=1}^{T}\mathbf{1}\{p_t = M_i\}y_t/N_i^T\right|\right]}{T}}_{E_1} + \underbrace{\frac{\sum_{i\in[m]:|p-M_i|\leqslant\epsilon}\mathbb{E}\left[N_i^T\left|M_i - \sum_{t=1}^{T}\mathbf{1}\{p_t = M_i\}y_t/N_i^T\right|\right]}{T}}_{E_2}.$$

We will show that

$$E_1 = O(\text{poly}(\log T)/T),$$

and

$$E_2 = \epsilon + O(1/T),$$

which will complete the argument. To this end, define A^T as the number of times that the forecast is ϵ-close to p, until time T:

$$A^T := \sum_{i \in [m]: |p - M_i| \leqslant \epsilon} N_i^T. \tag{15}$$

Lemma 17 shows that PI-F99 satisfies $\mathbb{E}\left[A^T\right] = T - O(\text{poly}(\log T))$. This immediately leads to the bound for E_1; note that $\left|M_i - \sum_{t=1}^T \mathbf{1}\{p_t = M_i\} y_t / N_i^T\right| \leqslant 1$, and thus

$$E_1 \leqslant \frac{\sum_{i \in [m]: |p - M_i| > \epsilon} \mathbb{E}\left[N_i^T\right]}{T} = 1 - \frac{\mathbb{E}\left[A^T\right]}{T} = O(\text{poly}(\log T)/T).$$

Bounding E_2 takes more work. The proof relies on the following 'good' event occurring with high probability:

$$G \equiv G_T : \text{for every } i \in [m], \left|p - \sum_{t=1}^T \mathbf{1}\{p_t = M_i\} y_t / N_i^T\right| \leqslant \epsilon/2.$$

Due to the pre-initialization steps in PI-F99, it can be guaranteed that $\Pr(G) = \Pr(G_T) = 1 - O(1/T)$. For the details, we refer the reader to the proof of Lemma 17 (see case (a) in the proof), where we show a stronger version of this fact (namely, with $\epsilon/\log T_k$ instead of $\epsilon/2$), for PI-F99(T_k) using a time-uniform concentration inequality (due to the time uniformity, the implication holds for PI-F99 as well). We now do case work to bound E_2.

(a) Suppose there exists an index $j \in [m]$ such that $|p - M_j| \leqslant \epsilon/2$. This index must be unique since the M_j's are 2ϵ apart. Further, no $i \neq j$ can satisfy $|p - M_i| \leqslant \epsilon$. We now obtain the following (below, we use abs$(\cdot)$ instead of $|\cdot|$ to avoid confusion with the conditioning operator):

$$\begin{aligned}
T \cdot E_2 &= \mathbb{E}\left[N_j^T \text{abs}(M_j - \sum_{t=1}^T \mathbf{1}\{p_t = M_j\} y_t / N_j^T)\right] \\
&\leqslant \mathbb{E}\left[N_j^T \text{abs}(M_j - \sum_{t=1}^T \mathbf{1}\{p_t = M_j\} y_t / N_j^T) \mid G\right] + (1 - \Pr(G)) \cdot T \\
&= \mathbb{E}\left[N_j^T \text{abs}(M_j - \sum_{t=1}^T \mathbf{1}\{p_t = M_j\} y_t / N_j^T) \mid G\right] + O(1) \\
&\leqslant T \cdot \mathbb{E}\left[\text{abs}(M_j - \sum_{t=1}^T \mathbf{1}\{p_t = M_j\} y_t / N_j^T) \mid G\right] + O(1) \\
&\leqslant T \cdot \mathbb{E}\left[|M_j - p| + \text{abs}(p - \sum_{t=1}^T \mathbf{1}\{p_t = M_i\} y_t / N_j^T) \mid G\right] + O(1) \\
&\leqslant T(\epsilon/2 + \epsilon/2) + O(1) = T\epsilon + O(1),
\end{aligned}$$

where the inequality in the last line follows by the case assumption and the definition of G. Thus for this case, we have shown that $E_2 = \epsilon + O(1/T)$.

(b) Suppose p is such that for every i, $|p - M_i| > \epsilon/2$. To bound E_2, we are interested in the M_i's for which $|p - M_i| \leqslant \epsilon$. There can be at most two such M_i's: an M_j that satisfies $M_j \in (p + \epsilon/2, p + \epsilon]$, and an M_l that satisfies $M_j \in [p - \epsilon, p - \epsilon/2)$.

Suppose there is an M_j satisfying $M_j \in (p+\epsilon/2, p+\epsilon]$. Set $R = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{p_t = M_j\}(M_j - y_t)$ and note that $R \in [-1, 1]$. By Lemma 18, $\mathbb{E}[|R|] \leqslant \mathbb{E}[R] + 2 \cdot \Pr(R < 0)$. Note that $\Pr(R \geqslant 0) \geqslant \Pr(G)$, since if G holds,

$$\begin{aligned}
R \cdot T &= \sum_{t=1}^{T} \mathbf{1}\{p_t = M_j\}(M_j - y_t) \\
&= \sum_{t=1}^{T} \mathbf{1}\{p_t = M_j\}((M_j - p) + (p - y_t)) \\
&\geqslant \sum_{t=1}^{T} \mathbf{1}\{p_t = M_j\}(\epsilon/2 + (p - y_t)) \\
&= N_i^T \epsilon/2 + \sum_{t=1}^{T} \mathbf{1}\{p_t = M_j\}(p - y_t) \\
&\geqslant N_i^T \epsilon/2 - \left|\sum_{t=1}^{T} \mathbf{1}\{p_t = M_j\}(p - y_t)\right| \\
&\geqslant N_i^T \epsilon/2 - N_i^T \epsilon/2 = 0,
\end{aligned}$$

where the inequality in the last line is implied by G. Thus, $\Pr(R < 0) \leqslant 1 - \Pr(G) = O(1/T)$. Next, we bound $\mathbb{E}[R]$.

$$\begin{aligned}
\mathbb{E}[R] &= \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[\mathbf{1}\{p_t = M_j\}(M_j - y_t)] \\
&= \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[\mathbf{1}\{p_t = M_j\}(M_j - \mathbb{E}[y_t \mid (y_1, \ldots, y_{t-1}), (p_1, \ldots, p_t)])] \\
&= \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}[\mathbf{1}\{p_t = M_j\}(M_j - p)] \\
&= \frac{\mathbb{E}N_j^T(M_j - p)}{T} \leqslant \frac{\mathbb{E}N_j^T}{T} \cdot \epsilon.
\end{aligned}$$

Putting it together, we obtain $\mathbb{E}[|R|] \leqslant \frac{\mathbb{E}N_j^T}{T} \cdot \epsilon + O(1/T)$.

Similarly, suppose there is an M_l satisfying $M_l \in [p - \epsilon, p - \epsilon/2)$. For this l, define $S = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{p_t = M_l\}(M_l - y_t)$. An identical argument as the one used for R goes through; we use the inequality $\mathbb{E}[|S|] \leqslant \mathbb{E}[-S] + 2 \cdot \Pr(S > 0)$ (from Lemma 18) and the relationship of $\Pr(S > 0)$ to G to obtain $\mathbb{E}[|S|] \leqslant \frac{\mathbb{E}N_l^T}{T} \cdot \epsilon + O(1/T)$.

Finally, we conclude if both M_k and M_l with the given relationship to p exist, then $E_2 \leqslant \mathbb{E}[|R| + |S|]$; if only M_j exists, then $E_2 \leqslant \mathbb{E}[|R|]$; if only M_l exists, then $E_2 \leqslant \mathbb{E}[|S|]$. In each case,

$$E_2 \leqslant \frac{\mathbb{E}\left[A^T\right] \cdot \epsilon}{T} + O(1/T) \leqslant \epsilon + O(1/T).$$

Since the two cases considered are exhaustive, this completes the proof.

■

Lemma 17 (for proving Proposition 16) *PI-F99 satisfies* $\mathbb{E}\left[A^T\right] = T - O(poly(\log T))$, *where* A^T *is defined in the proof of Proposition 16.*

Proof We first show PI-F99(T_k) satisfies $\mathbb{E}\left[A^T\right] = T - O(\text{poly}(\log T))$ for $T \leqslant T_k$ once k is large enough. We do so via two cases.

(a) For the first case, suppose $p = r_j = l_{j+1}$ for some $j \in [m-1]$, that is, the bias of the Bernoulli is exactly at the common endpoint of two intervals. In other words, $A^T = N_j^T + N_{j+1}^T$. We show that with high probability, the forecaster will *learn* this index j in the initialization phase of each PI-F99(T_k), and continue playing either M_j or M_{j+1} until he switches to PI-F99(T_{k+1}).

Consider the strategy PI-F99(T_k) for some $k \geqslant 0$. From time $t = mK_k$ onwards, each M_i has been forecasted at least K_k times, so that the value of p_i^t is *close* to p. To formalize close, we will use a time-uniform sub-Gaussian concentration inequality shown by Howard et al. (2021, equation (3.4)). We use their inequality, replacing each instance of t with $K_k/4$, since a Bernoulli is (1/4)-sub-Gaussian and each M_i has been forecasted at least K_k times. Additionally, we replace α with $1/mT_k^2$. It can be verified that the final deviation term inside the brackets is at most $\epsilon/\log T_k$; in other words, with probability at least $1 - 1/T_k^2$, the following 'good' event occurs:

$$\text{for all times } mK_k \leqslant t \leqslant T_k, \max_{i \in [m]} \left|p_i^t - p\right| \leqslant \epsilon/\log T_k \leqslant \epsilon.$$

The radius of each interval is ϵ. Thus if the above event occurs, it follows that for intervals $i < j$, the right-endpoint $r_i < p - \epsilon \leqslant p_i^t$, so we have an excess ($e_i^T > 0$) until T_k; and for intervals $i > j+1$, the left-endpoint $l_i > p + \epsilon \geqslant p_i^t$, so we have a deficit ($d_i^T > 0$) until T_k. For interval j, either both $d_j^t, e_j^t \leqslant 0$ or $e_j^t > 0$; for interval $j+1$, either both $d_j^t, e_j^t \leqslant 0$ or $d_j^t > 0$. Overall, with probability at least $1 - 1/T_k^2$, for times $mK_k < t \leqslant T_k$, Foster's algorithm randomizes between M_j and M_{j+1} (possibly playing one of them deterministically).

(b) The other case is when p belongs to the interior of some interval I_j, $j \in [m]$, or $p \in \{0, 1\}$. Then, there exists some $\delta > 0$ such that $|p - M_i| \geqslant \epsilon + \delta$ for all $i \neq j$. For a sufficiently large value of $\widetilde{k} \in \mathbb{N}$, $\delta > \epsilon/\log(T_{\widetilde{k}})$. Consider the strategy PI-F99(T_k) for $k \geqslant \widetilde{k}$. As noted in the previous case, our choice of K_k ensures that with probability at least $1 - 1/T_k^2$, for all times $mK_k \leqslant t \leqslant T_k$, $\max_{i \in [m]} \left|p_i^t - p\right| \leqslant \epsilon/\log T_k < \delta$. Using triangle inequality, we conclude that $\left|p_j^t - M_j\right| \leqslant \epsilon$ and $\left|p_i^t - M_{i \neq j}\right| > \epsilon$. It follows

that for every $i \neq j$, there is either a deficit or an excess, and for j there is neither. Thus with probability at least $1 - 1/T_k^2$, Foster's algorithm plays M_j after time mK_k.

Cases (a) and (b) lead to a lower bound on $\mathbb{E}\left[A^T\right]$ for $\texttt{PI-F99}(T_k)$, the expected number of times an M_i is forecasted that is ϵ-close to p. Namely, we obtain that for $k \geqslant \widetilde{k}$, for the strategy $\texttt{PI-F99}(T_k)$ that plays assuming a horizon of T_k from $t = 1$ itself, we have for $T \leqslant T_k$:

$$\begin{aligned}\mathbb{E}\left[A^T\right] &\geqslant (1 - 1/T_k^2)(T - mK_k) \\ &\geqslant T(1 - 1/T_k^2) - O(\text{poly}(\log T_k)) \\ &\geqslant T - O(\text{poly}(\log T)). \end{aligned} \tag{16}$$

The final inequality above holds since $T \leqslant T_k$.

We derive the implication for the overall strategy that is actually played, PI-F99. Recall the notation $T^{(0)} = 0$ and $T^{(k)} = T_0 + T_1 + \ldots + T_{k-1}$. In PI-F99, the $\texttt{PI-F99}(T_k)$ strategy is played from time $T^{(k-1)} + 1$ to time $T^{(k-1)} + T_k = T^{(k)}$. Let T be such that $T \in [T^{(k)} + 1, T^{(k+1)}]$ for any $k \geqslant \widetilde{k}$. Then by (16),

$$\mathbb{E}\left[A^T - A^{T^{(k)}}\right] \geqslant T - T^{(k)} - O(\text{poly}(\log T)).$$

Again by (16), the above holds with $T \leftarrow T^{(k)}$, $T^{(k)} \leftarrow T^{(k-1)}$, if $k \geqslant \widetilde{k} + 1$ ($\leftarrow$ corresponds to replacing the term on the left with the term on the right):

$$\mathbb{E}\left[A^{T^{(k)}} - A^{T^{(k-1)}}\right] \geqslant T^{(k)} - T^{(k-1)} - O(\text{poly}(\log T)).$$

Instantiating this recursively for all $k \geqslant \widetilde{k} + 1$, and adding the inequalities together gives us:

$$\mathbb{E}\left[A^T\right] \geqslant T - T^{(\widetilde{k})} - \log(T) \cdot O(\text{poly}(\log T)) = T - O(\text{poly}(\log T)),$$

since $\widetilde{k}$ is some fixed constant (given p). This completes the argument.

■

Lemma 18 *For any bounded random variable $R \in [-a, a]$,*

$$\mathbb{E}\left[|R|\right] \leqslant \min(\mathbb{E}\left[R\right] + 2a \cdot Pr(R < 0), \mathbb{E}\left[-R\right] + 2a \cdot Pr(R > 0)). \tag{17}$$

Proof Note that,

$$\begin{aligned}\mathbb{E}\left[|R|\right] &= \mathbb{E}\left[R \cdot \mathbf{1}\{R \geqslant 0\} - R \cdot \mathbf{1}\{R < 0\}\right] \\ &= \mathbb{E}\left[R \cdot \mathbf{1}\{R \geqslant 0\} + R \cdot \mathbf{1}\{R < 0\} - 2R \cdot \mathbf{1}\{R < 0\}\right] \\ &= \mathbb{E}\left[R - 2R \cdot \mathbf{1}\{R < 0\}\right] \\ &\leqslant \mathbb{E}\left[R + 2a \cdot \mathbf{1}\{R < 0\}\right] \\ &= \mathbb{E}\left[R\right] + 2a \cdot \Pr(R < 0). \end{aligned}$$

In the above proof, we can replace R with $-R$ everywhere, since $-R \in [-a, a]$ as well. Thus we also obtain,

$$\mathbb{E}\left[|R|\right] = \mathbb{E}\left[|-R|\right] \leqslant \mathbb{E}\left[-R\right] + 2a \cdot \Pr(R > 0).$$

■

Proposition 19 *PI-F99 achieves a calibration rate of $O(1/\sqrt{T})$ against any strategy of nature.*

Proof Define $T^{(0)} = 0$ and $T^{(k)} = T_0 + T_1 + \ldots + T_{k-1} = (2^k - 1)T_0$, for $k \geqslant 1$. Further, define the cumulative (non-normalized) calibration error corresponding only to the times $t = t_1 + 1$ to t_2 as follows:

$$\mathrm{CE}(t_1, t_2) := \mathbb{E}\left[\max\left(\sum_{i=1}^{m}\left|\sum_{t=t_1+1}^{t_2} \mathbf{1}\{p_t = M_i\}(M_i - y_t)\right|, \epsilon(t_2 - t_1)\right)\right].$$

From (14), PI-F99 satisfies, for every $k \in \mathbb{N}$ and $T^{(k-1)} < t \leqslant T^{(k)}$,

$$\mathrm{CE}(T^{(k-1)}, t) \leqslant \epsilon(t - T^{(k-1)}) + C\sqrt{T_{k-1}} \tag{18}$$

for some universal constant C that does not depend on k.

Now for a given $T > T^{(2)}$, let $k \geqslant 3$ be such that $T^{(k-1)} < T \leqslant T^{(k)}$. By triangle inequality and (18),

$$\begin{aligned}
\mathbb{E}\left[\max\left(\sum_{i=1}^{m}\left|\sum_{t=1}^{T} \mathbf{1}\{p_t = M_i\}(M_i - y_t)\right|, \epsilon T\right)\right] &\leqslant \sum_{i=1}^{k-1} \mathrm{CE}(T^{(i-1)}, T^{(i)}) + \mathrm{CE}(T^{(k-1)}, T) \\
&\leqslant \sum_{i=1}^{k-1}(\epsilon(T^{(i)} - T^{(i-1)}) + C\sqrt{T_{i-1}}) \\
&\qquad + \epsilon(T - T^{(k-1)}) + C\sqrt{T_{k-1}} \\
&= \epsilon T + \sum_{i=1}^{k} C\sqrt{T_{i-1}} \\
&= \epsilon T + C(\sqrt{T_0} + \sqrt{2T_0} + \sqrt{4T_0} + \ldots + \sqrt{2^{k-1}T_0}) \\
&\leqslant \epsilon T + C\sqrt{T_0} \cdot \frac{\sqrt{2^k} - 1}{\sqrt{2} - 1} \\
&\leqslant \epsilon T + C \cdot \frac{\sqrt{2^k T_0}}{\sqrt{2} - 1}. \\
&\leqslant \epsilon T + C'\sqrt{T},
\end{aligned}$$

where $C' = C \cdot 2/(\sqrt{2} - 1)$. The final inequality holds since for $k \geqslant 3$,

$$\sqrt{2^k T_0} \leqslant \sqrt{4(2^{k-1} - 1)T_0} = \sqrt{4T^{(k-1)}} < 2\sqrt{T}.$$

Dividing by T and taking ϵ to the left-hand-size, we get that for all $T > T^{(2)}$,

$$\mathbb{E}\left[\max\left(\sum_{i=1}^{m}\left|\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}\{p_t = M_i\}(M_i - y_t)\right| - \epsilon, 0\right)\right] \leqslant C'\sqrt{T} = O(1/\sqrt{T}),$$

as needed. ■

Appendix C. Generalization of POTC-Cal to bounded outputs

If the output is bounded instead of binary (see Remark 1), then POTC-Cal can be modified as follows. The forecaster maintains p_i^T as in the original algorithm, but these are now the mean of the v_t values instead of y_t values. The choice of the index i and the final forecast (p_{t0}, p_{t1}) is made identically to the original POTC-Cal. Finally, the forecaster plays

$$p_t = p_{t0} \text{ if } v_t \leqslant r_i, \text{ and } p_t = p_{t1} \text{ if } v_t > r_i. \tag{19}$$

Note that $r_i = l_{i+1}$ is the right (left) endpoint of interval i (interval $i+1$), and thus a natural threshold for deciding which of the two intervals to play.

Lemmas 3 and 4 hold for this modified setup and algorithm, and thus the $O(1/T)$ rate showed by Theorem 2 also holds. Lemma 3 goes through since the set of equations (6) can be modified as follows:

$$\left|d_i^{t+1} - d_i^t\right| = \left|e_i^{t+1} - e_i^t\right| = \left|\frac{v_{t+1} - p_i^t}{N_i^t + 1}\right| \leqslant \frac{1}{N_i^{t+1}}.$$

In the proof of Lemma 4, we assumed without loss of generality that $y_{t+1} = 0$. This assumption can be modified to $v_{t+1} \leqslant r_i$ in keeping with the forecaster's updated strategy (19). The case $e_i^{t+1} < d_i^{t+1}$ goes through since it is a consequence of the set of equations (6). For the case $e_i^{t+1} \geqslant d_i^{t+1}$, we have

$$\begin{aligned} N_i^{t+1}\max(d_i^{t+1}, e_i^{t+1}) &= N_i^t p_i^t + v_{t+1} - N_i^{t+1} r_i \\ &= N_i^t e_i^t + (v_{t+1} - r_i) \leqslant N_i^t \max(d_i^t, e_i^t), \end{aligned}$$

where the last inequality follows by the case assumption $v_{t+1} \leqslant r_i$.