Games and Economic Behavior 109 (2018) 271–293

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Smooth calibration, leaky forecasts, finite recall, and Nash dynamics ☆

Dean P. Foster [a,b], Sergiu Hart [c,*]

[a] Amazon Inc., New York City, United States
[b] University of Pennsylvania, Philadelphia, United States
[c] Institute of Mathematics, Department of Economics, and Center for the Study of Rationality, The Hebrew University of Jerusalem, Israel

ARTICLE INFO

Article history:
Received 30 March 2017
Available online 23 March 2018

JEL classification:
C7
D8
C1

Keywords:
Calibration
Nash dynamics
Fixed points
Deterministic calibration
Smooth calibration
Finite recall

ABSTRACT

We propose to smooth out the calibration score, which measures how good a forecaster is, by combining nearby forecasts. While regular calibration can be guaranteed only by randomized forecasting procedures, we show that *smooth calibration* can be guaranteed by *deterministic* procedures. As a consequence, it does not matter if the forecasts are *leaked*, i.e., made known in advance: smooth calibration can nevertheless be guaranteed (while regular calibration cannot). Moreover, our procedure has finite recall, is stationary, and all forecasts lie on a finite grid. To construct the procedure, we deal also with the related setups of online linear regression and weak calibration. Finally, we show that smooth calibration yields uncoupled finite-memory *dynamics* in n-person games—"smooth calibrated learning"—in which the players play approximate *Nash equilibria* in almost all periods (by contrast, calibrated learning, which uses regular calibration, yields only that the time averages of play are approximate correlated equilibria).

1. Introduction

How good is a forecaster? Assume for concreteness that every day the forecaster issues a forecast of the type "the chance of rain tomorrow is 30%." A simple test one may conduct is to calculate the proportion of rainy days out of those days for which the forecast was 30%, and compare it to 30%; and do the same for all other forecasts. A forecaster is said to be *calibrated* if, in the long run, the differences between the actual proportions of rainy days and the forecasts are small—no matter what the weather really was (see Dawid, 1982).

What if rain is replaced by an event that is under the control of another agent? If the forecasts are made public before the agent decides on his action—we refer to this setup as "*leaky forecasts*"—then calibration *cannot* be guaranteed; for example, the agent can make the event happen if and only if the forecast is less than 50%, and so the forecasting error (that is, the "calibration score") is always at least 50%. However, if in each period the forecast and the agent's decision are made "simultaneously"—which means that neither one knows the other's decision before making his own—then calibration

☆ Previous versions: July 2012, February 2015, March 2017. Research of the second author was partially supported by a European Research Council (ERC) Advanced Investigator grant. The authors thank Yakov Babichenko for useful comments, and the editor, associate editor, and referees for their very careful reading and helpful suggestions.

* Corresponding author.
E-mail addresses: dean@foster.net (D.P. Foster), hart@huji.ac.il (S. Hart).
URLs: http://deanfoster.net/ (D.P. Foster), http://www.ma.huji.ac.il/hart (S. Hart).

https://doi.org/10.1016/j.geb.2017.12.022
0899-8256/

can be guaranteed; see Foster and Vohra (1998). The procedure that yields calibration no matter what the agent's decisions are requires the use of *randomizations* (e.g., with probability 1/2 the forecaster announces 30%, and with probability 1/2 he announces 60%). Indeed, as the discussion at the beginning of this paragraph suggests, one cannot have a deterministic procedure that is calibrated (see Dawid, 1985 and Oakes, 1985).

Now the standard calibration score is overly fastidious: the days when the forecast was, say, 30.01% are considered separately from the days when the forecast was 29.99% (formally, the calibration score is a highly discontinuous function of the data, i.e., the forecasts and the actions). This suggests that one first combines all days when the forecast was *close to* 30%, and only then compares the 30% with the proportion of rainy days. If, say, there were 200 days with a forecast of 30.01%, out of which 10 were rainy, and another 100 days with a forecast of 29.99%, out of which 80 were rainy, then the forecaster is very far from being calibrated; however, he is smoothly calibrated, as his forecasts were all close to 30%, and there were $90/300 = 30\%$ rainy days. Undershooting at 29.99% and overshooting at 30.01% is now balanced out. Formally, what this amounts to is applying a so-called "smoothing" operation to the forecasting errors (which makes smooth calibration easier to obtain than calibration).[1]

Perhaps surprisingly, once we consider smooth calibration, there is no longer a need for randomization when making the forecasts: we will show that there exist *deterministic* procedures that guarantee smooth calibration, no matter what the agent does. In particular, it follows that it does not matter if the forecasts are made known to the agent before his decision, and so smooth calibration can be guaranteed even when forecasts may be leaked.[2] This may come as a surprise, because, as pointed out above, an agent who knows the forecast *before* deciding on the weather will choose rain when the forecast is less than 50% and no rain otherwise, giving a calibration error of 50% or more, no matter what the forecaster does. However, against such an agent one can easily be *smoothly* calibrated, by forecasting 50.01% on odd days and 49.99% on even days (the resulting weather will then alternate between rain and no rain, and so half the days will be rainy days—and all the forecasts are indeed close to 50%). What this proves is only that one can be smoothly calibrated against this specific strategy of the agent (this is the strategy that shows that it is impossible to have calibration with deterministic leaky procedures); our result shows that one can in fact *guarantee* smooth calibration with a deterministic strategy, against *any* strategy of the agent.

The forecasting procedure that we construct and that guarantees smooth calibration has moreover finite recall (i.e., only the forecasts and actions of the last R periods are taken into account, for some fixed finite R), and is stationary (i.e., independent of "calendar time": the forecast is the same any time that the "window" of the past R periods is the same).[3] Finally, we can have all the forecasts lie on some finite fixed grid.

The construction starts with the "online linear regression" problem, introduced by Foster (1991), where one wants to generate every period a good linear estimator based only on the data up to that point. We provide a finite-recall stationary algorithm for this problem; see Section 3. We then use this algorithm, together with a fixed-point argument, to obtain "weak calibration", a concept introduced by Kakade and Foster (2004) and Foster and Kakade (2006); see Section 4. Section 5 shows that weak and smooth calibration are essentially equivalent, which yields the existence of smoothly calibrated procedures. Finally, these procedures are used to obtain dynamics ("smoothly calibrated learning") that are uncoupled, have finite memory, and are close to Nash equilibria most of the time (while the similar dynamics that are based on regular calibration yield only the time average becoming close to correlated equilibria; see Foster and Vohra, 1997).

1.1. Literature

The *calibration problem* has been extensively studied, starting with Dawid (1982), Oakes (1985), and Foster and Vohra (1998); see Olszewski (2015) for a comprehensive survey of the literature. Kakade and Foster (2004) and Foster and Kakade (2006) introduced the notion of *weak calibration*, which shares many properties with smooth calibration. In particular, both can be guaranteed by deterministic procedures, and both are of the "general fixed point" variety: they can find fixed points of arbitrary continuous functions (see for instance the last paragraph in Section 2.3).[4] However, while weak calibration may be at times technically more convenient to work with, smooth calibration is the more natural concept, easier to interpret and understand; it is, after all, just a standard smoothing of regular calibration.

The *online regression problem*—see Section 3 for details—was introduced by Foster (1991); for further improvements, see J. Foster (1999), Vovk (2001), Azoury and Warmuth (2001), and the book of Cesa-Bianchi and Lugosi (2006).

2. Calibration: model and result

In this section we present the calibration game in its standard and "leaky" versions, introduce the notion of smooth calibration, and state our main results.

[1] Corollary 12 in Section 4 will formally show that regular calibration implies smooth calibration.

[2] When the forecasting procedure is deterministic it can be simulated by the agent, and so it is irrelevant whether the agent *observes* the forecasts, or just *computes* them by himself, before taking his action.

[3] Another, seemingly less elegant, way to obtain this is by restarting the procedure once in a while; see, e.g., Lehrer and Solan (2009).

[4] They are thus more "powerful" than the standard calibration procedures (such as those based on Blackwell's approachability), which find *linear* fixed points (such as eigenvectors and invariant probabilities).

2.1. The calibration game

Let[5] $C \subseteq \mathbb{R}^m$ be a compact convex set, and let $A \subseteq C$ (for example, C could be the set of probability distributions $\Delta(A)$ over a finite set A, which is identified with the set of unit vectors in C, or a product of such sets). The *calibration game* has two players: the "action" player—the "A-player" for short—and the "conjecture" (or "calibrating") player—the "C-player" for short. At each time period $t = 1, 2, \ldots$, the C-player chooses $c_t \in C$ and the A-player chooses $a_t \in A$. There is full monitoring and perfect recall: at time t both players know the realized history $h_{t-1} = (c_1, a_1, \ldots, c_{t-1}, a_{t-1}) \in (C \times A)^{t-1}$.

In the *standard* calibration game, c_t and a_t are chosen simultaneously (perhaps in a mixed, i.e., randomized, way). In the *leaky* calibration game, a_t is chosen after c_t has been chosen and revealed; thus, c_t is a function of h_{t-1}, whereas a_t is a function of h_{t-1} *and* c_t. Formally, a pure strategy of the C-player is $\sigma : \cup_{t\geq 1}(C \times A)^{t-1} \to C$, and a pure strategy of the A-player is $\tau : \cup_{t\geq 1}(C \times A)^{t-1} \to A$ in the standard game, and $\tau : \cup_{t\geq 1}(C \times A)^{t-1} \times C \to A$ in the leaky game. A pure strategy of the C-player will also be referred to as *deterministic*.

The calibration score—which the C-player wants to minimize—is defined at time $T \geq 1$ as follows. For every forecast c in C let $n(c) \equiv n_T(c) := |\{1 \leq t \leq T : c_t = c\}$ be the number of times that it has been used, and let

$$\bar{a}(c) \equiv \bar{a}_T(c) := \frac{1}{n(c)} \sum_{s=1}^{T} \mathbf{1}_{c_s = c}\, a_s$$

be the average of the actions in the periods when the forecast was c, where we write $\mathbf{1}_{x=y}$ for the indicator that $x = y$ (i.e., $\mathbf{1}_{x=y} = 1$ when $x = y$ and $\mathbf{1}_{x=y} = 0$ otherwise); $\bar{a}(c)$ is defined only when c appears in the sequence $c_1, \ldots, c_T$, i.e., $n(c) > 0$. The *calibration score* at time T is then defined as[6]

$$K_T := \sum_{c \in C} \frac{n(c)}{T} \|\bar{a}(c) - c\|.$$

Thus K_T is the mean distance between the forecast c and the average $\bar{a}(c)$ of the actions a in those periods where the forecast was c, weighted proportionately to how often each forecast c has been used in those T periods. An alternative formulation is easily seen to be

$$K_T = \frac{1}{T} \sum_{t=1}^{T} \|\bar{a}_t - c_t\|, \tag{1}$$

where $\bar{a}_t := \bar{a}(c_t)$, i.e.,

$$\bar{a}_t := \frac{\sum_{s=1}^{T} \mathbf{1}_{c_s = c_t}\, a_s}{\sum_{s=1}^{T} \mathbf{1}_{c_s = c_t}};$$

indeed, for each c there are $n(c)$ identical terms in (1) that each equal $\|\bar{a}(c) - c\|$.

2.2. Smooth calibration

We introduce the notion of "smooth calibration". A *smoothing function* is a function $\Lambda : C \times C \to [0, 1]$ with $\Lambda(c, c) = 1$ for every c. Its interpretation is that $\Lambda(c', c)$ gives the weight that we assign to c' when we are at c. We will use $\Lambda(c', c)$ instead of the indicator $\mathbf{1}_{c'=c}$ to "smooth" out the forecasts and the average actions. Specifically, put

$$\bar{a}_t^{\Lambda} := \frac{\sum_{s=1}^{T} \Lambda(c_s, c_t)\, a_s}{\sum_{s=1}^{T} \Lambda(c_s, c_t)} \quad \text{and} \quad c_t^{\Lambda} := \frac{\sum_{s=1}^{T} \Lambda(c_s, c_t)\, c_s}{\sum_{s=1}^{T} \Lambda(c_s, c_t)}.$$

The Λ-*smoothed calibration score* at time T is then defined as

$$K_T^{\Lambda} = \frac{1}{T} \sum_{t=1}^{T} \|\bar{a}_t^{\Lambda} - c_t^{\Lambda}\|. \tag{2}$$

A standard (and useful) assumption is a Lipschitz condition: there exists $L < \infty$ such that $|\Lambda(c', c) - \Lambda(c'', c)| \leq L\|c' - c''\|$ for all $c, c', c'' \in C$. Thus, the functions $\Lambda(\cdot, c)$ are uniformly Lipschitz: $\mathcal{L}(\Lambda(\cdot, c)) \leq L$ for every $c \in C$, where $\mathcal{L}(f) := \sup\{\|f(x) - f(y)\| / \|x - y\| : x, y \in X,\ x \neq y\}$ denotes the *Lipschitz constant* of the function f (if f is not a Lipschitz function then $\mathcal{L}(f) = +\infty$; when $\mathcal{L}(f) \leq L$ we say that f is *L-Lipschitz*).

[5] We denote by $\mathbb{R}^m$ the m-dimensional Euclidean space, with the usual ℓ_2-norm $\|\cdot\|$.

[6] The summation is over all c with $n(c) > 0$, and we use the Euclidean norm (the squared distance $\|\bar{a}(c) - c\|^2$ may well be used instead, in line with standard statistics usage).

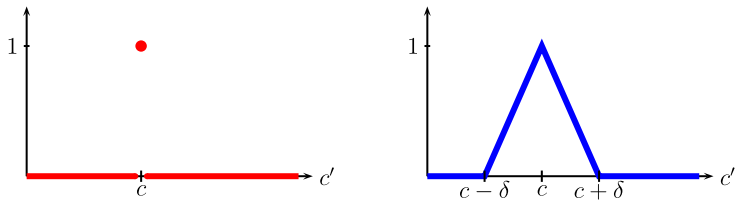

Fig. 1. *Left*: The indicator function $\mathbf{1}_{x=c}$. *Right*: The δ-tent smoothing function $\Lambda(x,c) = [1 - \|x-c\|/\delta]_+$ with Lipschitz bound $L = 1/\delta$.

Two classic examples of Lipschitz smoothing functions are: (i) the so-called *tent function* $\Lambda(c',c) = [1 - \|c'-c\|/\delta]_+$ for $\delta > 0$, where $[z]_+ := \max\{z, 0\}$; thus, only points c' within distance δ of c are considered, and their weight is proportional to the distance from c (see Fig. 1 *Right* for this function Λ, and compare it with the indicator function in Fig. 1 *Left*); and (ii) the so-called *Gaussian* (or *normal*) smoothing function $\Lambda(c',c) = \exp(-\|c'-c\|^2/(2\sigma^2))$.

Remarks. *(a)* The original calibration score K_T is obtained when Λ is the indicator function, i.e., $\Lambda(c',c) = \mathbf{1}_{c'=c}$ for all $c, c' \in C$.

(b) The normalization $\Lambda(c,c) = 1$ pins down the Lipschitz constant (otherwise one could replace Λ with $\alpha\Lambda$ for small $\alpha > 0$, and so lower the Lipschitz constant without affecting the score).

(c) Smoothing both $\bar{a}_t$ and c_t and then taking the difference is the same as smoothing the difference: $\bar{a}_t^\Lambda - c_t^\Lambda = (\bar{a}_t - c_t)^\Lambda$. Moreover, smoothing a_t is the same as smoothing $\bar{a}_t$, i.e., $\bar{a}_t^\Lambda = a_t^\Lambda$.

(d) An alternative score smoothes only the average action $\bar{a}_t$, but not the forecast c_t:

$$\tilde{K}_T^\Lambda = \frac{1}{T}\sum_{t=1}^{T} \|\bar{a}_t^\Lambda - c_t\|.$$

If the smoothing function puts positive weight only in small neighborhoods, i.e., there is $\delta > 0$ such that $\Lambda(c',c) > 0$ only when $\|c'-c\| \le \delta$, then the difference between K_T^Λ and $\tilde{K}_T^\Lambda$ is at most δ (because in this case $\|c_t^\Lambda - c_t\| \le \delta$ for every t). More generally, $|K_T^\Lambda - \tilde{K}_T^\Lambda| \le \delta$ when $(1/T)\sum_{t=1}^T \|c_t^\Lambda - c_t\| \le \delta$ for any collection of points $c_1, \ldots, c_T \in C$, which is indeed the case, for instance, for the Gaussian smoothing with small enough σ^2. The reason that we prefer to use K^Λ rather than $\tilde{K}^\Lambda$ is that K^Λ vanishes when there is perfect calibration (i.e., $\bar{a}_t = c_t$ for all t), whereas $\tilde{K}^\Lambda$ remains positive; clean statements such as $K_T^\Lambda \le \varepsilon$ become $\tilde{K}_T^\Lambda \le \varepsilon + \delta$.

Finally, given $\varepsilon > 0$ and $L < \infty$, we will say that a strategy of the C-player—which is also called a "procedure"—is (ε, L)-*smoothly calibrated* if there is $T_0 \equiv T_0(\varepsilon, L)$ such that

$$K_T^\Lambda = \frac{1}{T}\sum_{t=1}^{T} \left\|\bar{a}_t^\Lambda - c_t^\Lambda\right\| \le \varepsilon \tag{3}$$

holds almost surely, for every strategy of the A-player, every $T > T_0$, and every smoothing function $\Lambda : C \times C \to [0,1]$ that is L-Lipschitz in the first coordinate. Unlike standard calibration, which can be guaranteed only with high probability, smooth calibration may be obtained by deterministic procedures—as will be shown below—in which case we may well require (3) to always hold (rather than just almost surely).

2.3. Leaky forecasts

We will say that a procedure (i.e., a strategy of the C-player) is *leaky* (smoothly) calibrated if it is (smoothly) calibrated also in the leaky setup, that is, against an A-player who may choose his action a_t at time t depending on the forecast c_t made by the C-player at time t (i.e., the A-player moves after the C-player). While, as we saw in the Introduction, there are no leaky calibrated procedures, we will show that there are *leaky smoothly calibrated* procedures.

Deterministic procedures (i.e., pure strategies of the C-player) are clearly leaky: the A-player can use the procedure at each period t to compute c_t as a function of the history h_{t-1}, and only then determine his action a_t. Thus, in particular, there cannot be deterministic calibrated procedures (because there are no leaky calibrated procedures); see Dawid (1985) and Oakes (1985).

In the case of smooth calibration, the procedure that we construct is deterministic, and thus leaky smoothly calibrated. However, there are also randomized leaky smoothly calibrated procedures. One example is the simple calibrated procedure of D. Foster (1999) in the one-dimensional case (where $A = \{$"rain", "no rain"$\}$ and $C = [0,1]$): the forecast there is "almost deterministic", in the sense that the randomization is only between two very close forecasts (such as 29.99% and 30.01%), and so can be shown to be leaky smoothly calibrated. For another example, see footnote 22 in Section 4 below.

A particular instance of the leaky setup is one where the A-player uses a fixed reaction function $g : C \to A$ that is a continuous mapping of forecasts to actions; thus, $a_t = g(c_t)$ (independently of time t and history h_{t-1}). In this case, leaky smooth calibration implies that most of the forecasts that are used must be approximate fixed points of g; indeed, in every period in which the forecast is c the action is the same, namely, $g(c)$, and so the average of the actions in all the periods where the forecast is (close to) c is (close to) $g(c)$ (use the continuity of g here); formally, see the arguments in part (iv) of the proof of Theorem 15 in Section 6). Thus, leaky procedures find (approximate) fixed points for arbitrary continuous functions g, and so must in general be more complex than the procedures that yield calibration (such as those obtained by Blackwell's approachability); cf. the complexity class PPAD (Papadimitriou, 1994) in the computer science literature (see also Hazan and Kakade, 2012 for the connection to calibration).

2.4. Result

A strategy σ has *finite recall* and is *stationary* if there exists a finite integer $R \geq 1$ and a function $\tilde{\sigma} : (C \times A)^R \to C$ such that

$$\sigma(h_{T-1}) = \tilde{\sigma}(c_{T-R}, a_{T-R}, c_{T-R+1}, a_{T-R+1}, \ldots, c_{T-1}, a_{T-1})$$

for every $T > R$ and history $h_{T-1} = (c_t, a_t)_{1 \leq t \leq T-1}$. Thus, only the "window" consisting of the last R periods matters; the rest of the history, as well as the calendar time T, do not. Finally, a finite set $D \subseteq C$ is a δ-*grid* for C if for every $c \in C$ there is $d \equiv d(c) \in D$ such that $\|d - c\| \leq \delta$.

Our result is:

Theorem 1. *For every $\varepsilon > 0$ and $L < \infty$ there is an (ε, L)-smoothly calibrated procedure. Moreover, the procedure may be taken to be:*

- *deterministic;*
- *leaky;*
- *with finite recall and stationary; and*
- *with all the forecasts lying on a finite grid.*[7]

The proof will proceed as follows. First, we construct deterministic finite-recall algorithms for the *online linear regression* problem (cf. Foster, 1991; Azoury and Warmuth, 2001); see Theorem 2 in Section 3. Next, we use these algorithms to get deterministic finite-recall *weakly calibrated* procedures (cf. Foster and Kakade 2004, 2006); see Theorem 10 in Section 4. Finally, we obtain smooth calibration from weak calibration; see Section 5.

3. Online linear regression

Classical linear regression tries to predict a variable y from a vector x of d variables (and so $y \in \mathbb{R}$ and $x \in \mathbb{R}^d$). There are observations $(x_t, y_t)_t$, and one typically assumes that[8] $y_t = \theta' x_t + \epsilon_t$, where ϵ_t are (zero-mean normally distributed) error terms. The optimal estimator for θ is then given by the least squares method; i.e., θ minimizes $(1/T)\sum_{t=1}^{T} \psi_t(\theta)$ with

$$\psi_t(\theta) := (y_t - \theta' x_t)^2$$

for every t.

In the *online linear regression* problem (Foster, 1991; see Section 1.1), the observations arrive sequentially, and at each time period t we want to determine θ_t given the information at that time, namely, $(x_1, y_1), \ldots, (x_{t-1}, y_{t-1})$ and x_t only. The goal is to bound the difference between the mean square errors in the online case and the offline case (i.e., "in hindsight"); namely,

$$\frac{1}{T}\sum_{t=1}^{T} \psi_t(\theta_t) - \frac{1}{T}\sum_{t=1}^{T} \psi_t(\theta).$$

Thus, an online linear-regression *algorithm* takes as input a sequence $(x_t, y_t)_{t \geq 1}$ in $\mathbb{R}^d \times \mathbb{R}$ and gives as output a sequence $(\theta_t)_{t \geq 1}$ in $\mathbb{R}^d$, such that θ_t is a function only of $x_1, y_1, \ldots, x_{t-1}, y_{t-1}, x_t$, for each t.

Our result is:

Theorem 2. *Let X, Y be positive reals, and $\varepsilon > 0$. Then there exists a positive integer $R_0 \equiv R_0(\varepsilon, X, Y, d)$ such that for every $R > R_0$ there is an R-recall stationary deterministic algorithm that gives $(\theta_t)_{t \geq 1}$, such that*

[7] The sizes R of the recall and δ of the grid depend on ε, L, the dimension m, and the bound on the compact set C.

[8] Vectors are viewed as column vectors, and θ' denotes the transpose of θ (thus $\theta' x$ is the scalar product $\theta \cdot x$ of θ and x).

$$\frac{1}{R}\sum_{t=T-R+1}^{T}[\psi_t(\theta_t)-\psi_t(\theta)]\le\varepsilon(1+\|\theta\|^2)\quad \textit{and} \tag{4}$$

$$\frac{1}{T}\sum_{t=1}^{T}[\psi_t(\theta_t)-\psi_t(\theta)]\le\varepsilon(1+\|\theta\|^2) \tag{5}$$

hold for every $T\ge R$, every $\theta\in\mathbb{R}^d$, and every sequence $(x_t,y_t)_{t\ge1}$ in $\mathbb{R}^d\times\mathbb{R}$ with $\|x_t\|\le X$ and $|y_t|\le Y$ for all t.

When in addition θ is bounded,[9] say, $\|\theta\|\le M$, the mean square error of our online algorithm is guaranteed not to exceed the optimal offline mean square error by more than $\varepsilon(1+M^2)$.

The proof of Theorem 2 in the remainder of this section proceeds as follows.

(i) *Forward algorithm* (Section 3.1). We start with the "forward algorithm" of Azoury and Warmuth (2001) and the resulting bound on the sum of regrets $\psi_t(\theta_t)-\psi_t(\theta)$ (Theorem 3).
(ii) *Discounted forward algorithm* (Section 3.2). We modify the procedure by introducing a λ-discount factor, which gives bounds on the discounted sum of regrets (Propositions 4 and 5).
(iii) *Windowed discounted forward algorithm* (Section 3.3). We further modify the procedure by restricting the history to a window consisting only of the last R periods, which gives bounds on the sum of regrets over that window (Proposition 8).
(iv) *Conclusion* (Section 3.3). One of the useful properties of discounting is that the weight of the initial segment from 1 up to $T-R$ is small relative to the whole sum from 1 to T, and so dropping that initial segment has little effect on the procedure and the resulting estimates. We can thus choose an appropriate discount factor λ and a window size R yielding the desired bounds on the windowed sum of regrets, and thus also on the time average of the regrets (Proposition 9, which yields Theorem 2).

3.1. Forward algorithm

The starting point is the following algorithm of Azoury and Warmuth (2001, Section 5.4). For each $a>0$, the *a-forward algorithm* gives[10] $\theta_t=Z_t^{-1}v_t$, where

$$Z_t=aI+\sum_{q=1}^{t}x_qx_q'\quad\text{and}\quad v_t=\sum_{q=1}^{t-1}y_qx_q. \tag{6}$$

Theorem 3 *(Azoury and Warmuth, 2001). For every $a>0$, the a-forward algorithm yields*

$$\sum_{t=1}^{T}[\psi_t(\theta_t)-\psi_t(\theta)]\le a\,\|\theta\|^2+\sum_{t=1}^{T}y_t^2\left(1-\frac{\det(Z_{t-1})}{\det(Z_t)}\right) \tag{7}$$

for every $T\ge1$, every $\theta\in\mathbb{R}^d$, and every sequence $(x_t,y_t)_{t\ge1}$ in $\mathbb{R}^d\times\mathbb{R}$.

Proof. Theorem 5.6 and Lemma A.1 in Azoury and Warmuth (2001), where Z_t denotes their η_t^{-1} matrix; the second term in their formula (5.17) is nonnegative since η_t is a positive definite matrix.[11] □

3.2. Discounted forward algorithm

Let $a>0$ and $0<\lambda<1$. The *λ-discounted a-forward algorithm* gives $\theta_t=Z_t^{-1}v_t$, where

$$Z_t=aI+\sum_{q=1}^{t}\lambda^{t-q}x_qx_q'\quad\text{and}\quad v_t=\sum_{q=1}^{t-1}\lambda^{t-q}y_qx_q. \tag{8}$$

[9] For example, θ lies in the unit simplex of $\mathbb{R}^d$.
[10] Z_t^{-1} is the inverse of the $d\times d$ matrix Z_t (which is invertible because $a>0$), and I denotes the identity matrix.
[11] Our statement is different from theirs because ψ_t equals twice L_t, and there is a misprinted sign in the first line of their formula (5.17).

Proposition 4. *For every $a > 0$ and $0 < \lambda < 1$, the λ-discounted a-forward algorithm yields*

$$\sum_{t=1}^{T} \lambda^{T-t} [\psi_t(\theta_t) - \psi_t(\theta)] \le a \|\theta\|^2 + \sum_{t=1}^{T} \lambda^{T-t} y_t^2 \left(1 - \lambda^d \frac{\det(Z_{t-1})}{\det(Z_t)}\right) \tag{9}$$

for every $T \ge 1$, every $\theta \in \mathbb{R}^d$, and every sequence $(x_t, y_t)_{t\ge1}$ in $\mathbb{R}^d \times \mathbb{R}$.

Proof. Let $b := \sqrt{a(1-\lambda)}$. From the sequence $(x_t, y_t)_{t\ge1}$ we construct a sequence $(\tilde{x}_s, \tilde{y}_s)_{s\ge1}$ in blocks as follows. For every $t \ge 1$, the t-th block B_t is of size $d+1$ and consists of $(\lambda^{-t/2} b e^{(1)}, 0), \ldots, (\lambda^{-t/2} b e^{(d)}, 0), (\lambda^{-t/2} x_t, \lambda^{-t/2} y_t)$, where $e^{(i)}$ is the i-th unit vector in $\mathbb{R}^d$. The a-forward algorithm applied to $(\tilde{x}_s, \tilde{y}_s)_{s\ge1}$ yields the following.

For $s = (d+1)t$, i.e., at the end of the B_t block, we have[12] $\sum_{s\in B_t} \tilde{x}_s \tilde{x}_s' = b^2 \lambda^{-t} \sum_{i=1}^{d} e^{(i)} \left(e^{(i)}\right)' + \lambda^{-t} x_t x_t' = \lambda^{-t}(b^2 I + x_t x_t')$; thus

$$\begin{aligned}\tilde{Z}_{(d+1)t} &= aI + \sum_{q=1}^{t} \sum_{s\in B_t} \tilde{x}_s \tilde{x}_s' = aI + \sum_{q=1}^{t} \lambda^{-q} (b^2 I + x_q x_q') \\ &= \lambda^{-t} \left(aI + \sum_{q=1}^{t} \lambda^{t-q} x_q x_q' \right) = \lambda^{-t} Z_t\end{aligned}$$

(since $\sum_{i=1}^{d} e^{(i)} \left(e^{(i)}\right)' = I$ and $b^2 = (1-\lambda)a$; recall (8)). Together with $\tilde{v}_{(d+1)t} = \sum_{q=1}^{t} \sum_{s\in B_t} \tilde{y}_s \tilde{x}_s = \sum_{q=1}^{t} \lambda^{-q} y_q x_q = \lambda^{-t} v_t$ (only the first entry in each block has a nonzero $\tilde{y}$), it follows that $\tilde{\theta}_{(d+1)t}$ indeed equals $\theta_t = Z_t^{-1} v_t$ as given by (8).

Next, for every t we have $\sum_{s\in B_t} \tilde{\psi}_s(\tilde{\theta}_s) \ge \lambda^{-t} \psi_t(\theta_t)$ (all terms in the sum are nonnegative, and we drop all except the last one). Also, for every $\theta \in \mathbb{R}^d$,

$$\sum_{s\in B_t} \tilde{\psi}_s(\theta) = \lambda^{-t} \left(b^2 \sum_{i=1}^{d} (\theta' e^{(i)})^2 + \psi_t(\theta) \right) = \lambda^{-t} \left(b^2 \|\theta\|^2 + \psi_t(\theta) \right).$$

Thus the left-hand side of (7) evaluated at the end of the T-th block B_T satisfies

$$\begin{aligned}LHS &\ge \sum_{t=1}^{T} \lambda^{-t} \psi_t(\theta_t) - a\|\theta\|^2 - b^2 \|\theta\|^2 \sum_{t=1}^{T} \lambda^{-t} - \sum_{t=1}^{T} \lambda^{-t} \psi_t(\theta) \\ &= \sum_{t=1}^{T} \lambda^{-t} [\psi_t(\theta_t) - \psi_t(\theta)] - \lambda^{-T} a \|\theta\|^2.\end{aligned}$$

On the right-hand side we get

$$RHS = \sum_{t=1}^{T} \sum_{s\in B_t} \tilde{y}_s^2 \left(1 - \frac{\det(\tilde{Z}_{s-1})}{\det(\tilde{Z}_s)}\right) = \sum_{t=1}^{T} \lambda^{-t} y_t^2 \left(1 - \frac{\det(\tilde{Z}_{(d+1)t-1})}{\det(\tilde{Z}_{(d+1)t})}\right)$$

(again, only the last term in each block has nonzero $\tilde{y}_s$). We have seen above that $\tilde{Z}_{(d+1)t} = \lambda^{-t} Z_t$; thus $\tilde{Z}_{(d+1)t-1} = \tilde{Z}_{(d+1)t} - \lambda^{-t} x_t x_t' = \lambda^{-t}(Z_t - x_t x_t') = \lambda^{-t+1} Z_{t-1} + (\lambda^{-t} - \lambda^{-t+1}) aI$. Therefore $\det(\tilde{Z}_{(d+1)t-1}) \ge \det(\lambda^{-t+1} Z_{t-1})$ (indeed, if B is a positive definite matrix and $\beta > 0$ then[13] $\det(B + \beta I) > \det(B)$). Therefore we obtain

$$\begin{aligned}RHS &\le \sum_{t=1}^{T} \lambda^{-t} y_t^2 \left(1 - \frac{\det(\lambda^{-t+1} Z_{t-1})}{\det(\lambda^{-t} Z_t)}\right) \\ &= \sum_{t=1}^{T} \lambda^{-t} y_t^2 \left(1 - \lambda^d \frac{\det(Z_{t-1})}{\det(Z_t)}\right)\end{aligned}$$

(the matrices Z_t are of size $d \times d$, and so $\det(cZ_t) = c^d \det(Z_t)$). Recalling that $LHS \le RHS$ by (7) and multiplying by λ^T yields the result. □

[12] The notation $\tilde{Z}_s$, $\tilde{\psi}_s$, ... pertains to the $(\tilde{x}_s, \tilde{y}_s)_{s\ge1}$ problem.

[13] Let $\beta_1, \ldots, \beta_d > 0$ be the eigenvalues of B; then the eigenvalues of $B + cI$ are $\beta_1 + c, \ldots, \beta_d + c$, and so $\det(B + cI) = \prod_i (\beta_i + c) > \prod_i \beta_i = \det(B)$.

Remark. From now on it will be convenient to assume that $\|x_t\| \le 1$ and $|y_t| \le 1$ (i.e., $X = Y = 1$); for general X and Y, multiply $x_t, y_t, \theta_t, \psi_t, a$ by $X, Y, Y/X, Y^2, X^2$, respectively, in the appropriate formulas.

Proposition 5. *For every $a > 0$ and $1/4 \le \lambda < 1$ there exists a constant $D_1 \equiv D_1(a, \lambda, d)$ such that the λ-discounted a-forward algorithm yields*

$$\sum_{t=1}^{T} \lambda^{T-t} [\psi_t(\theta_t) - \psi_t(\theta)] \le a \|\theta\|^2 + D_1, \tag{10}$$

for every $T \ge 1$, every $\theta \in \mathbb{R}^d$, and every sequence $(x_t, y_t)_{t \ge 1}$ in $\mathbb{R}^d \times \mathbb{R}$ with $\|x_t\| \le 1$ and $|y_t| \le 1$ for all t.

Proof. Let $K \ge 1$ be an integer such that $1/4 \le \lambda^K \le 1/2$. Given $T \ge 1$, let the integer $m \ge 1$ satisfy $(m-1)K < T \le mK$. Writing ζ_t for $\det(Z_t)$, we have

$$\begin{aligned}
\sum_{t=1}^{T} \lambda^{T-t} \left(1 - \lambda^d \frac{\zeta_{t-1}}{\zeta_t}\right) &\le \sum_{t=1}^{mK} \lambda^{T-t} \left(1 - \lambda^d \frac{\zeta_{t-1}}{\zeta_t}\right) \\
&= \lambda^T \sum_{j=1}^{m-1} \sum_{t=jK+1}^{(j+1)K} \lambda^{-t} \ln\left(\lambda^{-d} \frac{\zeta_t}{\zeta_{t-1}}\right) \\
&\le \lambda^T \sum_{j=1}^{m-1} \lambda^{-(j+1)K} \sum_{t=jK+1}^{(j+1)K} \ln\left(\lambda^{-d} \frac{\zeta_t}{\zeta_{t-1}}\right) \\
&\le \lambda^T \sum_{j=1}^{m-1} \lambda^{-(j+1)K} \ln\left(\lambda^{-dK} \frac{\zeta_{(j+1)K}}{\zeta_{jK}}\right)
\end{aligned} \tag{11}$$

(in the second line we have used $1 - 1/u \le \ln u$ for $0 < u \le 1$, as in (4.21) in Azoury and Warmuth, 2001; in the third line, $\lambda^{-t} \le \lambda^{-(j+1)K}$ since $t \le (j+1)K$ and $\lambda < 1$).

Let $B = (b_{ij})$ be a $d \times d$ symmetric positive definite matrix with $|b_{ij}| \le \beta$ for all i, j, and let $a > 0$. Then $a^d \le \det(aI + B) \le d!(a + \beta)^d$. Indeed, the second inequality follows easily since the determinant is the sum of $d!$ products of d elements each. For the first inequality, let $\beta_1, ..., \beta_d > 0$ be the eigenvalues of B; then the eigenvalues of $aI + B$ are $a + \beta_1, ..., a + \beta_d$, and so $\det(aI + B) = \Pi_{i=1}^d (a + \beta_i) > a^d$. Applying this to Z_t (using (8), $|x_{t,i} x_{t,j}| \le \|x_t\|^2 \le 1$, and $\sum_{t=1}^{T} \lambda^{T-t} < 1/(1-\lambda)$) yields

$$a^d \le \zeta_t \equiv \det(Z_t) \le d! \left(a + \frac{1}{1-\lambda}\right)^d.$$

Therefore, since $\lambda^{-K} \le 4$, we get

$$\lambda^{-dK} \frac{\zeta_{(j+1)K}}{\zeta_{jK}} \le 4^d d! \left(1 + \frac{1}{a(1-\lambda)}\right)^d =: D,$$

and so (11) is

$$\begin{aligned}
&\le \lambda^T \sum_{j=1}^{m-1} \left(\lambda^{-K}\right)^{j+1} \ln D \le \lambda^T \frac{\lambda^{-K(m+1)} - \lambda^{-2K}}{\lambda^{-K} - 1} \ln D \\
&\le \lambda^T \frac{\lambda^{-T} \lambda^{-K} - 0}{2 - 1} \ln D = 4 \ln D
\end{aligned}$$

(since $2 \le \lambda^{-K} \le 4$ and $K(m+1) < T + K$). Substituting this in (9) and putting

$$D_1 := 4 \left(\ln d! + d \ln 4 + d \ln\left(1 + \frac{1}{a(1-\lambda)}\right)\right) \tag{12}$$

completes the proof. □

3.3. Windowed discounted forward algorithm

From now on it is convenient to put $(x_t, y_t, \theta_t) = (0, 0, 0)$ for all $t \le 0$.

Let $a > 0$, $0 < \lambda < 1$, and integer $R \ge 1$. The *R-windowed λ-discounted a-forward algorithm* gives $\theta_t = Z_t^{-1} v_t$, where[14]

$$Z_t = aI + \sum_{q=t-R+1}^{t} \lambda^{t-q} x_q x_q' \quad \text{and} \quad v_t = \sum_{q=t-R+1}^{t-1} \lambda^{R-q} y_q x_q. \tag{13}$$

Lemma 6. *For every $a > 0$ and $0 < \lambda < 1$ there exists a constant $D_2 \equiv D_2(a, \lambda, d)$ such that if $(\tilde\theta_t)_{t\ge1}$ is given by the λ-discounted a-forward algorithm, and $(\theta_t)_{t\ge1}$ is given by the R-windowed λ-discounted a-forward algorithm for some integer $R \ge 1$, then*

$$\left|\psi_t(\tilde\theta_t) - \psi_t(\theta_t)\right| \le D_2 \lambda^R \tag{14}$$

for every[15] $t \ge 1$ and every sequence $(x_t, y_t)_{t\ge1}$ in $\mathbb{R}^d \times \mathbb{R}$ with $\|x_t\| \le 1$ and $|y_t| \le 1$ for all t.

To prove this lemma we use the following basic result. The norm of a matrix A is $\|A\| := \max_{z\ne0} \|Az\| / \|z\|$.

Lemma 7. *For $k = 1, 2$, let $c_k = A_k^{-1} b_k$, where A_k is a $d \times d$ symmetric matrix whose eigenvalues are all greater than or equal to some $\alpha > 0$, and $\|b_k\| \le M$. Then $\|c_k\| \le M/\alpha$ and*

$$\|c_1 - c_2\| \le \frac{1}{\alpha} \|b_1 - b_2\| + \frac{M}{\alpha^2} \|A_1 - A_2\|.$$

Proof. First, $\|c_k\| \le \left\|A_k^{-1}\right\| \|b_k\| \le (1/\alpha)M$ since $\left\|A_k^{-1}\right\|$ equals the maximal eigenvalue of A_k^{-1}, which is the reciprocal of the minimal eigenvalue of A_k, and so $\left\|A_k^{-1}\right\| \le 1/\alpha$.

Second, express $c_1 - c_2$ as $A_1^{-1}(b_1 - b_2) + A_1^{-1}(A_2 - A_1)A_2^{-1}b_2$, to get

$$\|c_1 - c_2\| \le \left\|A_1^{-1}\right\| \|b_1 - b_2\| + \left\|A_1^{-1}\right\| \|A_2 - A_1\| \left\|A_2^{-1}\right\| \|b_2\|$$

and the proof is complete. □

Proof of Lemma 6. For $t \le R$ we have $\tilde\theta_t \equiv \theta_t$, and so consider $t > R$. We have[16] $\left\|\tilde v_t\right\|, \|v_t\| \le \sum_{q=1}^{\infty} \lambda^q = 1/(1-\lambda)$. The matrices $\tilde Z_t$ and Z_t are the sum of aI and a positive-definite matrix, and so their eigenvalues are $\ge a$. Next,

$$\left\|\tilde v_t - v_t\right\| = \left\|\sum_{q=1}^{t-R} \lambda^{t-q} y_q x_q\right\| \le \frac{\lambda^R}{1-\lambda};$$

similarly, for each element $(\tilde Z_t - Z_t)_{ij}$ of $\tilde Z_t - Z_t$ we have

$$\left|(\tilde Z_t - Z_t)_{ij}\right| = \left|\sum_{q=1}^{t-R} \lambda^{t-q} x_{q,i} x_{q,j}\right| \le \frac{\lambda^R}{1-\lambda},$$

and so[17] $\left\|\tilde Z_t - Z_t\right\| \le d\lambda^R/(\lambda - 1)$. Using Lemma 7 yields

$$\left\|\tilde\theta_t - \theta_t\right\| \le \frac{1}{a}\frac{\lambda^R}{1-\lambda} + \frac{1}{a^2}\frac{d\lambda^R}{1-\lambda} = \frac{\lambda^R(a+d)}{(1-\lambda)a^2}.$$

[14] The sums below effectively start at $\min\{t - R + 1, 1\}$ (because we put $x_q = 0$ for $q \le 0$).

[15] For $t \le R$ we have $\tilde\theta_t = \theta_t$ since they are given by the same formula.

[16] Notation: $\tilde v_t$ and $\tilde Z_t$ pertain to the sequence $\tilde\theta_t$ given by the λ-discounted a-forward algorithm, whereas v_t and Z_t pertain to the sequence θ_t given by the R-windowed λ-discounted a-forward algorithm.

[17] Because $\|A\| \le d \max_{i,j} |a_{ij}|$ for any $d \times d$ matrix A.

Hence

$$\begin{aligned}\left|\psi_t(\tilde{\theta}_t)-\psi_t(\theta_t)\right| &= \left|(y_t-\tilde{\theta}_t'x_t)^2-(y_t-\theta_t'x_t)^2\right| \\ &= \left|(\tilde{\theta}_t'-\theta_t')x_t\cdot\left(2y_t-(\tilde{\theta}_t'+\theta_t')x_t\right)\right| \\ &\le \left\|\tilde{\theta}_t-\theta_t\right\|\left(2+\left\|\tilde{\theta}_t\right\|+\|\theta_t\|\right) \\ &\le \frac{\lambda^R(a+d)}{(1-\lambda)a^2}\left(2+\frac{2}{(1-\lambda)a}\right)=D_2\lambda^R,\end{aligned}$$

where

$$D_2 := \frac{2\,(a+d)\,(a(1-\lambda)+1)}{a^3(1-\lambda)^2}; \tag{15}$$

this completes the proof. □

Proposition 8. *For every $a>0$ and $1/4\le\lambda<1$ there exist constants $D_1\equiv D_1(a,\lambda,d)$ and $D_2\equiv D_2(a,\lambda,d)$ such that for every integer $R\ge 1$ the R-windowed λ-discounted a-forward algorithm yields*

$$\frac{1}{R}\sum_{t=T-R+1}^{T}[\psi_t(\theta_t)-\psi_t(\theta)]\le(a\,\|\theta\|^2+D_1)\left(1-\lambda+\frac{\lambda}{R}\right)+\frac{(\|\theta\|+1)^2}{R(1-\lambda)}+D_2\lambda^R \tag{16}$$

for every $T\ge 1$, every $\theta\in\mathbb{R}^d$, and every sequence $(x_t,y_t)_{t\ge1}$ in $\mathbb{R}^d\times\mathbb{R}$ with $\|x_t\|\le 1$ and $|y_t|\le 1$ for all t.

Proof. Let $\tilde{\theta}_t$ be given by the λ-discounted a-forward algorithm. Put $g_t:=\psi_t(\theta_t)-\psi_t(\theta)$ (where θ_t is given by the R-windowed λ-discounted a-forward algorithm) and $\tilde{g}_t:=\psi_t(\tilde{\theta}_t)-\psi_t(\theta)$. Apply (10) at T, and also at each one of $T-R+1, T-R+2, \ldots, T-1$; multiply those by $1-\lambda$ and add them all up, to get

$$\begin{aligned}\sum_{t=1}^{T}\lambda^{T-t}\tilde{g}_t+(1-\lambda)\sum_{r=1}^{R-1}\sum_{t=1}^{T-r}\lambda^{T-r-t}\tilde{g}_t &\le (a\,\|\theta\|^2+D_1)(1+(R-1)(1-\lambda)) \\ &= (a\,\|\theta\|^2+D_1)(R-R\lambda+\lambda).\end{aligned}$$

For $t\le T-R$, the total coefficient of $\tilde{g}_t$ on the left-hand side above is $\lambda^{T-t}+(1-\lambda)\sum_{r=1}^{R-1}\lambda^{T-r-t}=\lambda^{T-R+1-t}$; for $T-R+1\le t\le T$, it is $\lambda^{T-t}+(1-\lambda)\sum_{r=1}^{T-t}\lambda^{T-r-t}=1$. Therefore

$$\sum_{t=1}^{T-R}\lambda^{T-R+1-t}\tilde{g}_t+\sum_{t=T-R+1}^{T}\tilde{g}_t\le(a\,\|\theta\|^2+D_1)(R-R\lambda+\lambda).$$

Now $\tilde{g}_t\ge-\psi_t(\theta)\ge-(\|\theta\|\,\|x_t\|+|y_t|)^2\ge-(\|\theta\|+1)^2$, and so

$$\begin{aligned}\sum_{t=T-R+1}^{T}\tilde{g}_t &\le (a\,\|\theta\|^2+D_1)(R-R\lambda+\lambda)+(\|\theta\|+1)^2\sum_{t=1}^{T-R}\lambda^{T-R+1-t} \\ &\le (a\,\|\theta\|^2+D_1)(R-R\lambda+\lambda)+\frac{(\|\theta\|+1)^2}{1-\lambda}.\end{aligned}$$

Divide by R and use $g_t\le\tilde{g}_t+D_2\lambda^R$ (by Proposition 6). □

Choosing appropriate λ and R allows us to bound the right-hand side of (16).

Proposition 9. *For every $\varepsilon>0$ and $a>0$ there is $\lambda_0\equiv\lambda_0(\varepsilon,a,d)<1$ such that for every $\lambda_0<\lambda<1$ there is $R_0\equiv R_0(\varepsilon,a,d,\lambda)\ge 1$ such that for every $R\ge R_0$ the R-windowed λ-discounted a-forward algorithm yields*

$$\frac{1}{R}\sum_{t=T-R+1}^{T}[\psi_t(\theta_t)-\psi_t(\theta)]\le\varepsilon(1+\|\theta\|^2), \tag{17}$$

$$\frac{1}{T}\sum_{t=1}^{T}[\psi_t(\theta_t)-\psi_t(\theta)]\le\varepsilon(1+\|\theta\|^2), \tag{18}$$

for every $T\ge R$, every $\theta\in\mathbb{R}^d$, and every sequence $(x_t,y_t)_{t\ge1}$ in $\mathbb{R}^d\times\mathbb{R}$ with $\|x_t\|\le 1$ and $|y_t|\le 1$ for all t.

Proof. The right-hand side of (16) is

$$\leq \left(D_1(1-\lambda) + \frac{D_1}{R} + \frac{2}{R(1-\lambda)} + D_2\lambda^R \right) + \|\theta\|^2 \left(a(1-\lambda) + \frac{a}{R} + \frac{2}{R(1-\lambda)} \right)$$

(use $\lambda/R \leq 1/R$ and $(\|\theta\|+1)^2 \leq 2\|\theta\|^2 + 2$). First, take $1/4 \leq \lambda_0 < 1$ close enough to 1 so that $a(1-\lambda_0) \leq \varepsilon/4$ and $D_1(a, \lambda_0, d) \cdot (1-\lambda_0) \leq \varepsilon/4$ (recall formula (12) for D_1 and use $\lim_{x\to 0^+} x \ln x = 0$). Then, given $\lambda \in [\lambda_0, 1)$, take $R_0 \geq 1$ large enough so that $a/R_0 \leq \varepsilon/4$, $D_1(a, \lambda, d)/R_0 \leq \varepsilon/4$, $2/(R_0(1-\lambda)) \leq \varepsilon/4$, and $D_2(a, \lambda, d)\lambda^{R_0} \leq \varepsilon/4$. This shows (17) for every $T \geq 1$.

In particular, for $T' < R$ we get $(1/R)\sum_{t=1}^{T'} [\psi_t(\theta_t) - \psi_t(\theta)] \leq \varepsilon(1 + \|\theta\|^2)$ (because $(x_t, y_t, \theta_t) = (0, 0, 0)$ for all $t \leq 0$). For $T \geq R$, add up the inequalities (17) for the disjoint blocks of size R that end at $t = T$, together with the above inequality for the initial smaller block of size $T' < R$ when T is not a multiple of R, to get $(1/R)\sum_{t=1}^{T} [\psi_t(\theta_t) - \psi_t(\theta)] \leq \lceil T/R \rceil \varepsilon(1 + \|\theta\|^2) \leq 2(T/R)\varepsilon(1 + \|\theta\|^2)$. Replacing ε with $\varepsilon/2$ yields (18). □

Remark. Similar arguments show that, for $\lambda_0 \leq \lambda < 1$, the discounted average is also small:

$$\frac{1-\lambda}{1-\lambda^T} \sum_{t=1}^{T} \lambda^{T-t} [\psi_t(\theta_t) - \psi_t(\theta)] \leq \varepsilon(1 + \|\theta\|^2).$$

Proposition 9 yields the main result of this section, Theorem 2.

Proof of Theorem 2. Use Proposition 9 (with, say, $a = 1$), and rescale everything by X and Y appropriately (see the remark before Proposition 5). □

4. Weak calibration

The notion of "weak calibration" was introduced by Kakade and Foster (2004) and Foster and Kakade (2006). The idea is as follows. Given a "test" function $w : C \to \{0, 1\}$ that indicates which forecasts c to consider, let the corresponding score be[18] $S_T^w := \|(1/T)\sum_{t=1}^{T} w(c_t)(a_t - c_t)\|$. It can be shown that if S_T^w is small for every such w, then the calibration score K_T is also small.[19]

Now instead of the discontinuous indicator functions, *weak calibration* requires that S_T^w be small for Lipschitz continuous "weight" functions $w : C \to [0, 1]$. Specifically, let $\varepsilon > 0$ and $L < \infty$. A procedure (i.e., a strategy of the C-player in the calibration game) is (ε, L)-*weakly calibrated* if there is $T_0 \equiv T_0(\varepsilon, L)$ such that

$$S_T^w = \left\| \frac{1}{T} \sum_{t=1}^{T} w(c_t)(a_t - c_t) \right\| \leq \varepsilon \tag{19}$$

holds for every strategy of the A-player, every $T > T_0$, and every weight function $w : C \to [0, 1]$ that is L-Lipschitz (i.e., $\mathcal{L}(w) \leq L$).

The importance of weak calibration is that, unlike regular calibration, it can be guaranteed by deterministic procedures (which are thus leaky): Kakade and Foster (2004) and Foster and Kakade (2006) have proven the existence of deterministic (ε, L)-weakly calibrated procedures. Moreover, as we will show in the next section, weak calibration is essentially equivalent to smooth calibration.

We now provide a deterministic (ε, L)-weakly calibrated procedure that in addition has finite recall and is stationary.

Theorem 10. *For every $\varepsilon > 0$ and $L < \infty$ there exists an (ε, L)-weakly calibrated deterministic procedure that has finite recall and is stationary; moreover, all its forecasts may be taken to lie on a finite grid.*

The proof uses the result of Theorem 2. The basic idea is to use the forecast itself as part of the input to the forecast—which adds a fixed-point construct to the regression. Assume for starters that a and c are one-dimensional, and also that we have only a single weight function w. Consider the online linear regression problem with $x_t = (c_t, w(c_t))$ and $y_t = a_t$. Given the history h_{t-1} up to and including time $t-1$, if we knew the value of c_t then we would get a forecast $\hat{a}_t := \theta_t' x_t$ for

[18] The S_T scores are norms of averages, rather than averages of norms like the K_T scores. "Windowed" versions of the scores may also be considered (with the average taken over the last R periods only; cf. (4)).

[19] Specifically, if $S_T^w \leq \varepsilon$ for all $w : C \to \{0, 1\}$ then $K_T \leq 2m\varepsilon$. Indeed, for each coordinate $i = 1, \ldots, m$, let C_+^i be the set of all c_t such that $\bar{a}_{t,i} > c_{t,i}$, and C_-^i the set of all c_t such that $\bar{a}_{t,i} < c_{t,i}$. Taking w to be the indicator of C_+^i yields $S_T^w = (1/T)\sum_t [\bar{a}_{t,i} - c_{t,i}]_+ \leq \varepsilon$ (where $[z]_+ := \max\{z, 0\}$); similarly, the indicator of C_-^i yields $(1/T)\sum_t [\bar{a}_{t,i} - c_{t,i}]_- \leq \varepsilon$. Adding the two inequalities gives $(1/T)\sum_t |\bar{a}_{t,i} - c_{t,i}| \leq 2\varepsilon$. Since this holds for each one of the m coordinates, it follows that $K_T \leq 2m\varepsilon$.

which, by equation (5), the regret is small. But we do not know c_t, as it is going in fact to be our forecast: that is, we want to choose c_t so that the resulting $\hat{a}_t$ satisfies $\hat{a}_t = c_t$. This requires solving a fixed-point problem (which is possible since the mapping H from c_t to $\hat{a}_t$ is continuous), and indeed yields the desired c_t. Now equation (5) yields, for an appropriate $\varepsilon > 0$,

$$\frac{1}{T}\sum_{t=1}^{T}(a_t - \hat{a}_t)^2 \le \frac{1}{T}\sum_{t=1}^{T}(a_t - \theta' x_t)^2 + \varepsilon$$

for all θ. But $\hat{a}_t = c_t$, and so taking $\theta = (1, \sqrt{\varepsilon})$ gives

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}(a_t - c_t)^2 &\le \frac{1}{T}\sum_{t=1}^{T}(a_t - c_t - \sqrt{\varepsilon}\, w(c_t))^2 + \varepsilon \\ &\le \frac{1}{T}\sum_{t=1}^{T}(a_t - c_t)^2 - 2\sqrt{\varepsilon}\frac{1}{T}\sum_{t=1}^{T} w(c_t)(a_t - c_t) + \varepsilon + \varepsilon\end{aligned}$$

(in the second line we have used $w(c_t) \in [0,1]$). Therefore

$$\frac{1}{T}\sum_{t=1}^{T} w(c_t)(a_t - c_t) \le \sqrt{\varepsilon};$$

together with the similar computation for $\theta = (1, -\sqrt{\varepsilon})$ we get

$$S_T^w = \left|\frac{1}{T}\sum_{t=1}^{T} w(c_t)(a_t - c_t)\right| \le \sqrt{\varepsilon},$$

as desired. To deal with m-dimensional a and c we use m separate online regressions, one for each coordinate; to deal with all the L-Lipschitz weight functions w, we take an appropriate finite grid.

Proof. *(i) Preliminaries.* Without loss of generality assume that $A \subseteq C \subseteq [0,1]^m$ (one can always translate the sets A and C—which does not affect (19)—and rescale them—which just rescales the Lipschitz constant); assume also that $L \ge 1$ (as L increases there are more Lipschitz functions) and $\varepsilon \le 1$.

For every $b \in \mathbb{R}^m$ let $\gamma(b) := \arg\min_{c\in C} \|c - b\|$ be the closest point to b in C (it is well defined and unique since C is a convex compact set); then

$$\|c - b\| \ge \|c - \gamma(b)\| \tag{20}$$

for every $c \in C$ (because

$$\|c - b\|^2 = \|c - \gamma(b)\|^2 + \|b - \gamma(b)\|^2 - 2(b - \gamma(b)) \cdot (c - \gamma(b))$$

and the third term is ≤ 0); moreover, $\gamma(b) = b$ when $b \in C$.

Let $\varepsilon_1 := \varepsilon/(2\sqrt{m})$. Denote by W_L the set of weight functions $w : C \to [0,1]$ with $\mathcal{L}(w) \le L$. By Lemma 18 in Appendix A, there exist d functions $f_1, \ldots, f_d$ in W_L such that for every $w \in W_L$ there is a vector $\varpi \equiv \varpi_w \in [0,1]^d$ with[20]

$$\max_{c\in C}\left|w(c) - \sum_{i=1}^{d}\varpi_i f_i(c)\right| \le \varepsilon_1. \tag{21}$$

Denote $F(c) := (f_1(c), \ldots, f_d(c)) \in [0,1]^d$; thus $\|F(c)\| \le \sqrt{d}$. Without loss of generality we assume that the set $\{f_1, \ldots, f_d\}$ includes the "j-th coordinate function", which maps each $c \in C$ to its j-th coordinate c_j; say, $f_j(c) = c_j$ for $j = 1, \ldots, m$ (thus $d > m$; in fact d is much larger than m).

Let $\varepsilon_2 := \varepsilon/(m + m(1+d)^2 + d^2)$ (where d is given by Lemma 18 in Appendix A, and depends on ε, m, and L) and $\varepsilon_3 := (\varepsilon_2)^2$.

(ii) The function H. Let λ and R be given by Theorem 2 and Proposition 9 for $a = 1$, $X = \sqrt{d}$, $Y = 1$, and $\varepsilon = \varepsilon_3$. For each $j = 1, \ldots, m$ consider the sequence $(x_t, y_t^{(j)})_{t\ge1} = (F(c_t), a_{t,j})_{t\ge1}$ in $\mathbb{R}^d \times \mathbb{R}$, where $a_t \in A$ is determined by the A-player, and $c_t \in C$ is constructed inductively as follows.

[20] Since W_L is compact in the sup norm, there are $f_1, \ldots, f_d \in W_L$ such that for every $w \in W_L$ there is $1 \le i \le d$ with $\max_{c\in C}|w(c) - f_i(c)| \le \varepsilon_1$. Lemma 18 improves on this, in getting a much smaller d by using linear combinations with bounded coefficients.

Let the history be $h_{t-1} = (c_1, a_1, ..., c_{t-1}, a_{t-1})$. For each $c \in \mathbb{R}^m$, let[21]

$$Z_t(c) = I + \sum_{q=1}^{R-1} \lambda^{R-q} x_q x_q' + F(c)F(c)' \in \mathbb{R}^{d\times d},$$

$$v_t^{(j)} = \sum_{q=1}^{R-1} \lambda^{R-q} a_{q,j} x_q \in \mathbb{R}^d,$$

$$H_{t,j}(c) = \left(Z_t(c)^{-1} v_t^{(j)}\right)' F(c) \in \mathbb{R},$$

$$H_t(c) = (H_{t,1}(c), ..., H_{t,m}(c)) \in \mathbb{R}^m$$

(where $x_q = F(c_q)$ for $q < t$). Finally, we extend the function H_t to all of $\mathbb{R}^m$ by putting $H_t(b) := H_t(\gamma(b))$ for every $b \in \mathbb{R}^m$; i.e., we project b to its closest point $\gamma(b)$ in C, and then apply H_t to it.

(iii) Fixed point of H. For every $c \in C$ we have $\left\|v_t^{(j)}\right\| \le \sqrt{d}\lambda/(1-\lambda)$ (since $|a_{q,j}| \le 1$ and $\|x_q\| = \|F(c_q)\| \le \sqrt{d}$), and so $\left\|Z_t(c)^{-1} v_t^{(j)}\right\| \le \sqrt{d}\lambda/(1-\lambda)$ by Lemma 7 ($Z_t(c)$ is positive definite and its eigenvalues are ≥ 1), which finally implies that $\left|H_{t,j}(c)\right| \le \sqrt{d}\lambda/(1-\lambda) \cdot \sqrt{d} = d\lambda/(1-\lambda) =: K$. Therefore the restriction of H_t to the compact and convex set $[-K, K]^m$, which is clearly a continuous function (since, again, $Z_t(c)$ is positive definite and its eigenvalues are ≥ 1), has a fixed point (by Brouwer's fixed-point theorem), which we denote b_t (any fixed point will do)[22]; put $c_t := \gamma(b_t) \in C$. Thus

$$c_t = \gamma(b_t) \quad \text{and} \quad b_t = H_t(b_t) = H_t(c_t).$$

Define $x_t := F(\gamma(b_t)) = F(c_t)$ and $\theta_t^{(j)} := Z_t(c_t)^{-1} v_t^{(j)} \in \mathbb{R}^d$. Then $Z_t(c_t) = I + \sum_{q=1}^{R} \lambda^{R-q} x_q x_q'$, and thus it corresponds to the R-windowed λ-discounted 1-forward algorithm (see (13)). Therefore, for every $j = 1, ..., m$ and every $\theta^{(j)} \in \mathbb{R}^d$ we have by (5)

$$\frac{1}{T}\sum_{t=1}^{T}\left[\psi_t^{(j)}(\theta_t^{(j)}) - \psi_t^{(j)}(\theta^{(j)})\right] \le \varepsilon_3\left(1 + \left\|\theta^{(j)}\right\|^2\right) \tag{22}$$

for all $T \ge T_0 \equiv R$, where $\psi_t^{(j)}(\theta) = (a_{t,j} - \theta' x_t)^2$, and thus $\psi_t^{(j)}(\theta_t^{(j)}) = (a_{t,j} - b_{t,j})^2$ (recall that $b_{t,j} = H_{t,j}(b_t) = \left(\theta_t^{(j)}\right)' F(\gamma(b_t)) = \left(\theta_t^{(j)}\right)' x_t$). Summing over j yields

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{m}\left[\psi_t^{(j)}(\theta_t^{(j)}) - \psi_t^{(j)}(\theta^{(j)})\right] \le \varepsilon_3\left(m + \sum_{j=1}^{m}\left\|\theta^{(j)}\right\|^2\right).$$

Now $\sum_{j=1}^m \psi_t^{(j)}(\theta_t^{(j)}) = \sum_{j=1}^m (a_{t,j} - b_{t,j})^2 = \|a_t - b_t\|^2 \ge \left\|a_t - \gamma(b_t)\right\|^2 = \|a_t - c_t\|^2 = \sum_{j=1}^m (a_{t,j} - c_{t,j})^2$ (by the definition of $\gamma(b_t)$ and (20), since $a_t \in A \subseteq C$), and therefore

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{m}\left[(a_{t,j} - c_{t,j})^2 - \psi_t^{(j)}(\theta^{(j)})\right] \le \varepsilon_3\left(m + \sum_{j=1}^{m}\left\|\theta^{(j)}\right\|^2\right). \tag{23}$$

(iv) Estimating S_T^w. Given a weight function $w \in W_L$, let the vector $\varpi \equiv \varpi_w \in [0,1]^d$ satisfy (21), i.e., $\left|w(c) - \varpi' F(c)\right| \le \varepsilon_1$ for all $c \in C$. Take $u = (u_j)_{j=1,...,m} \in \mathbb{R}^m$ with $\|u\| = 1$. For every $j = 1, ..., m$, take $\theta^{(j)} = e^{(j)} + \varepsilon_2 u_j \varpi \in \mathbb{R}^d$, where $e^{(j)} \in \mathbb{R}^d$ is the j-th unit vector; thus $\left\|\theta^{(j)}\right\| \le 1 + \varepsilon_2 d \le 1 + d$ (since $\varepsilon_2 \le \varepsilon \le 1$). We have

$$(\theta^{(j)})' x_t = (\theta^{(j)})' F(c_t) = c_{t,j} + \varepsilon_2\left(\varpi' F(c_t)\right) u_j$$

(since $f_j(c) = c_j$ for $j \le m$), and hence

$$\begin{aligned}(a_{t,j} - c_{t,j})^2 - \psi_t^{(j)}(\theta^{(j)}) &= (a_{t,j} - c_{t,j})^2 - (a_{t,j} - c_{t,j} - \varepsilon_2\left(\varpi' F(c_t)\right) u_j)^2 \\ &= 2\varepsilon_2\left(\varpi' F(c_t)\right) u_j (a_{t,j} - c_{t,j}) - (\varepsilon_2 \varpi' F(c_t))^2 u_j^2.\end{aligned}$$

[21] A subscript j stands for the j-th coordinate (e.g., $a_{t,j}$ is the j-th coordinate of a_t), whereas a superscript j refers to the j-th procedure (e.g., $v_t^{(j)}$).

[22] There may be more than one fixed point here, in which case we may choose the fixed point at random, and obtain a *randomized* procedure that satisfies everything the deterministic procedure does. Using it yields in Theorem 1 a randomized procedure that is *leaky* smoothly calibrated (cf. Section 2.3).

Summing over $j = 1, \ldots, m$ yields

$$\begin{aligned}\sum_{j=1}^{m}\left[(a_{t,j}-c_{t,j})^2-\psi_t^{(j)}(\theta^{(j)})\right] &= 2\varepsilon_2\left(\varpi' F(c_t)\right)u'(a_t-c_t)-(\varepsilon_2\varpi' F(c_t)^2\,\|u\|^2\\ &\geq 2\varepsilon_2\left(\varpi' F(c_t)\right)u'(a_t-c_t)-(\varepsilon_2)^2 d^2\\ &\geq 2\varepsilon_2 w(c_t)u'(a_t-c_t)-\varepsilon_1\cdot 2\varepsilon_2\,\|u\|\,\|a_t-c_t\|-(\varepsilon_2)^2 d^2\\ &\geq 2\varepsilon_2 w(c_t)u'(a_t-c_t)-2\varepsilon_1\varepsilon_2\sqrt{m}-(\varepsilon_2)^2 d^2\end{aligned}$$

(since: $\|u\|=1$, $|\varpi' F(c)|\leq d$ [the coordinates of ϖ are between -1 and 1 and those of $F(c)$ between 0 and 1], $\|a_t-c_t\|\leq\sqrt{m}$ (since $a_t, c_t\in[0,1]^m$, and recall (21)).

Together with (23) we get (recall that $\varepsilon_3=(\varepsilon_2)^2$ and $\varepsilon_1=\varepsilon/(2\sqrt{m})$):

$$2\varepsilon_2\cdot\frac{1}{T}\sum_{t=1}^{T}w(c_t)u'(a_t-c_t)\leq(\varepsilon_2)^2(m+m(1+d)^2+d^2)+\varepsilon\varepsilon_2;$$

hence, dividing by $2\varepsilon_2$ and recalling that $\varepsilon_2=\varepsilon/(m+m(1+d)^2+d^2))$:

$$u\cdot\frac{1}{T}\sum_{t=1}^{T}w(c_t)(a_t-c_t)\leq\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

Since $u\in\mathbb{R}^m$ with $\|u\|=1$ was arbitrary, the proof of (19) is complete.

(v) Grid. For the "moreover" statement, let $\varepsilon_4:=\varepsilon/(L\sqrt{m}+1)$, and take $D\subseteq C$ to be a finite ε_4-grid in C; i.e., for every $c\in C$ there is $d(c)\in D$ with $\|d(c)-c\|\leq\varepsilon_4$. Replace the forecast c_T obtained above with $\tilde{c}_T:=d(c_T)$; then, for every $a_T\in A$, we have

$$\|w(c_T)(a_T-c_T)-w(\tilde{c}_T)(a_T-\tilde{c}_T)\|\leq L\varepsilon_4\sqrt{m}+\varepsilon_4=\varepsilon.$$

Therefore the score S_T^w changes by at most ε, and so it is at most 2ε. □

5. Smooth calibration

In this section we show (Propositions 13 and 14) that weak calibration and smooth calibration are essentially equivalent (albeit with different constants ε, L). The existence of weakly calibrated procedures (Theorem 10, proved in the previous section) then implies the existence of smoothly calibrated procedures, which proves Theorem 1.

We first show how to go from weak to smooth calibration. When comparing the two scores, we see that the smooth calibration score uses weighted averages rather than sums: $\sum_{s=1}^{T}\Lambda(c_s,c_t)(a_s-c_s)$ is divided by $\sum_{s=1}^{T}\Lambda(c_s,c_t)$. The following useful lemma shows how to bound the latter using the former.

Lemma 11. *There exists a constant $0<\gamma\equiv\gamma_C<\infty$ that depends only on the dimension m and the diameter α of C such that for any L-smoothing weight function Λ, any $c_1,\ldots,c_T\in C$, and any*[23] *$b_1,\ldots,b_T\in C-C$, putting*

$$B_t:=\sum_{s=1}^{T}\Lambda(c_s,c_t)b_s\quad\text{and}\quad W_t:=\sum_{s=1}^{T}\Lambda(c_s,c_t)$$

for all $1\leq t\leq T$, we have

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\|B_t\|}{W_t}\leq\gamma L^{m/2}\left(\frac{1}{T}\max_{1\leq t\leq T}\|B_t\|\right)^{1/2}.$$

Proof. Let $D\subset\mathbb{R}^m$ be an m-dimensional cube with sides of length α that contains C (such a cube exists because the diameter of C is α). Let $D_1,\ldots,D_M$ be a partition of D into disjoint cubes with sides of length $1/(2L\sqrt{m})$; the diameter of each such cube is thus $1/(2L)$, and the number of cubes is $M=\lceil 2\alpha L\sqrt{m}\rceil^m$.

Put $\kappa:=(1/T)\max_{1\leq t\leq T}\|B_t\|$ and $\bar{b}_t\equiv\bar{b}_t^{\Lambda}:=B_t/W_t$. When W_t is large, the inequality $\|B_t\|/W_t\leq\kappa T/W_t$ provides a good bound; we will show that, for a large proportion of indices t, this is indeed the case.

[23] The set $C-C$ consists of all $b=b'-b''$ with $b',b''\in C$.

Given $\eta > 0$ (which will be specified later), call a cube D_j *good* if it contains at least ηT elements of the sequence $c_1, ..., c_T$ (i.e., $|\{t \le T : c_t \in D_j\}| \ge \eta T$), and *bad* otherwise; call an index $t \le T$ *good* if c_t belongs to some good cube D_i, and *bad* otherwise.

If c_t and c_s belong to the same cube D_j then $\|c_s - c_t\| \le \text{diam}(D_j) = 1/(2L)$, and so $1 - \Lambda(c_s, c_t) = \Lambda(c_t, c_t) - \Lambda(c_s, c_t) \le L\|c_s - c_t\| \le 1/2$, which implies that $\Lambda(c_s, c_t) \ge 1/2$. Therefore for every good t we have $W_t = \sum_{s \le T} \Lambda(c_s, c_t) \ge (1/2)\eta T$, and thus $\left\|\bar{b}_t\right\| \le 2T\kappa/(\eta T) = 2\kappa/\eta$, which then gives

$$\frac{1}{T} \sum_{\text{good } t \le T} \|\bar{b}_t\| \le \frac{2\kappa}{\eta}.$$

The number of bad t is less than $M \cdot \eta T$ (because each bad cube contains less than ηT elements of $c_1, ..., c_T$, and there are M cubes). For every s we have $\|b_s\| \le \text{diam}(C) = \alpha$, and so $\bar{b}_t$, which is a weighted average of the b_s, satisfies $\|\bar{b}_t\| \le \alpha$ as well. Thus

$$\frac{1}{T} \sum_{\text{bad } t \le T} \|\bar{b}_t\| \le \frac{1}{T} M \cdot \eta T \cdot \alpha = \alpha \eta M.$$

Adding the last two displayed inequalities and choosing $\eta = \sqrt{(2\kappa)/(\alpha M)}$ yields

$$\frac{1}{T} \sum_{t \le T} \|\bar{b}_t\| \le 2\sqrt{2\kappa\alpha M};$$

recalling that $M = \left\lceil 2\alpha L \sqrt{m} \right\rceil^m$ gives the result, with γ essentially equal to $2^{(m+3)/2} \alpha^{m/2} m^{m/4}$. □

An immediate consequence is that smooth calibration is a weaker requirement than calibration.

Corollary 12. *Calibration implies smooth calibration.*

Proof. We will show that[24] $K_T^\Lambda = O\left(\sqrt{K_T}\right)$ for each fixed L (where K_T^Λ is the Λ-smoothly calibrated score for any L-Lipschitz Λ, and K_T is the regular calibration score). Indeed, for every $c_t \in C$ we have (use $0 \le \Lambda(\cdot, \cdot) \le 1$)

$$\begin{aligned} \frac{1}{T} \left\| \sum_{s=1}^{T} \Lambda(c_s, c_t)(a_s - c_s) \right\| &\le \frac{1}{T} \sum_{s=1}^{T} \Lambda(c_s, c_t) \|(a_s - c_s)\| \\ &\le \frac{1}{T} \sum_{s=1}^{T} \|a_s - c_s\| = K_T. \end{aligned}$$

Now apply Lemma 11 with $b_s = a_s - c_s$ for all s. □

Returning to our proof, we can finally obtain smoothly calibrated procedures with the desired properties, proving Theorem 1.

Proposition 13. *An (ε, L)-weakly calibrated procedure is (ε', L)-smoothly calibrated for $\varepsilon' = \gamma L^{m/2} \varepsilon^{1/2}$ (with the constant $\gamma \equiv \gamma_C$ given by Lemma 11).*

Proof. For any L-Lipschitz smoothing function Λ, Lemma 11 with $b_s = a_s - c_s$ for all s yields

$$K_T^\Lambda \le \gamma L^{m/2} \left(\sup_{w \in W_L} S_T^w \right)^{1/2}$$

(because $\Lambda(\cdot, c_t) \in W_L$ for all c_t; recall the definition (19) of S_T^w in Section 4). Therefore $\sup_{w \in W_L} S_T^w \le \varepsilon$ implies $K_T^\Lambda \le \gamma L^{m/2} \varepsilon^{1/2} = \varepsilon'$ □

Proof of Theorem 1. Apply Theorem 10 and Proposition 13, and recall (Section 2.3) that for deterministic procedures leaks do not matter. □

[24] The notations $f(x) = O(g(x))$, $f(x) = \Omega(g(x))$, and $f(x) = \Theta(g(x))$ mean, as usual, that there are constants $C < \infty$ and $c > 0$ such that for all x we have, respectively, $f(x) \le Cg(x)$, $f(x) \ge cg(x)$, and $cg(x) \le f(x) \le Cg(x)$. In our case x stands for (ε, L); the dimension m is assumed fixed.

As an aside, we now show how to go from smooth to weak calibration.

Proposition 14. *An (ε, L)-smoothly calibrated procedure is (ε', L')-weakly calibrated, where*[25] $\varepsilon' = \Omega\left(\varepsilon^{1/2} L^{m/2}\right)$ *and* $L' = O\left(\varepsilon^{1/2} L^{(m+2)/2}\right)$.

Proof. Let $(D_j)_{j=1,\ldots,M}$ be a partition of $[0,1]^m \supseteq C$ into disjoint cubes with sides $\delta := 1/(L\sqrt{m})$; the diameter of each cube is thus $\delta\sqrt{m} = 1/L$, and the number of cubes is $M = \delta^{-m} = L^m m^{m/2}$. Let $\varepsilon_1 := \sqrt{\varepsilon L^m}$ and $\varepsilon_2 := \varepsilon_1 m^{m/2}/M = \sqrt{\varepsilon/L^m}$. Take $L' := L\varepsilon_1/2 = \sqrt{\varepsilon L^{m+2}}/2$ and $\varepsilon' = \varepsilon_1(1 + \sqrt{m} + \sqrt{m^{m+1}}) = \sqrt{\varepsilon L^m}(1 + \sqrt{m} + \sqrt{m^{m+1}})$.

Fix $a_t, c_t \in C \subseteq [0,1]^m$ for $t = 1, \ldots, T$, and a weight function w in $W_{L'}$. Assume that $K_T^\Lambda \le \varepsilon$ holds for every smoothing function Λ that is L-Lipschitz in the first coordinate; we will show that $S_T^w \le \varepsilon'$ (where K_T^Λ and S_T^w are given by (2) and (19), respectively).

Let $V \subseteq \{1, \ldots, T\}$ be the set of indices t such that the cube D_j that contains c_t includes at least a fraction ε_2 of $c_1, \ldots, c_T$, i.e., $|\{s \le T : c_s \in D_j\}| \ge \varepsilon_2 T$. Then

$$T - |V| = |\{t \le T : t \notin V\}| < M \cdot \varepsilon_2 T = \varepsilon_2 M T, \tag{24}$$

because there are at most M cubes containing less than $\varepsilon_2 T$ points each.

We distinguish two cases.

Case 1: $\max_{t\in V} w(c_t) < \varepsilon_1$. Since $\|a_t - c_t\| \le \sqrt{m}$, we have $\|\sum_{t\in V} w(c_t)(a_t - c_t)\| \le |V| \cdot \varepsilon_1 \cdot \sqrt{m} \le \varepsilon_1 \sqrt{m} T$ (use $|V| \le T$), and $\|\sum_{t\notin V} w(c_t)(a_t - c_t)\| \le (T - |V|) \cdot 1 \cdot \sqrt{m} = \varepsilon_2 M \sqrt{m} T$ (use (24)). Adding and dividing by T yields

$$S_T^w \le (\varepsilon_1 + \varepsilon_2 M)\sqrt{m} = \varepsilon_1 \sqrt{m}(1 + m^{m/2}) < K\varepsilon_1 = \varepsilon'.$$

Case 2: $\max_{t\in V} w(c_t) \ge \varepsilon_1$. Let $s \in V$ be such that $w(c_s) = \max_{t\in V} w(c_t) \ge \varepsilon_1$, and let $R \subseteq V$ be the set of indices r such that c_r lies in the same cube D_j as c_s.

For each r in R, proceed as follows. First, we have $|w(c_s) - w(c_r)| \le L'\|c_s - c_r\| \le L' \cdot \delta\sqrt{m} = L\varepsilon_1/2 \cdot (1/L) = \varepsilon_1/2$, and so

$$w(c_r) \ge w(c_s) - \frac{\varepsilon_1}{2} \ge \varepsilon_1 - \frac{\varepsilon_1}{2} = \frac{\varepsilon_1}{2}. \tag{25}$$

Next, put $w^r(c) := \min\{w(c), w(c_r)\}$ and $\Lambda(c, c_r) := w^r(c)/w(c_r)$ for $r \in R$ (and, for $t \notin R$, put, say, $\Lambda(c, c_t) = 1$ for all c); then $\mathcal{L}(\Lambda(\cdot, c_r)) \le \mathcal{L}(w)/w(c_r) \le L'/(\varepsilon_1/2) = L$, and so, by our assumption

$$\frac{1}{T}\sum_{r\in R} \|\bar{a}_r^\Lambda - c_r^\Lambda\| \le \frac{1}{T}\sum_{t\le T} \|\bar{a}_t^\Lambda - c_t^\Lambda\| \le K_T^\Lambda \le \varepsilon. \tag{26}$$

We will now show that S_T^w is close to an appropriate multiple of $\|\bar{a}_r^\Lambda - c_r^\Lambda\|$, for each r in R. For t in V we have $w(c_t) \le w(c_s)$, and so $0 \le w(c_t) - w^r(c_t) \le w(c_s) - w(c_r) \le \varepsilon_1/2$ (recall (25)), which gives

$$\left\| \sum_{t\in V} \left(w(c_t) - w^r(c_t)\right)(a_t - c_t) \right\| \le |V| \cdot \frac{\varepsilon_1}{2} \cdot \sqrt{m} \le \frac{1}{2}\varepsilon_1 \sqrt{m} T.$$

For $t \notin V$ we have $0 \le w(c_t) - w^r(c_t) \le 1$, and so (recall (24))

$$\left\| \sum_{t\notin V} \left(w(c_t) - w^r(c_t)\right)(a_t - c_t) \right\| \le (T - |V|) \cdot 1 \cdot \sqrt{m} \le \varepsilon_2 M \sqrt{m} T.$$

Adding the two inequalities and dividing by T yields

$$\begin{aligned} S_T^w = \left\| \frac{1}{T}\sum_{t=1}^T w(c_t)(a_t - c_t) \right\| &\le \left\| \frac{1}{T}\sum_{t=1}^T w^r(c_t)(a_t - c_t) \right\| + \sqrt{m}\left(\frac{\varepsilon_1}{2} + \varepsilon_2 M\right) \\ &\le \|\bar{a}_r^\Lambda - c_r^\Lambda\| + \sqrt{m}\left(\frac{\varepsilon_1}{2} + \varepsilon_2 M\right), \end{aligned}$$

because $\sum_{t\le T} w^r(c_t)(a_t - c_t) = \left(\sum_{t\le T} w^r(c_t)\right)(\bar{a}_r^\Lambda - c_r^\Lambda)$ and $\sum_{t\le T} w^r(c_t) \le T$. The set R contains at least $\varepsilon_2 T$ points (these are all the points in the same cube as c_s), i.e., $|R| \ge \varepsilon_2 T$; averaging over all r in the set R and then recalling (26) finally gives

[25] We have not tried to optimize the estimates for ε' and L'.

$$
\begin{aligned}
S_T^w &\le \frac{1}{|R|}\sum_{r\in R}\|\bar{a}_r^\Lambda - c_r^\Lambda\| + \sqrt{m}\left(\frac{\varepsilon_1}{2}+\varepsilon_2 M\right)\\
&\le \frac{1}{\varepsilon_2}\frac{1}{T}\sum_{r\in R}\|\bar{a}_r^\Lambda - c_r^\Lambda\| + \sqrt{m}\left(\frac{\varepsilon_1}{2}+\varepsilon_2 M\right)\\
&\le \frac{1}{\varepsilon_2}\varepsilon + \sqrt{m}\left(\frac{\varepsilon_1}{2}+\varepsilon_2 M\right) = \frac{\varepsilon\sqrt{L^m}}{\sqrt{\varepsilon}} + \varepsilon_1\left(\frac{\sqrt{m}}{2}+m^{(m+1)/2}\right)\\
&= \varepsilon_1\left(1+\frac{\sqrt{m}}{2}+m^{(m+1)/2}\right) < K\varepsilon_1 = \varepsilon',
\end{aligned}
$$

completing the proof. □

6. Nash equilibrium dynamics

In this section we use our results on smooth calibration to obtain dynamics in n-person games that are in the long run close to Nash equilibria most of the time.

A (finite) *game* is given by a finite set of players N, and, for each player $i \in N$, a finite set of actions[26] A^i and a payoff function $u^i : A \to \mathbb{R}$, where $A := \prod_{i\in N} A^i$ denotes the set of action combinations of all players. Let $n := |N|$ be the number of players, $m^i := |A^i|$ the number of pure actions of player i, and $m := \sum_{i\in N} m^i$; also, let U be a bound on payoffs, i.e., $|u^i(a)| \le U$ for all $a \in A$ and $i \in N$. The set of mixed actions of player i is $X^i := \Delta(A^i)$, the unit simplex (i.e., the set of probability distributions) on A^i; we identify the pure actions in A^i with the unit vectors of X^i, and so $A^i \subseteq X^i$. Put $C \equiv X := \prod_{i\in N} X^i$ for the set of mixed-action combinations. The payoff functions u^i are multilinearly extended to X, and thus $u^i : X \to \mathbb{R}$.

For each player i, a combination of mixed actions of the other players $x^{-i} = (x^j)_{j\neq i} \in \prod_{j\neq i} X^j =: X^{-i}$, and $\varepsilon \ge 0$, let $\mathrm{BR}^i_\varepsilon(x^{-i}) := \{x^i \in X^i : u^i(x^i, x^{-i}) \ge \max_{y^i\in X^i} u^i(y^i, x^{-i}) - \varepsilon\}$ denote the set of *ε-best replies* of i to x^{-i}. A (mixed) action combination $x \in X$ is a *Nash ε-equilibrium* if $x^i \in \mathrm{BR}^i_\varepsilon(x^{-i})$ for every $i \in N$; let $\mathrm{NE}(\varepsilon) \subseteq X$ denote the set of Nash ε-equilibria of the game.

A (discrete-time) *dynamic* consists of each player $i \in N$ playing a pure action $a_t^i \in A^i$ at each time period $t = 1, 2, \ldots$; put $a_t = (a_t^i)_{i\in N} \in A$. There is perfect monitoring: at the end of period t all players observe a_t. The dynamic is *uncoupled* (Hart and Mas-Colell, 2003, 2006, 2013) if the play of every player i may depend only on player i's payoff function u^i (and not on the other players' payoff functions). Formally, such a dynamic is given by a mapping for each player i from the history $h_{t-1} = (a_1, \ldots, a_{t-1})$ and his own payoff function u^i into $X^i = \Delta(A^i)$ (player i's choice may be random); we will call such mappings *uncoupled*. Let $x_t^i \in X^i$ denote the mixed action that player i plays at time t, and put $x_t = (x_t^i)_{i\in N} \in X$.

The dynamics we consider are smooth variants of the "calibrated learning" introduced by Foster and Vohra (1997). *Calibrated learning* consists of each player best-replying to calibrated forecasts on the other players' actions; it results in the joint distribution of play converging in the long run to the set of correlated equilibria of the game. Kakade and Foster (2004) defined *publicly calibrated learning*, where each player approximately best-replies to a public weakly calibrated forecast on the joint actions of all players, and proved that most of the time the play is an approximate Nash equilibrium. We consider instead *smooth calibrated learning*, where weak calibration is replaced with the more natural smooth calibration; it amounts to taking calibrated learning and smoothing out both the forecasts and the best replies. Moreover, our forecasts are n-tuples of mixed strategies (in $\prod_i \Delta(A^i)$), rather than correlated mixtures (in $\Delta(\prod_i A^i)$).

Formally, a *smooth calibrated learning* dynamic is given by:

(D1) An (ε_c, L_c)-smoothly calibrated deterministic procedure, which yields at time t a forecast $c_t \in X$ on the distribution of actions of each player.

(D2) For each player $i \in N$ an L_g-Lipschitz ε_g-approximate best-reply mapping $g^i : X \to X^i$; i.e., $g^i(x) \in \mathrm{BR}^i_{\varepsilon_g}(x^{-i})$ for every $x^{-i} \in X^{-i}$.

(D3) Each player runs the procedure in (D1), generating at time t a forecast $c_t \in X$; then each player i plays at period t the mixed action[27] $x_t^i := g^i(c_t) \in X^i$, where g^i is given by (D2). All players observe the action combination $a_t = (a_t^i)_{i\in N} \in A$ that has actually been played, and remember it.

The existence of a deterministic smoothly calibrated procedure in (D1) is given by Theorem 1. For each player i, the payoff function u^i is linear in x^i, and $|u^i(x^i, x^{-i}) - u^i(y^i, x^{-i})| \le \sqrt{m^i}U\|x^i - y^i\|$, and so $\mathcal{L}(u^i) \le \sqrt{m^i}U \le \sqrt{m}U$; the existence

[26] We refer to one-shot choices as "actions" rather than "strategies", the latter term being reserved for repeated interactions.

[27] Thus $\mathbb{P}[a_t = a \mid h_{t-1}] = \prod_{i\in N} x_t^i(a^i)$ for every $a = (a^i)_{i\in N} \in A$, where h_{t-1} is the history and $x_t^i(a^i)$ is the probability that $x_t^i \in \Delta(A^i)$ assigns to the pure action $a^i \in A^i$.

of Lipschitz approximate best-reply mappings in (D2) is then given by Lemma 19 in Appendix A (in particular, for ε_g and L_g such that $L_g \geq \nu_m(\sqrt{m}U/\varepsilon_g)^{m+1}$).

Since for each player i the approximate best reply condition in (D2) makes use *only* of player i's payoff function u^i, we can without loss of generality choose g^i so as to depend only on u^i, which makes the dynamic *uncoupled* (see above).

Our result is:

Theorem 15. *Fix the finite set of players N, the finite action spaces A^i for all $i \in N$, and the payoff bound $U < \infty$. For every $\varepsilon > 0$, any smooth calibrated learning dynamic with appropriate parameters[28] is an uncoupled dynamic that satisfies*

$$\liminf_{T\to\infty} \frac{1}{T} |\{t \leq T : x_t \in \mathrm{NE}(\varepsilon)\}| \geq 1 - \varepsilon \quad \text{(a.s.)}$$

for every finite game with payoff functions $(u^i)_{i\in N}$ that are bounded by U (i.e., $|u^i(a)| \leq U$ for all $i \in N$ and $a \in A$).

The idea of the proof is as follows. First, assume that the forecasts c_t are in fact calibrated (rather than just smoothly calibrated) and, moreover, that they are calibrated with respect to the mixed plays x_t (rather than with respect to the actual plays a_t). Because x_t is given by a fixed function of c_t, namely, $x_t = g(c_t) \equiv (g^i(c_t))_{i\in N}$, the sequence of mixed plays in those periods when the forecast was a certain c is the constant sequence $g(c), \ldots, g(c)$, whose average is $g(c)$, and calibration then implies that $g(c)$ must be close to c (most of the time, i.e., for forecasts that appear with positive frequency). But we have only smooth calibration; however, because g is a continuous function, if c and $g(c)$ are far from one another then so are c' and $g(c')$ for any c' close to c, and so the average of such $g(c')$ is also far from c, contradicting smooth calibration. Thus, most of the time $g(c_t)$ is close to c_t, and hence $g(g(c_t))$ is close to $g(c_t)$ (because g is continuous)—which says that $g(c_t)$ is close to an approximate best reply to itself, i.e., $g(c_t)$ is an approximate Nash equilibrium. Finally, an appropriate use of a strong law of large numbers shows that if the actual plays a_t are (smoothly) calibrated then so are their expectations, i.e., the mixed plays x_t. Two crucial features of our dynamic—which are needed to get Nash equilibria, and cannot be obtained with standard, probabilistic, calibration—are, first, that all players always have the same forecast, and second, that (smooth) calibration is preserved despite the fact that the actions depend on the forecasts (leakiness).

Proof. This proof goes along similar lines to the proof of Kakade and Foster (2004) for publicly calibrated dynamics (which is the only other calibration-based Nash dynamic to date[29]).

Recall that $m^i := |A^i|$ and $m := \sum_{i\in N} m^i$, and so $X \subset [0,1]^m$. Put $g(c) := (g^i(c))_{i\in N}$ for every $c \in X$; thus $g : X \to X$ is a Lipschitz function with $\mathcal{L}(g) \leq nL_g$ (because $\mathcal{L}(g^i) \leq L_g$ for each i).

Take Λ to be the L_c-tent smoothing function: $\Lambda(c', c) = [1 - L_c\|c' - c\|]_+$ for all $c, c' \in X$.

For each period t, let $c_t \in X$ be the *forecast*, $x_t = g(c_t) \in X$ the *behavior* (i.e., mixed actions), and $a_t \in A$ the realized pure *actions* (c_t, x_t, and a_t all depend on the history). We divide the proof into the following steps: (i) smoothed average actions $\bar{a}_t^\Lambda$ and forecasts $\bar{c}_t^\Lambda$ are close (by smooth calibration); (ii) smoothed average actions $\bar{a}_t^\Lambda$ and behaviors $\bar{x}_t^\Lambda$ are close (by the law of large numbers); (iii) forecasts c_t and behaviors x_t are close (because smoothing had little effect there); (iv) behaviors x_t are close to Nash equilibria. Finally, (v) shows how to tweak the parameters to get the desired result.

(i) Smoothed average actions and smoothed forecasts are close.

Let T_0 be such that the smooth calibration score $K_T^\Lambda \leq \varepsilon_c$ for all $T > T_0$, i.e.,

$$\frac{1}{T}\sum_{t=1}^{T} \|\bar{a}_t^\Lambda - c_t^\Lambda\| \leq \varepsilon_c \tag{27}$$

for all $T > T_0$.

(ii) Smoothed average actions and smoothed average behaviors are close.

Let $D \subset X$ be a finite ε_1-grid of X. For each $d \in D$ we have $\mathbb{E}[\Lambda(c_s, d)a_s \mid h_{s-1}] = \Lambda(c_s, d)x_s$ (given h_{s-1}, only a_s is random, and its conditional expectation is $\mathbb{E}[a_s \mid h_{s-1}] = g(c_s) = x_s$). By the Strong Law of Large Numbers for Dependent Random Variables (see Loève, 1978, Theorem 32.1.E: $(1/T)\sum_{s=1}^{T}(X_s - \mathbb{E}[X_s|h_{s-1}]) \to 0$ as $T \to \infty$ a.s., for random variables X_s that are, in particular, uniformly bounded; note that there are finitely many $d \in D$) we get

$$\lim_{T\to\infty} \frac{1}{T}\sum_{s=1}^{T} \Lambda(c_s, d)(a_s - x_s) = 0 \quad \text{for all } d \in D \quad \text{(a.s.)}. \tag{28}$$

[28] Such as those given in (33).

[29] Recall footnote 4.

Thus, for each one of the (almost all) infinite histories h_∞ where (28) holds, there is a finite $T_1 \equiv T_1(h_\infty)$ such that $(1/T)\left\|\sum_{t=1}^{T}\Lambda(c_s,d)(a_s-x_s)\right\| \le \varepsilon_1$ for all $T > T_1$ and all $d \in D$. Now for every $c \in X$ there is $d \in D$ with $\|d-x\| \le \varepsilon_1$, and so $|\Lambda(c_s,d)-\Lambda(c_s,c)| \le L_c\|c-d\| \le L_c\varepsilon_1$; together with[30] $\|a_s-x_s\| \le \sqrt{m}$ it follows that

$$\frac{1}{T}\left\|\sum_{s=1}^{T}\Lambda(c_s,c)(a_s-x_s)\right\| \le (1+\sqrt{m}L_c)\varepsilon_1 \quad \text{for all } T > T_1 \text{ and all } c \in X.$$

Taking in particular $c = c_t$ for all $t \le T$, and then applying Lemma 11 to the set $[0,1]^m$ and $b_s = a_s - x_s$, yields

$$\frac{1}{T}\sum_{t=1}^{T}\|\bar{a}_t^\Lambda - \bar{x}_t^\Lambda\| \le \gamma_m L_c^{m/2}\sqrt{(1+\sqrt{m}L_c)\varepsilon_1} =: \varepsilon_2, \tag{29}$$

where the constant γ_m depends only on m.

(iii) Behaviors and forecasts are close.

Because $\Lambda(c_s,c_t) > 0$ only when $\|c_s - c_t\| < 1/L_c$, it follows that c_t^Λ, as a weighted average of such c_s, satisfies $\|c_t^\Lambda - c_t\| < 1/L_c$. Moreover, $\|x_s - x_t\| = \|g(c_s) - g(c_t)\| \le nL_g/L_c$, and so $\|\bar{x}_t^\Lambda - x_t\| \le nL_g/L_c$, which together with (27) and (29) gives

$$\frac{1}{T}\sum_{t=1}^{T}\|x_t - c_t\| \le \varepsilon_c + \varepsilon_2 + \frac{1}{L_c} + \frac{nL_g}{L_c} =: \varepsilon_3 \tag{30}$$

for almost every infinite history h_∞ and for every $T > \max\{T_0, T_1(h_\infty)\}$.

(iv) Behaviors are close to Nash equilibria. From (30) it immediately follows that, for every $\varepsilon_4 > 0$,

$$\frac{1}{T}\left|\{t \le T : \|g(c_t) - c_t\| > \varepsilon_4\}\right| \le \frac{1}{\varepsilon_4}\frac{1}{T}\sum_{t=1}^{T}\|(g(c_t) - c_t)\| \le \frac{\varepsilon_3}{\varepsilon_4}. \tag{31}$$

If $\|g(c_t) - c_t\| \le \varepsilon_4$ then

$$\left\|g^i(x_t) - x_t^i\right\| = \left\|g^i(g(c_t)) - g^i(c_t)\right\| \le L_g\varepsilon_4,$$

and so

$$\begin{aligned} u^i(x_t) &\ge u^i(g^i(x_t), x_t^{-i}) - \sqrt{m^i}U\left\|g^i(x_t) - x_t^i\right\| \\ &\ge \max_{y^i \in \Delta(A^i)} u^i(y^i, x_t^{-i}) - \varepsilon_g - \sqrt{m^i}UL_g\varepsilon_4 \end{aligned}$$

(for the second inequality we have used $g^i(x) \in \mathrm{BR}^i_{\varepsilon_g}(x^{-i})$). Therefore $\|g(c_t) - c_t\| \le \varepsilon_4$ implies that $x_t \in \mathrm{NE}(\varepsilon_5)$, where

$$\varepsilon_5 := \varepsilon_g + \sqrt{m}UL_g\varepsilon_4 \tag{32}$$

(recall that $m = \sum_i m^i \ge m^i$), and so, from (30) and (31) we get

$$\frac{1}{T}\left|\{t \le T : x_t \notin \mathrm{NE}(\varepsilon_5)\}\right| \le \frac{\varepsilon_3}{\varepsilon_4}$$

for all large enough T, for almost every infinite history.

(v) Tweaking the parameters. To bound both $\varepsilon_3/\varepsilon_4$ and ε_5 by, say, 3ε, one may take, for instance (see (29)–(32) and recall Lemma 19 in Appendix A),

$$\begin{aligned} &\varepsilon_g = \varepsilon, \quad L_g = \nu_m\left(\frac{\sqrt{m}U}{\varepsilon}\right)^{m+1}, \\ &\varepsilon_4 = \frac{2\varepsilon}{\sqrt{m}UL_g}, \\ &\varepsilon_c = \varepsilon\varepsilon_4, \quad L_c = \frac{1+nL_g}{\varepsilon\varepsilon_4}, \\ &\varepsilon_2 = \varepsilon\varepsilon_4, \quad \varepsilon_1 = \frac{\varepsilon_2^2}{\gamma_m^2 L_c^m(1+\sqrt{m}L_c)}, \end{aligned} \tag{33}$$

because we then get $\varepsilon_5 = \varepsilon + 2\varepsilon = 3\varepsilon$ and $\varepsilon_3 = \varepsilon\varepsilon_4 + \varepsilon\varepsilon_4 + \varepsilon\varepsilon_4 = 3\varepsilon\varepsilon_4$. □

[30] Because $a_s, x_s \in [0,1]^m$.

Remarks. *(a) Nash dynamics.* Uncoupled dynamics where Nash ε-equilibria are played $1-\varepsilon$ of the time were first proposed by Foster and Young (2003), followed by Kakade and Foster (2004), Foster and Young (2006), Hart and Mas-Colell (2006), Germano and Lugosi (2007), Young (2009), Babichenko (2012), and others (see also Remark (h) below).

(b) Coordination. All players need to coordinate before playing the game on the smoothly calibrated procedure that they will run; thus, at every period t they all generate the same forecast t. By contrast, in the original calibrated learning dynamic of Foster and Vohra (1997)—which leads to correlated equilibria—every player may use his own calibrated procedure.

This fits the so-called *Conservation Coordination Law* for game dynamics, which says that some form of "coordination" must be present, either in the limit static equilibrium concept (such as correlated equilibrium) or in the dynamic leading to it (such as Nash equilibrium dynamics). See Hart and Mas-Colell (2000, 2001, 2003, footnote 19) and Hart (2005, footnote 19).

(c) Deterministic calibration. In order for all the players to generate the same forecasts, it is not enough that they all use the same procedure; in addition, the forecasts must be deterministic (otherwise the randomizations, which are carried out independently by the players, may lead to different actual forecasts). This is the reason that we use smoothly calibrated procedures rather than fully calibrated ones (cf. Oakes, 1985 and Foster and Vohra, 1998).

(d) Leaky calibration. One may use a common *randomized* smoothly calibrated procedure, provided that the randomizations are carried out publicly (i.e., they must be leaked!). Alternatively, a "central bureau of statistics" may be used each period to provide the forecast to all the players.

(e) Finite memory. In (D1) one may use a smoothly calibrated procedure that has finite recall and is stationary (see Theorem 1). However, while in the calibration game of Section 5 both the actions a_t and the forecasts c_t are monitored and thus become part of the recall window, in the n-person game only a_t is monitored (while the forecast c_t is computed by each player separately, but is *not* played). Therefore, in order to run the calibrated procedure, in the n-person game each player needs to remember at time T, in addition to the last R action combinations $a_{T-R}, \ldots, a_{T-1}$, also the last R forecasts $c_{T-R}, \ldots, c_{T-1}$. "Finite recall" of size R in the calibration procedure therefore becomes "finite memory" of size $2R$ in the game dynamic: the memory contains R elements of A and R elements of[31] C.

Alternatively, to get finite recall rather than finite memory one may introduce an artificial player, say, player 0, with action set $A^0 := X$ and constant payoff function $u^0 \equiv 0$, who plays at each period t the forecast c_t, i.e., $a_t^0 = c_t$; this way the forecasts become part of the recall of all players.

(f) Forecasting joint play. In (D1) one may use a procedure that forecasts the joint play: the forecasts c_t lie in $\Delta(A)$, rather than in $\prod_i \Delta(A^i))$ (the dimension is then larger, $\prod_i m^i$ instead of $\sum_i m^i$). The approximate best reply functions g^i can then be defined over $\Delta(A)$, and the proof carries through essentially without change. Thus most of the time the play is close to Nash equilibrium, despite the fact that the forecasts are allowed to be correlated; in fact, the forecasts turn out to be close to being independent (because $g(c_t) \in X$ is independent, and c_t is close to $g(c_t)$).

(g) Separate forecasts. One cannot simplify (D1) by replacing the forecasting procedure that yields $c_t = (c_t^i)_{i \in N} \in X$ with n separate forecasting procedures that yield $c_t^i \in X^i$ for each $i \in N$, because then behaviors x_t and forecasts c_t need no longer be close (in part (iii) of the proof, when c_s^i is now close to c_t^i for some i, it does not follow that c_s^j and c_t^j for $j \neq i$ are also close, and so neither are $g(c_s)$ and $g(c_t)$).

(h) Continuous approximate best reply. In (D2) one may take the functions g^i to be continuous rather than Lipschitz and carry the proof with the modulus of continuity instead of the Lipschitz bound (for uncoupledness one would need to require uniform equicontinuity).

(i) Exhaustive search. Dynamics that perform exhaustive search can also be used to get the result of Theorem 15.[32] Take for instance a finite grid on X, say, $D = \{d_1, \ldots, d_M\} \subset X$, that is fine enough so that there always is a pure Nash ε-equilibrium on the grid. Let the dynamic go over the points $d_1, d_2, \ldots$ in sequence until the first time that $d_T^i \in \mathrm{BR}_\varepsilon^i(d_T^{-i})$ for all i, following which d_T is played forever. This is implemented by having for every player i a distinct action $a_0^i \in A^i$ that is played at time t only when $d_t^i \in \mathrm{BR}_\varepsilon^i(d_t^{-i})$ (otherwise a different action is played); once the action combination $a_0 = (a_0^i)_{i \in N} \in A$ is played, say, at time T, each player i plays d_T^i at all $t > T$. This dynamic is uncoupled (each player only considers $\mathrm{BR}_\varepsilon^i$) and has memory of size 2 (i.e., 2 elements of X): for $t \leq T$, it consists of d_{t-1} and a_{t-1} (the last checked point and the last played action combination); for $t > T$, it consists of d_T and a_0. Of course, all players need to coordinate before playing the game on the sequence $d_1, d_2, \ldots, d_M$ and the action combination a_0.

(j) Continuous action spaces. The result of Theorem 15 easily extends to continuous action spaces and approximate *pure* Nash equilibria. Assume that for each player $i \in N$ the set of actions A^i is a convex compact subset of some Euclidean space (such games arise, for instance, from exchange economies where the actions are net trades; see, e.g., Hart and Mas-Colell, 2015). Thus $A = \prod_{i \in N} A^i$ is a compact convex set in some Euclidean space, say, $\mathbb{R}^m$.

For every $\varepsilon \geq 0$, the set of *pure ε-best replies* of player i to $a^{-i} \in a^{-i}$ is $\mathrm{PBR}_\varepsilon^i(a^{-i}) := \{a^i \in A^i : u^i(a^i, a^{-i}) \geq \max_{b^i \in A^i} u^i(b^i, a^{-i}) - \varepsilon\}$. An action combination $a \in A$ is a *pure Nash ε-equilibrium* if $a^i \in \mathrm{PBR}_\varepsilon^i(a^{-i})$ for every $i \in N$; let $\mathrm{PNE}(\varepsilon) \subseteq A$ denote the set of pure Nash ε-equilibria.

[31] For a similar transition from finite recall to finite memory, see Theorem 7 in Hart and Mas-Colell (2006).

[32] We thank Yakov Babichenko for suggesting this.

Smooth calibrated learning is defined as above, except that now the approximate best replies are pure actions (the play is $a_t = g(c_t)$, and it is monitored by all players). Our result here is:

Theorem 16. *Fix the finite set of players N, the convex compact action spaces A^i for all $i \in N$, and the Lipschitz bound $L < \infty$. For every $\varepsilon > 0$, and any smooth calibrated learning dynamic with appropriate parameters, there is $T_0 \equiv T_0(\varepsilon, L)$ such that for every $T \geq T_0$,*

$$\frac{1}{T} \left| \{ t \leq T : a_t \in \mathrm{PNE}(\varepsilon) \} \right| \geq 1 - \varepsilon$$

for every game with payoff functions $(u^i)_{i \in N}$ that are L-Lipschitz (i.e., $\mathcal{L}(u^i) \leq L$) and quasi-concave in one's own action (i.e., $u^i(a^i, a^{-i})$ is quasi-concave in $a^i \in A^i$ for every $a^{-i} \in A^{-i}$), for all $i \in N$.

Proof. We now have $A^i = X^i$ and $a_t = x_t = g(c_t)$, and everything is deterministic. Proceed as in the proof of Theorem 15, skipping part (ii) (the use of the Law of Large Numbers) and taking $\varepsilon_1 = 0$ and $T_1 = 0$. □

(k) Reaction function and fixed points. The proof of Theorem 15 shows that in the leaky calibration game, if the A-player uses a stationary strategy given by a Lipschitz "reaction" function g (i.e., he plays $g(c_t)$ at time t), then smooth calibration implies that the forecasts c_t are close to fixed points of g most of the time.

Appendix A

Let C be a compact subset of $\mathbb{R}^m$ and let $\varepsilon > 0$. A *maximal 2ε-net* in C is a maximal collection of points $z_1, \ldots, z_K \in C$ such that $\|z_k - z_j\| \geq 2\varepsilon$ for all $k \neq j$; maximality implies $\cup_{k=1}^K B(z_k, 2\varepsilon) \supseteq C$. Let $\alpha_k(x) := [3\varepsilon - \|x - z_k\|]_+$, and put $\bar{\alpha}(x) := \sum_{k=1}^K \alpha_k(x)$. For every $x \in C$ we have $0 \leq \alpha_k(x) \leq 3\varepsilon$ and $\bar{\alpha}(x) \geq \varepsilon$ (since $\alpha_k(x) \geq \varepsilon$ when $x \in B(z_k, 2\varepsilon)$, and the union of these balls covers C). Finally, define $\beta_k(x) := \alpha_k(x)/\bar{\alpha}(x)$.

Lemma 17. *The functions $(\beta_k)_{1 \leq k \leq K}$ satisfy the following properties:*

- **(i)** *$\beta_k(x) \geq 0$ for all $x \in C$ and all k.*
- **(ii)** *$\sum_{k=1}^K \beta_k(x) = 1$ for all $x \in C$.*
- **(iii)** *$\beta_k(x) = 0$ for all $x \notin B(z_k, 3\varepsilon)$.*
- **(iv)** *For each $x \in C$ there are at most[33] 4^m indices k such that $\beta_k(x) > 0$.*
- **(v)** *$\mathcal{L}(\beta_k) \leq 4^{m+2}/\varepsilon$ for every k.*

Proof. (i) and (ii) are immediate. For (iii), we have $\beta_k(x) > 0$ iff $\alpha_k(x) > 0$ iff $\|x - z_k\| < 3\varepsilon$. This implies that $B(z_k, \varepsilon) \subseteq B(x, 4\varepsilon)$. The open balls of radius ε with centers at z_k are disjoint (because $\|z_k - z_j\| \geq 2\varepsilon$ for $k \neq j$), and so there can be at most 4^m such balls included in $B(x, 4\varepsilon)$ whose volume is 4^m times larger; this proves (iv). For every $x, y \in C$:

$$\begin{aligned}
|\beta_k(x) - \beta_k(y)| &\leq \left| \frac{\alpha_k(x)}{\bar{\alpha}(x)} - \frac{\alpha_k(y)}{\bar{\alpha}(x)} \right| + \left| \frac{\alpha_k(y)}{\bar{\alpha}(x)} - \frac{\alpha_k(y)}{\bar{\alpha}(y)} \right| \\
&\leq \frac{1}{\bar{\alpha}(x)} |\alpha_k(x) - \alpha_k(y)| + \frac{\alpha_k(y)}{\bar{\alpha}(x)\bar{\alpha}(y)} \sum_{j=1}^K |\alpha_j(x) - \alpha_j(y)| \\
&\leq \frac{1}{\varepsilon} \|x - y\| + \frac{3\varepsilon}{\varepsilon \cdot \varepsilon} 2 \cdot 4^m \|x - y\| \leq \frac{4^{m+2}}{\varepsilon} \|x - y\|
\end{aligned}$$

(since $\bar{\alpha}(x) \geq \varepsilon$, $\alpha_k(x) \leq 3\varepsilon$, and there are at most $2 \cdot 4^m$ indices j where neither $\alpha_j(x)$ nor $\alpha_j(y)$ vanishes); this proves (v). □

Thus, the functions $(\beta_k)_{1 \leq k \leq K}$ constitute a *Lipschitz partition of unity* that is subordinate to the maximal 2ε-net $z_1, \ldots, z_K$. Next, we obtain a basis for the Lipschitz functions on C.

Lemma 18. *Let W_L be the set of functions $w : C \to [0, 1]$ with $\mathcal{L}(w) \leq L$. Then for every $\varepsilon > 0$ there exist d functions $f_1, \ldots, f_d \in W_L$ such that for every $w \in W_L$ there is a vector $\varpi \equiv \varpi_w \in [0, 1]^d$ satisfying*

$$\max_{x \in C} \left| w(x) - \sum_{i=1}^d \varpi_i f_i(x) \right| < \varepsilon.$$

Moreover, one can take $d = \mathrm{O}(L^m/\varepsilon^{m+1})$.

[33] We have not tried to get the best bounds in (iv) and (v); indeed, they may be easily reduced.

Proof. Put $\varepsilon_1 := \varepsilon/(3L)$. Let $z_1, ..., z_K$ be a maximal $2\varepsilon_1$-net on C, and let $\beta_1, ..., \beta_K$ be the corresponding Lipschitz partition of unity given by Lemma 17 (for ε_1).

Given $w \in W_L$, let $v(x) := \sum_{k=1}^{N} w(z_k)\beta_k(x)$; then $w(z_k) \in [0,1]$ and we have

$$|w(x) - v(x)| = \left| \sum_{k=1}^{N} (w(x) - w(z_k))\, \beta_k(x) \right| \le \sum_{k:\beta_k(x)>0} \beta_k(x)\, |w(x) - w(z_k)|$$
$$\le \sum_{k:\beta_k(x)>0} \beta_k(x) 3\varepsilon_1 L = 3\varepsilon_1 L,$$

since $\beta_k(x) > 0$ implies $\|x - z_k\| < 3\varepsilon_1$ and thus $|w(x) - w(z_k)| \le L\,\|x - z_k\| \le L \cdot 3\varepsilon_1$ (because $\mathcal{L}(w) \le L$).

Now $\mathcal{L}(\beta_k) \le 4^{m+2}/\varepsilon_1$ by (v) of Lemma 17; we thus replace each β_k by the sum of $Q = \lceil 4^{m+2}/(\varepsilon_1 L) \rceil$ identical copies of $(1/Q)\beta_k$—denote them $f_{k,1}, ..., f_{k,Q}$—which thus satisfy $\mathcal{L}(f_{k,q}) = (1/Q)\mathcal{L}(\beta_k) \le L$, and so

$$\left| w(x) - \sum_{k=1}^{K} \sum_{q=1}^{Q} w(z_k) f_{k,q}(x) \right| = |w(x) - v(x)| \le 3\varepsilon_1 L = \varepsilon.$$

The $d = KQ$ functions $(f_{k,q})_{1\le k\le K, 1\le q\le Q}$ yield our result.

Finally, $K = O(\varepsilon_1^{-m})$ (because C contains the K disjoint open balls of radius ε_1 centered at the z_k) and $Q \le 4^{m+2}/(\varepsilon_1 L) + 1$, and so $d = KQ = O(\varepsilon_1^{-m-1} L^{-1}) = O(\varepsilon^{-m-1} L^m)$. □

In the game setup we construct ε-best reply functions that are Lipschitz. The following lemma applies when the action spaces are finite (as in Theorem 15), and also when they are continuous (as in Theorem 16). In the former $C = X = \prod_{i\in N} X^i$ where $X^i = \Delta(A^i)$, and in the latter $C = X = A = \prod_{i\in N} A^i$, and the set $\Delta(A^i)$ is identified with A^i; also, $\mathrm{BR}^i_\varepsilon$ stands for $\mathrm{PBR}^i_\varepsilon$, the set of *pure* ε-best replies.

Lemma 19. *Assume that for each player $i \in N$ the function $u^i : X \to \mathbb{R}$ is a Lipschitz function with $\mathcal{L}(u^i) \le L$, and $u^i(\cdot, c^{-i})$ is quasi-concave on X^i for every fixed $c^{-i} \in X^{-i}$. Then for every $\varepsilon > 0$ there is a Lipschitz function $g^i : X \to X^i$ such that $g^i(c) \in \mathrm{BR}^i_\varepsilon(c^{-i})$ for all $c \in X$, and $\mathcal{L}(g^i) \le \nu_m (L/\varepsilon)^{m+1}$ where the constant ν_m depends only on the dimension m.*

Proof. Put $\varepsilon_1 := \varepsilon/(6L)$. Let $z_1, ..., z_K \in C$ be a maximal $2\varepsilon_1$-net on C, and let $\beta_1, ..., \beta_K$ be the subordinated Lipschitz partition of unity given by Lemma 17. For each $i \in N$ and $1 \le k \le K$ take $x^i_k \in \mathrm{BR}^i_0(z^{-i}_k)$, and define $g^i(c) := \sum_{k=1}^{K} \beta_k(c) x^i_k$. Because $\beta_k(c) > 0$ if and only if $\|c - z_k\| < 3\varepsilon_1$, it follows that $x^i_k \in \mathrm{BR}^i_\varepsilon(c^{-i})$ (indeed, for every $y^i \in \Delta(A^i)$ we have $u^i(x^i_k, c^{-i}) > u^i(x^i_k, z^{-i}_k) - 3L\varepsilon_1 \ge u^i(y^i, z^{-i}_k) - 3L\varepsilon_1 > u^i(y^i, c^{-i}) - 6L\varepsilon_1 = \varepsilon$, where we have used $\mathcal{L}(u^i) \le L$ twice, and $x^i_k \in \mathrm{BR}^i_0(z^{-i}_k)$). The set $\mathrm{BR}^i_\varepsilon(c^{-i})$ is convex by the quasi-concavity assumption, and so $g^i(c)$, as an average of such x^i_k, belongs to $\mathrm{BR}^i_\varepsilon(c^{-i})$.

Now $\max_{c\in C} \|c\| \le \sqrt{m}$ (because $C \subseteq [0,1]^m$), and so $\|x_k\| \le \sqrt{m}$ (where $x_k = (x^i_k)_{i\in N}$) for all k, and $K \le (\sqrt{m}/\varepsilon_1)^m$ (because $C \subseteq B(0, \sqrt{m})$ contains the K disjoint open balls of radius ε_1 centered at the points z_k). Therefore the Lipschitz constant of $g(c) = \sum_{k=1}^{K} \beta_k(c) x_k$ satisfies, by Lemma 17 (v), $\mathcal{L}(g) \le \sum_{k=1}^{K} \|x_k\|\, \mathcal{L}(\beta_k) \le (\sqrt{m}/\varepsilon_1)^m \sqrt{m}\, 4^{m+2}/\varepsilon_1 = \nu_m \varepsilon^{-m-1} L^{m+1}$ for $\nu_m = \sqrt{m}^{m+1} 4^{m+2} 6^{m+1}$. □

References

Azoury, K.S., Warmuth, M.K., 2001. Relative loss bounds for on-line density estimation with the exponential family of distributions. Mach. Learn. 43, 211–246.

Babichenko, Y., 2012. Completely uncoupled dynamics and Nash equilibria. Games Econ. Behav. 76, 1–14.

Cesa-Bianchi, N., Lugosi, G., 2006. Prediction, Learning, and Games. Cambridge University Press.

Dawid, A.P., 1982. The well-calibrated Bayesian. J. Amer. Statistical Assoc. 77, 605–613.

Dawid, A.P., 1985. Self-calibrating priors do not exist: comment. J. Amer. Statistical Assoc. 80, 340–341.

Foster, D.P., 1991. Prediction in the worst case. Ann. Statist. 19, 1084–1090.

Foster, D.P., 1999. A proof of calibration via Blackwell's approachability theorem. Games Econ. Behav. 29, 73–78.

Foster, J., 1999. On relative loss bounds in generalized linear regression. In: 12th International Symposium on Fundamentals of Computation Theory. FCT '99, pp. 269–280.

Foster, D.P., Kakade, S.M., 2006. Calibration via regression. In: IEEE Information Theory Workshop 2006.

Foster, D.P., Vohra, R.V., 1997. Calibrated learning and correlated equilibrium. Games Econ. Behav. 21, 40–55.

Foster, D.P., Vohra, R.V., 1998. Asymptotic calibration. Biometrika 85, 379–390.

Foster, D.P., Young, H.P., 2003. Learning, hypothesis testing, and Nash equilibrium. Games Econ. Behav. 45, 73–96.

Foster, D.P., Young, H.P., 2006. Regret testing: learning to play Nash equilibrium without knowing you have an opponent. Theoretical Econ. 1, 341–367.

Germano, F., Lugosi, G., 2007. Global Nash convergence of Foster and Young's regret testing. Games Econ. Behav. 60, 135–154.

Hart, S., 2005. Adaptive heuristics. Econometrica 73, 1401–1430; Chapter 11 of Hart and Mas-Colell (2013).

Hart, S., Mas-Colell, A., 2000. A simple adaptive procedure leading to correlated equilibrium. Econometrica 68, 1127–1150; Chapter 2 of Hart and Mas-Colell (2013).

Hart, S., Mas-Colell, A., 2001. A general class of adaptive strategies. J. Econ. Theory 98, 26–54; Chapter 3 of Hart and Mas-Colell (2013).
Hart, S., Mas-Colell, A., 2003. Uncoupled dynamics do not lead to Nash equilibrium. Amer. Econ. Rev. 93, 1830–1836; Chapter 7 of Hart and Mas-Colell (2013).
Hart, S., Mas-Colell, A., 2006. Stochastic uncoupled dynamics and Nash equilibrium. Games Econ. Behav. 57, 286–303; Chapter 8 of Hart and Mas-Colell (2013).
Hart, S., Mas-Colell, A., 2013. Simple Adaptive Strategies: From Regret-Matching to Uncoupled Dynamics. World Scientific.
Hart, S., Mas-Colell, A., 2015. Markets, correlation, and regret-matching. Games Econ. Behav. 93, 42–54.
Hazan, E., Kakade, S., 2012. (Weak) calibration is computationally hard. In: 25th Annual Conference on Learning Theory, COLT '12. JMLR: Workshop and Conference Proceedings 23, 3.1–3.10.
Kakade, S.M., Foster, D.P., 2004. Deterministic calibration and Nash equilibrium. In: 17th Annual Conference on Learning Theory, COLT '04; J. Comput. Syst. Sci. 74 (2008), 115–130.
Lehrer, E., Solan, E., 2009. Approachability with bounded memory. Games Econ. Behav. 66, 995–1004.
Loève, M., 1978. Probability Theory, vol. II, 4th edition. Springer.
Oakes, D., 1985. Self-calibrating priors do not exist. J. Amer. Statistical Assoc. 80, 339.
Olszewski, W., 2015. Calibration and expert testing. In: Young, H.P., Zamir, S. (Eds.), Handbook of Game Theory, vol. 4. Springer, pp. 949–984.
Papadimitriou, C., 1994. On the complexity of the parity argument and other inefficient proofs of existence. J. Comput. System Sci. 48, 498–532.
Vovk, V., 2001. Competitive on-line statistics. Int. Stat. Rev. 69, 213–248.
Young, H.P., 2009. Learning by trial and error. Games Econ. Behav. 65, 626–643.