Deterministic calibration and Nash equilibrium

Sham M. Kakade, Dean P. Foster

***Abstract*— [1] introduces a deterministic forecasting algorithm that satisfies a relaxed notion of calibration, termed weak calibration, and show that when players best respond to such forecasts, the empirical play converges to the set of convex combinations of approximate Nash equilibria. In this report, we develop intuitive explanations and arguments for the construction of the weakly calibrated forecaster, along with its associated proofs, and its role in ensuring no-regret behavior. In addition to analyzing the theory from the paper, we implement the associated forecast fixed-point algorithm in the context of a toy example, the Rock-Paper-Scissors game, where our empirical results confirm convergence to the unique mixed-strategy Nash equilibrium, thereby validating the theoretical guarantees.**

***Index Terms*— Nash Equilibria; Calibration; Game Theory; Learning in Games**

I. Introduction

A central problem in game theory is whether equilibrium concepts, such as the Nash equilibrium, can be arrived at from reasonable learning rules and dynamics. Despite its foundational role, the Nash equilibrium is often criticized for its lack of plausibility in decentralized learning settings. While correlated equilibria can arise naturally from simple learning protocols, the same is not true for Nash equilibria, where convergence guarantees are rare and often depend on restrictive assumptions or the use of randomized strategies. The most commonly studied learning algorithms are those in which players make predictions about the behavior of their opponents and respond with the best responses, an approach considered rational from a learning perspective. Fictitious Play is a classical example in this category [2].

This paper presents a natural learning process in which players' empirical joint distribution converges to the set of convex combinations of Nash equilibria, while ensuring that their actions are rational in a decentralized sense. The learning setting considered assumes that at each round, players observe a public forecast, a probability distribution over joint actions, that they use to determine their moves. Importantly, players do not make arbitrary predictions; instead, their forecasts are generated by a deterministic algorithm that satisfies a relaxed version of the calibration property. The players then independently calculate their (approximate) best responses to this forecast, without needing to know the payoffs or strategies of others. The forecast is common to all players, ensuring their responses are coordinated solely through the shared prediction. This setup allows for a robust notion of rationality: even if some players deviate from the protocol, any player who follows the learning rule still achieves "no-internal regret", i.e., they would not have gained by consistently switching from one action to another in hindsight. This follows from the result that if a player uses calibrated predictions and responds optimally to the forecast, then they necessarily experience no regret with respect to their own action choices [3].

Calibration, in this context, refers to the property that a player's predicted probabilities should match the observed empirical frequencies over time. A simple and intuitive example is the case of a weather forecaster. Suppose the forecaster announces a 70% chance of rain on multiple days. If the forecasts are well-calibrated, then it should rain approximately 70% of those days. Formally, calibration requires that the forecast probability and the conditional empirical frequency match asymptotically for forecasts that occur often enough [4]. However, deterministic forecasts have been shown not to achieve this in adversarial settings [5], [6], which led previous work to rely on randomized forecasting algorithms to ensure calibration [7]. In contrast, this paper introduces the notion of weak calibration, which relaxes the strict calibration requirement using tools from weak convergence, allowing the design of a deterministic forecasting algorithm that still retains meaningful guarantees. This relaxation ensures that, while exact calibration may not be achievable deterministically, the forecast still aligns with observed frequencies in a smoothed or averaged sense.

Using this weakly calibrated deterministic forecaster, the authors construct a learning process that yields strong convergence properties. Prior work, [8] demonstrated that calibration-based strategies can lead play to converge to the set of correlated equilibria, which is a superset of Nash equilibria. This paper shows that under their deterministic, weakly calibrated learning process, the empirical joint frequency of play converges to the set of convex combinations of Nash equilibria. Moreover, on most rounds, the actual distribution of play is close to some approximate Nash equilibrium, a significant improvement over previous calibration-based learning schemes. The process allows the possibility of switching among equilibria but ensures that play is never far from the equilibrium set. The forecast distribution and the empirical play distribution become increasingly close over time. This means that forecasts not only portray the outcomes accurately in the long run but also match the actual strategies played in most rounds.

In the remainder of this report, we first provide a mathematical formulation of calibration and weak calibration, followed by an example demonstrating the use of randomized rounding to achieve standard calibration. We then discuss

Report written by Srikar Babu Gadipudi, as a part of the EE6417: Incentive-Centered Design course in Jan-May semester 2025, IIT Madras: ee21b138@smail.iitm.ac.in

the setting of publicly calibrated learning, resulting in the convergence result toward Nash equilibria. Subsequently, we present the construction of a deterministic algorithm that achieves weak calibration and analyze its convergence properties. Throughout, we emphasize intuitive explanations over formal mathematical rigor and refer the reader to [1] for detailed proofs. Finally, we implement the proposed learning process on a toy example, the classic Rock-Paper-Scissors game, and empirically observe the convergence of empirical play to its unique Nash equilibrium.

II. Deterministic Calibration

Consider a finite outcome space $\Omega = \{1, 2, \ldots, |\Omega|\}$. Let X be the sequence of outcomes whose tth element is X_t, which indicates the outcome at time t. X_t is represented as a one-hot vector in $\{0,1\}^{|\Omega|}$, i.e., if the state i was realized at time t then $X_t(i) = 1$ and for every other index j, $X_t(j) = 0$ ($j \neq i$). The empirical frequency of outcomes up to time T is then given by

$$\bar{X}_T = \frac{1}{T}\sum_{t=1}^{T} X_t.$$

A forecasting method, F, is simply a function that maps past observations to a probability distribution over outcomes, i.e., the forecast at time t is

$$f_t = F(X_1, X_2, \ldots, X_{t-1}) \in \Delta(\Omega)$$

where $\Delta(\Omega)$ denotes the probability simplex over Ω. Each f_t is interpreted as the algorithm's prediction of the next outcome X_t.

A. Calibration

Intuitively, a forecasting algorithm is calibrated if its predictions match the observed outcomes in the long run, specifically on those rounds where it made similar predictions.

To formalize this, fix a forecast $p \in \Delta(\Omega)$ and a small threshold $\epsilon > 0$. Define the function $I_{p,\epsilon}$ as follows

$$I_{p,\epsilon}(f_t) = \begin{cases} 1 & \text{if } |f_t - p| \leq \epsilon \\ 0 & \text{otherwise} \end{cases}$$

where $|p - q|$ is the $L1$ norm between p and q. Then, the calibration error for F with respect p and ϵ over T rounds is defined as

$$\mu_T(I_{p,\epsilon}, X, F) := \frac{1}{T}\sum_{t=1}^{T} I_{p,\epsilon}(f_t)(X_t - f_t).$$

A forecast is said to be calibrated if this error vector converges to zero for all $p \in \Delta(\Omega)$ and all $\epsilon > 0$, i.e., $\mu_T(I_{p,\epsilon}, X, F) \to 0$ as $T \to \infty$. As discussed in Section I, it has been shown that no deterministic algorithm can be calibrated in this sense under adversarial conditions [6], [5].

B. Weak calibration

To overcome this impossibility, the paper introduces a relaxed notion called weak calibration, based on smoothing the calibration error using Lipschitz test functions.

Let $w : \Delta(\Omega) \to [0,1]$ be any Lipschitz continuous function[1]. The calibration error with respect to w is

$$\mu_T(w, X, F) = \frac{1}{T}\sum_{t=1}^{T} w(f_t)(X_t - f_t).$$

A forecasting algorithm F is weakly calibrated if for all sequences X and all Lipschitz functions w, the error $\mu_T(w, X, F) \to 0$ as $T \to \infty$.

To strengthen this, define W_λ to be the class of all Lipschitz functions with Lipschitz constant at most λ. Then, uniform weak calibration requires

$$\sup_{w \in W_\lambda} |\mu_T(w, X, F)| \to 0 \quad \text{as} \quad T \to \infty.$$

One of the primary results in the paper is as follows:

Theorem II.1. *There exists a deterministic forecasting algorithm F that is uniformly weakly calibrated.*

The construction of such an algorithm is provided in Section IV.

C. Randomized rounding and standard calibration

Although deterministic algorithms cannot be calibrated in the strict sense, the paper shows how randomized rounding of a weakly calibrated forecast can recover an appropriate standard calibration. The idea is to start with a weakly calibrated deterministic algorithm and then introduce controlled randomness by rounding the predictions to a nearby grid point in a way that preserves the expected values. This allows us to achieve approximate calibration in the traditional sense.

Example II.1. *Consider a scenario where a weather forecaster publicly announces a probability $f_t \in [0,1]$ of rain for each day t. Over five days, the forecasts might be* $f_1 = 0.8606, f_2 = 0.2387, f_3 = 0.5751, f_4 = 0.4005, f_5 = 0.0696$. *Suppose an observer rounds each forecast to two decimal places, i.e., to the nearest point in the grid $V = \{0.00, 0.01, \ldots, 1.00\}$. For example, for $f_1 = 0.8606$, we round it probabilistically to* 0.87 *with probability* 0.06 *and* 0.86 *with probability* 0.94*. This procedure ensures that the expected rounded value equals the original forecast. Under this scheme, the asymptotic calibration error of the observer will be small (even smaller if the observer rounds it to the third decimal digit), as seen in Corollary II.2.*

Formally, let $\Delta(\Omega)$ be the simplex over outcomes and let $V \subset \Delta(\Omega)$ be a finite triangulation (grid) of the simplex, such that every forecast $f \in \Delta(\Omega)$ lies in a simplex with

[1]Lipschitz functions are those functions that satisfy the following condition: Given $x, y \in dom(f)$ there exists a constant λ such that $|f(x) - f(y)| \leq \lambda|x - y|$.

vertices $v_1, v_2, \dots, v_k \in V$. We define weights $w_v(f)$ for each vertex $v \in V$ such that

$$f = \sum_{v \in V(f)} w_v(f) \cdot v, \quad \text{where} \quad \sum w_v(f) = 1, \; w_v(f) \geq 0.$$

The randomized rounding operator $Round_V(f)$ selects $v \in V(f)$ with probability $w_v(f)$. We round f to one of its grid points stochastically, while preserving the mean.

Corollary II.2. *If the diameter of each simplex in the triangulation is at most δ, i.e., for any q, q' in the same simplex then $|q - q'| \leq \delta$, and F is a weakly calibrated forecaster, then the randomized forecaster $\hat{F} = Round_V \circ F$ satisfies*

$$\lim_{T \to \infty} \sup |\mu_T(I_{p,\epsilon}, X, F)| \leq \delta.$$

By rounding a weakly calibrated forecast, one can achieve approximate calibration to within the triangulation error δ.

Proof: Let f_t be the deterministically weakly calibrated forecast at round t, and $\hat{f}_t = Round_V(f_t)$ the randomized rounded forecast. Since $\hat{f}_t$ is a random variable taking values in V, we can analyze calibration errors on each subset $J \subseteq V$ instead of ϵ-balls around p (which is the case with $I_{p,\epsilon}$). Let $I_J(\hat{f}_t)$ be the indicator function for whether $\hat{f}_t \in J$. The expected calibration error on this subset is

$$\mathbb{E}\Big[\frac{1}{T} \sum_{t=1}^{T} I_J(\hat{f}_t)(X_t - \hat{f}_t)\Big].$$

By design of the rounding $\mathbb{E}[\hat{f}_t | f_t] = f_t$ and f_t is δ-close to any $v \in V(f_t)$, so $|\hat{f}_t - f_t| \leq \delta$.

For any subset, the error made by the rounded forecast is close to the error made by the original forecast, because rounding only moves us by at most δ

$$|X_t - \hat{f}_t| \leq |X_t - f_t| + |f_t - \hat{f}_t| \leq |X_t - f_t| + \delta.$$

The expected error for each vertex $v \in V$ is computed by the weak calibration property of the base forecast F. Since F is weakly calibrated, this expectation error goes to zero over time. Therefore, the empirical error of $\hat{F}$ over any such subset converges to its expectation, which in turn is within δ of zero due to the rounding error. For any region J the rounded forecast yields empirical frequencies that are within δ of the predictions, ensuring approximate calibration.

Note that this randomized scheme is "almost deterministic" in the sense that the forecast made by $Round_V \circ F$ is δ-close to a deterministic forecast.

III. Publicly Calibrated Learning

In this section, we examine the learning setting considered in the paper, where players repeatedly interact in a game and make decisions based on publicly observed forecasts. Also, we will analyze how weak calibration can guide players toward rational behaviour and ultimately to Nash equilibria.

To begin, consider a finite n-player game, where each player $i \in \{1, 2, \dots, n\}$ has a finite set of actions $\mathcal{A}_i$ and the joint action space is denoted by $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \dots \mathcal{A}_n$. The payoff function of player i is given by $u_i : \mathcal{A} \to \mathbb{R}$. A Nash equilibrium is a distribution $p \in \Delta(\mathcal{A})$ such that every player is playing a best response to the marginal distribution over the opponents' strategies/ Formally, p is a Nash equilibrium if for all i and for all actions $a'_i \in \mathcal{A}_i$, we have

$$\mathbb{E}_{a \sim p}[u_i(a)] \geq \mathbb{E}_{a \sim p}[u_i(a'_i, a_{-i})],$$

where $a = (a_i, a_{-i})$. An ϵ-Nash equilibrium (NE_ϵ), allows for small deviations from this optimality condition. A distribution $p \in \Delta(\mathcal{A})$ is an ϵ-Nash equilibrium if, for all i and all $a'_i \in \mathcal{A}_i$

$$\mathbb{E}_{a \sim p}[u_i(a)] \geq \mathbb{E}_{a \sim p}[u_i(a'_i, a_{-i})] - \epsilon.$$

This relaxed definition is useful in learning concepts where players' actions may not exactly converge, but approximate equilibrium behaviour can still emerge over time.

In the learning setting considered in the paper, play proceeds in discrete rounds. On each round t, a public forecast $f_t \in \Delta(\mathcal{A})$, a probability distribution over joint actions, is generated and observed by all players. A deterministic weakly calibrated forecaster computes this forecast, and it is the only signal available to players. Each player i selects action $a^i_t \in \mathcal{A}_i$ in response to f_t, using a policy $\pi_i : \Delta(\mathcal{A} \to \mathcal{A}_i)$. Thus, we have $a^i_t = \pi_i(f_t)$, meaning the player's actions depend solely on the forecast. The assumption that players act based only on the public forecast may initially seem restrictive, but alternatively, one could assume that each player has the knowledge that the forecaster possesses to make their prediction. And since a forecaster only needs history of the play as an input, all of the players share the common knowledge of history, essentially.

Given this setup, the distribution over joint actions at time t becomes a product distribution $\pi(f_t) = \pi_1(f_t) \dots \pi_n(f_t) \in \Delta(\mathcal{A})$. A natural goal is for players to behave rationally by responding optimally to the forecast. Ideally, this would mean using best response policies, where for each player i, the strategy $\pi_i(f) \in \Delta(\mathcal{A}_i)$ satisfies

$$\mathbb{E}_{a \sim \pi(f)}[u_i(a)] \geq \mathbb{E}_{a \sim (\pi'_i, \pi_{-i}(f))}[u_i(a)]$$

for all $\pi'_i \in \Delta(\mathcal{A}_i)$. However, exact best response mappings are generally discontinuous in the forecast f, which conflicts with the continuity required for convergence guarantees under weak calibration.

To address this, the paper suggests that players use continuous ϵ-best response policies, for a small $\epsilon > 0$, where each $\pi_i : \Delta(\mathcal{A}) \to \Delta(\mathcal{A}_i)$ is continuous and satisfies

$$\mathbb{E}_{a \sim \pi(f)}[u_i(a)] \geq \mathbb{E}_{a \sim (\pi'_i, \pi_{-i}(f))}[u_i(a)] - \epsilon$$

for all $\pi_i \in \Delta(\mathcal{A}_i)$.

Such policies can be constructed using the randomized rounding method based on a triangulation of $\Delta(\mathcal{A})$. Let $V \subset \Delta(\mathcal{A})$ be the set of vertices of a triangulation and let $Round_V(f)$ denote a randomized rounding function that maps any forecast $f \in \Delta(\mathcal{A})$ to a nearby vertex $v \in V$, such that $\mathbb{E}[Round_V(f)] = f$. Each player then draws a vertex v according to this rounding, computes the marginal v_{-i} and plays a pure best response to v_{-i}. Since $|f - v| \leq \delta$

(as mentioned in Corollary II.2), the resulting action is guaranteed to be a 2δ-best response to f. Rounding to a nearby vertex and playing a best response to its marginal preserves approximate optimality. This construction yields continuous ϵ-best response functions, which are compatible with weak calibration and form the foundation for the following convergence results.

A. Nash convergence

Define the distance between a distribution p and a set q as $d(p, Q) = \inf_{q \in Q} |p - q|$, similarly between two distributions.

Theorem III.1. *Assume the forecasting rule F is weakly calibrated and that all players act according to the ϵ-best response functions $\pi_i(\cdot)$. The following three conditions hold:*

(i) *(Believing Nash)*

$$\frac{1}{T}\sum_{t=1}^{T} d(f_t, NE_\epsilon) \to 0 \quad as \quad T \to \infty.$$

(ii) *(Playing Nash)*

$$\frac{1}{T}\sum_{t=1}^{T} d(\pi(f_t), NE_\epsilon) \to 0 \quad as \quad T \to \infty.$$

(iii) *(Merging)*

$$\frac{1}{T}\sum_{t=1}^{T} d(f_t, \pi(f_t)) \to 0 \quad as \quad T \to \infty.$$

This result shows that the empirical forecasts and play distributions converge to the ϵ-Nash set. Moreover, at most time steps, play is close to some ϵ-Nash equilibrium, solving the coordination problem in a natural way, players can switch between equilibria but only among the ϵ-Nash set.

Corollary III.2. *Under the same assumption as Theorem III.1, the empirical frequency of play converges to the convex hull of ϵ-Nash equilibria*

$$\frac{1}{T}\sum_{t=1}^{T} \to Convex(NE_\epsilon)$$

with probability one.

This ensures that each player's average utility is at least that of some ϵ-Nash equilibrium, thus providing a stronger guarantee than convergence to the correlated equilibria set.

B. The proof

We provide the proof of Theorem III.1 using the following lemmas. Note, the proofs for the following lemmas and the theorem are presented with intuitive arguments, as opposed to using complete mathematical rigour; refer to [1] for exact proofs.

Lemma III.3. *For all Lipschitz continuous test functions w,*

$$\frac{1}{T}\sum_{t=1}^{T} w(f_t)(f_t - \pi(f_t)) \to 0 \quad as \quad T \to \infty.$$

Proof: Since $\mathbb{E}[X_t | f_t] = \pi(f_t)$, the calibration condition ensures that both

$$\frac{1}{T}\sum_{t=1}^{T} w(f_t)(f_t - X_t) \to 0 \text{ and } \frac{1}{T}\sum_{t=1}^{T} w(f_t)(X_t - \pi(f_t)) \to 0.$$

Combining both gives the result.

Lemma III.4. *If $f = \pi(f)$, then $f \in NE_\epsilon$.*

Proof: Since $\pi(f)$ is a product distribution and $\pi_i(f)$ is an ϵ-best response to f_{-i}, the condition for ϵ-Nash is satisfied.

These lemmas imply that forecasts satisfying $f = \pi(f)$ must be ϵ-Nash equilibria. The remainder of the proof shows that such forecasts dominate the empirical distribution.

Lemma III.5. *If $f \neq \pi(f)$, then f is asymptotically unused, i.e., there exists a continuous test function w with $w(f) = 1$, such that*

$$\frac{1}{T}\sum_{t=1}^{T} w(f_t) \to 0 \quad as \quad T \to \infty.$$

Proof: Use continuity of $\pi(\cdot)$ to construct a neighbourhood where forecasts diverge from play distribution. Lemma III.3, then implies that this region contributes vanishing weights.

Lemma III.6. *Let Q be a compact set of forecasts such that $\forall\ f \in Q, f \neq \pi(f)$. Then Q is asymptotically unused.*

Proof: Use compactness to extract a finite subcover and sum the associated test functions. Then apply Lemma III.5.

Now completing the proof of Theorem III.1. For part (iii), fix $\delta > 0$. Define

$$Q_\delta = \{q \in \Delta(\mathcal{A}) : d(q, \pi(f)) \geq \delta\}.$$

Since $f \neq \pi(f)$, Lemma III.6 implies Q_δ is asymptotically unused

$$\frac{1}{T}\sum_{t=1}^{T} I(f_t \in Q_\delta) \to 0.$$

This implies

$$\frac{1}{T}\sum_{t=1}^{T} d(f_t, \pi(f_t)) \to 0.$$

For parts (i) and (ii), whenever $f_t \neq \pi(f_t)$, from above we know that we can define a set Q_δ such that it is asymptotically unused. And $f \notin NE_\epsilon$ when $f \neq \pi(f)$ from Lemma III.4 and III.5, combining these two we can argue the proofs for parts (i) and (ii).

For Corollary III.2, since

$$\frac{1}{T}\sum_{t=1}^{T} X_t \to \frac{1}{T}\sum_{t=1}^{T} \pi(f_t) \quad \text{as} \quad T \to \infty$$

(by calibration with a constant test function), and each $\pi(f_t) \in NE_\epsilon$, the average must lie in the convex hull

$$\frac{1}{T}\sum_{t=1}^{T} X_t \to Convex(NE_\epsilon).$$

IV. Deterministically Calibrated Algorithm

To ensure the existence of a deterministic, uniformly weakly calibrated forecasting algorithm, we design and analyze the construction that explicitly avoids the limitations of deterministic forecasting in the classical sense. The idea is to work with a relaxed version of the probability simplex and use geometric triangulations to guide a fixed-point algorithm.

Given a finite outcome space Ω, a standard forecast is a distribution $f \in \Delta \subset \mathbb{R}^{|\Omega|}$, where

$$\Delta = \Big\{ f \in \mathbb{R}^{|\Omega|} : \sum_k f(k) = 1, f(k) \geq 0 \Big\}.$$

However, for technical reasons in the algorithm construction, allow forecasts to lie in a relaxed simplex

$$\tilde{\Delta} = \Big\{ f \in \mathbb{R}^{|\Omega|} : \sum_k f(k) = 1, f(k) \geq -\epsilon \Big\}$$

for some small $\epsilon > 0$.

This expansion becomes important, it simplifies the construction of a fixed-point mapping and ensures convexity and continuity in our update rules. Later, this simplex is projected back to the original simplex Δ, preserving convergence guarantees due to convexity.

We assume a triangulation over $\tilde{\Delta}$, defined as a partition into simplices such that any point $p \in \tilde{\Delta}$ lies in one such simplex. Let $V \subset \tilde{\Delta}$ denote the finite vertex set of this triangulation. For each $p \in \tilde{\Delta}$, let $V(p) \subset V$ be the set of vertices spanning the simplex that contains p, so

$$p = \sum_{v \in V(p)} w_v(p) v \quad \text{with} \sum_{v \in V(p)} w_v(p) = 1, \; w_v(p) \geq 0.$$

These weight functions $w_v(p)$ define the test functions used in calibration measurement

$$\mu_T(v) = \frac{1}{T} \sum_{t=1}^{T} w_v(f_t)(X_t - f_t)$$

where $f_t \in \tilde{\Delta}$.

A. The algorithm: Forecast the fixed point

We now describe the Forecast the fixed point algorithm, which is deterministic and ensures weak calibration by construction. For each vertex $v \in V$, define

$$\rho_T(v) = v + \mu_T(v).$$

This function attempts to "correct" the forecast by shifting v in the direction of its accumulated calibration error.

Extend ρ_T to arbitrary $p \in \tilde{\Delta}$ by interpolation

$$\rho_T(p) = \sum_{v \in V} w_v(p) \rho_T(v) = p + \sum_{v \in V} w_v(p) \mu_T(v).$$

At time $T = 1$, set $\mu_0(v) = 0$ for all $v \in V$. At each round T ($T \geq 2$), the algorithm chooses the forecast $f_T \in \tilde{\Delta}$ satisfying

$$f_T = \rho_{T-1}(f_T),$$

i.e., the forecast is fixed point of the update function ρ_{T-1}.

B. Analysis of forecast the fixed point algorithm

To prove the existence of a fixed point consider the following Lemma IV.1.

Lemma IV.1. *For any sequence of outcomes $X_1, X_2, \ldots, X_{T-1}$, a fixed point of ρ_{T-1} exists in $\tilde{\Delta}$. Moreover, for the chosen forecast f_T, it holds that*

$$\sum_{v \in V} w_v(f_T) \mu_{T-1}(v) = 0.$$

Proof: The result follows from Brouwer's fixed point theorem. Since the extended simplex $\tilde{\Delta}$ is convex and compact, and the function ρ_{T-1} is continuous (as it is constructed using linear operations and continuous coordinates), Browser;s theorem guarantees the existence of a fixed point $f_T \in \tilde{\Delta}$ such that $f_T = \rho_{T-1}(f_T)$. Substituting into the definition of ρ_{T-1}, we get $f_T = f_T + \sum_{v \in V} w_v(f_T) \mu_{T-1}(v)$, which implies the stated equality.

We obtain a bound on the calibration error. Define $|| \cdot ||$ as the $L2$ norm.

Lemma IV.2. *For any time T, the total squared calibration error satisfies*

$$\sum_{v \in V} ||\mu_T(v)||^2 \leq \frac{9}{T}.$$

Proof: We define $r_T(v) = T\mu_T(v) = \sum_{t=1}^{T} w_v(f_t)(X_t - f_t)$, which accumulates the weighted error for search vertex v up to time T. Consider how $||r_T(v)||^2$ evolves overt time. We can write $||r_T(v)||^2 = ||r_{T-1}(v)||^2 + w_v(f_T)^2 ||X_T - f_T||^2 + 2w_v(f_T) r_{T-1}(v)(X_T - f_T)$. When we sum this over all verties $v \in V$, the cross term (last term) vanishes due to Lemma IV.1, which states $\sum_v w_v(f_T)\mu_{T-1}(v) = 0$, and hence $\sum_v w_v(f_T) r_{T-1}(v)(X_T - f_T) = 0$. The second term is bounded above by 9, since $||X_T - f_T|| \leq 3$ (property of being in the same simplex, $\tilde{\Delta}$) and the weights $w_v(f_T)$ are non-negative and sum to 1. By recursion, we then have $\sum_v ||r_T(v)||^2 \leq \sum_v ||r_{T-1}(v)||^2 + 9$, and unrolling this bound yields $\sum_v ||r_T(v)||^2 \leq 9T$. Diving by T^2, we obtain the desired result.

Now consider the calibration error for any Lipschitz test function.

Lemma IV.3. *Let $g : \tilde{\Delta} \to \mathbb{R}$ be a Lipschitz function with constant λ_g. Then the calibration error under g satisfies*

$$|\mu_T(g)| \leq \sqrt{\frac{9|V|}{T}} + 3\lambda_g \epsilon.$$

Proof: Construct an approximation to g using the triangulation/. Define $\hat{g}(p) = \sum_{v \in V} g(v) w_v(p)$, which is a linear interpolant of g over the simplex containing p. By Lipschitz continuity of g, the interpolation error is bounded by $|g(p) - \hat{g}(p)| \leq \lambda_g \epsilon$. The calibration error for g is given by $\mu_T(g) = \frac{1}{T}\sum_{t=1}^{T} g(f_t)(X_t - f_t)$. We decompose this into $\mu_T(\hat{g}) + \mu_T(g - \hat{g})$. The second term is bounded in norm by $\frac{1}{T}\sum_{t=1}^{T} |g(f_t) - \hat{g}(f_t)| \cdot ||X_t - f_t|| \leq 3\lambda_g \epsilon$, using $||X_t - f_t|| \leq 3$ (property of being the same simplex). The first term $\mu_T(\hat{g})$ can be expresses as $\sum_{v \in V} g(v)\mu_T(v)$, and hence its norm is at most $\sum_{v \in V} |g(v) \cdot ||\mu_T(v)||$. By Cauchy-Schwartz inequality and the boundedness of g on a finite set V, we have $|\mu_T(\hat{g})| \leq \sqrt{|V| \sum_{v \in V} ||\mu_T(v)||^2}$, which, by Lemma IV.2, is at most $\sqrt{9|V|/T}$. Adding both components gives the result.

Now we tie everything up to show that the constructed algorithm is asymptotically uniform weak calibration.

Theorem IV.4. *There exists a deterministic forecasting algorithm such that for every Lipschitz continuous test function g, the calibration error $\mu_T(g) \to 0$ as $T \to \infty$, and the convergence is uniform over all such functions with bounded Lipschitz constant.*

Proof: Construct a sequence of triangulations $\{V_i\}$ of $\tilde{\Delta}$, each with decreasing maximum diameter $\epsilon_i \to 0$. For each i, define the class G_i of Lipschitz functions with constant at most $\lambda_i = 1/\sqrt{\epsilon_i}$. The algorithm runs in phases, where in phase i, the fixed point forecasting algorithm is applied using the triangulation V_i and is executed for at least $T_i = |V_i|/\epsilon_i$ rounds. From Lemma IV.3, the calibration error for any function $g \in G_i$ after phase i is at most $6\sqrt{\epsilon_i}$.

To ensure uniform convergence, we can guarantee that for any $g \in G_i$, the total calibration error at all times T after phase i remains below $12\sqrt{\epsilon_i}$. This is achieved by choosing the duration of phase $i+1$ to be long enough that the contributions to the cumulative error from the transition are small, relative to the entire time horizon. Since any Lipschitz function with bounded constant eventually belongs to some G_i, and since the error after that phase remains small, the algorithm achieves the desired uniform convergence over the entire class of test functions.

The proof is essentially complete except that the algorithm is defined on $\tilde{\Delta}$ and not on Δ. Now we consider a projection that projects $\tilde{\Delta}$ onto Δ.

Lemma IV.5. *Let $P : \tilde{\Delta} \to \Delta$ denote the projection onto the standard simplex. Define the projected forecast $f_t = P(\tilde{f}_t)$, where $\tilde{f}_t \in \tilde{\Delta}$ is the output of the fixed point algorithm. Then for any Lipschitz test function w, the calibration error of the composed forecast $P \circ F$ satisfies*

$$\mu_T(w, X, P \circ F) \leq \mu_T(w \circ P, X, F) + |\Omega|\epsilon.$$

Proof: Since each component of $\tilde{f}_t \in \tilde{\Delta}$ deviates from its projection $f_t \in \Delta$ by at most ϵ and since the dimension of the space is $|\Omega|$, the total deviation satisfies $||P(\tilde{f}_t) - \tilde{f}_t|| \leq |\Omega|\epsilon$. Then for any Lipschitz test function w, the difference in calibration error incurred by using the projected forecast rather than the unprotected one is bounded by $|\Omega|\epsilon$, due to the triangle inequality and boundedness of X_t. Furthermore, the function $w \circ P$ is still Lipschitz, and hence the convergence guarantees established for F carry over to the composition $P \circ F$. Therefore, the final algorithm, fixed-point forecasting followed by projection, remains deterministic and retains the property of uniform weak calibration.

V. Toy Example: Rock-Paper-Scissors

To demonstrate the deterministic weak calibration based learning, we implemented their algorithm in the classic two-player Rock-Paper-Scissors game. Although the Rock-Paper-Scissors game is simple in structure, it serves as an illustrative example to analyze convergence due to its unique Nash equilibrium in mixed strategies.

Rock-Paper-Scissors is a symmetric, zero-sum game with three pure strategies: Rock, Paper and Scissors. The payoff matrix for Player 1 is

$$Payoff_1 = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}, \; Payoff_2 = -Payoff_1.$$

We know that the unique mixed-strategy Nash equilibrium is for both players to play each action with equal probability, i.e., $(1/3, 1/3, 1/3)$.

The learning algorithm is such that both players iteratively improve their behaviour using only a shared public forecast about the other players' future actions. Each forecast is generated as a probabilistic prediction and is then rounded to a nearby point on a discretized grid of the probability simplex using a randomized rounding scheme, refer Section IV. This allows for approximate calibration while preserving the expected value of the forecast. Each player then responds with a best response to this rounded forecast. In this setup, players do not have to observe or infer each others' payoffs or strategies, only the shared forecast is needed. After each round, players update their calibration error by comparing their action (expressed as a one-hot vector) to the forecast and accumulating this deviation across all forecast points used. Refer to Figure 1 for empirical distributions of Rock, Paper and Scissors for the first 500 rounds of the simulation.

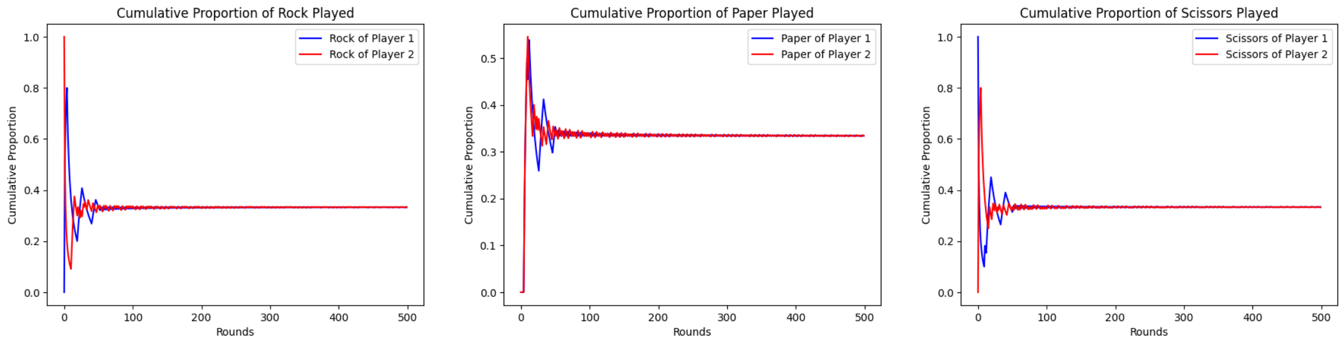
As the simulation proceeds over multiple rounds, each player builds up a history of empirical play. By averaging over the history, we compute the empirical strategy used by each player. When running the simulation for 20000 rounds with a grid resolution of 0.1, the empirical strategies of both players converge very closely to the Nash equilibrium. The output strategies (empirically obtained) are

$$\text{for Player 1:}[0.3333, 0.3333, 0.3334],$$
$$\text{for Player 2:}[0.3334, 0.3333, 0.3333].$$

These results provide empirical validation for the theoretical guarantees offered in the paper. In particular, the convergence to the Nash equilibrium shows that the deterministic weak calibration algorithm, paired with best-response behaviour, yields rational play and leads to a solution of the game. This convergence is not only to an approximate Nash equilibrium at each step but also, over time, to the convex hull of such equilibria, showing the merging behaviour described in the theoretical results.

This toy example effectively encapsulates the fundamental ideas introduced by the authors: that weak calibration can be achieved deterministically and that this is sufficient to guide rational learning behaviour toward Nash equilibrium. The code used for the simulation can be found here[2].

[2]https://github.com/srikarbabug/EE6417-IncentiveCenteredDesign-Project



(a) Empirical distribution of Rock (b) Empirical distribution of Paper (c) Empirical distribution of Scissors

Fig. 1: Empirical distribution of Rock, Paper and Scissors observed in the first 500 rounds of the simulation.

VI. Conclusion

This paper explored the connection between deterministic weak calibration and Nash equilibrium convergence by studying the public learning dynamics in the context of repeated games. We implemented their algorithm in a classic Rock-Paper-Scissors game and empirically observed the convergence of player strategies toward the Nash equilibrium. The results validate the theoretical results that weak calibration, when paired with best response behavior, provides a plausible learning mechanism that drives empirical play toward rational outcomes.

While the approach is conceptually elegant, it comes with computational challenges, as elaborated in the paper. The convergence rates, though provable, are not polynomial time, limiting the scalability of the method for larger games. This is in line with known hardness results for computing Nash equilibria. Furthermore, questions remain unanswered regarding how efficiently the triangulation or rounding schemes can be generalised to complex strategic settings or games with large action spaces. Addressing these open problems could improve the practical viability of deterministic calibration based learning in multi-agent systems.

References

[1] S. M. Kakade and D. P. Foster, “Deterministic calibration and nash equilibrium,” *Journal of Computer and System Sciences*, vol. 74, no. 1, pp. 115–130, 2008.

[2] G. W. Brown, “Iterative solution of games by fictitious play,” *Act. Anal. Prod Allocation*, vol. 13, no. 1, p. 374, 1951.

[3] D. P. Foster and R. Vohra, “Regret in the on-line decision problem,” *Games and Economic Behavior*, vol. 29, no. 1-2, pp. 7–35, 1999.

[4] A. P. Dawid, “The well-calibrated bayesian,” *Journal of the American statistical Association*, vol. 77, no. 379, pp. 605–610, 1982.

[5] ——, “Comment: The impossibility of inductive inference,” *Journal of the American Statistical Association*, vol. 80, no. 390, pp. 340–341, 1985.

[6] D. Oakes, “Self-calibrating priors do not exist,” *Journal of the American Statistical Association*, vol. 80, no. 390, pp. 339–339, 1985.

[7] D. P. Foster and R. V. Vohra, “Asymptotic calibration,” *Biometrika*, vol. 85, no. 2, pp. 379–390, 1998.

[8] ——, “Calibrated learning and correlated equilibrium,” *Games and Economic Behavior*, vol. 21, no. 1-2, pp. 40–55, 1997.