THE EXACT MIXABILITY OF FOLLOW-THE-PERTURBED-LEADER

Anonymous authors
Paper under double-blind review

ABSTRACT

Follow-the-perturbed-leader with Gumbel noise in online learning and perturb-and-MAP in structured prediction run the same device, an independent standard Gumbel for every variable-value pair of a decision set followed by a single call to the optimizer. One Gumbel per action would make the perturbed maximizer an exact sample of exponential weights, whose regret against a mixable loss stays bounded at every horizon, but the actions are exponentially many. Noise on the variables keeps the single call and loses that exactness, at a price in learning rate that neither literature has determined. We determine that price on every product decision set, the configuration space of a graphical model, where it is exactly the fractional certificate complexity of what the loss reads. Dividing the learning rate of exponential weights by that complexity keeps the regret bounded at every horizon, whatever the energies of the model, and no smaller divisor does. A measure from Boolean function complexity thereby acquires an operational meaning, as the cost of a single optimizer call.

1 INTRODUCTION

Perturbing the leader goes back to Hannan (1957), and Kalai & Vempala (2005) made it the generic algorithm for combinatorial decision sets: add random noise to the cumulative losses, then let an optimizer pick the best action. The actions are the feasible set of an optimization problem, exponentially many, and the optimizer is the learner's only access to them. The noise enters the objective through the variables: one independent standard Gumbel for every variable-value pair, so that an action is perturbed by the sum of the Gumbels of its values. Structured prediction runs the same device under another name: perturb-and-MAP turns a MAP solver into a sampler for an energy-based model by perturbing the energies before the call (Papandreou & Yuille, 2011; Hazan & Jaakkola, 2012), with the same noise, one Gumbel per variable and value. With one Gumbel per action instead, the perturbed maximizer would be a sample of exponential weights over the actions and the expected perturbed maximum would be the log-partition function up to a constant, the Gumbel-max trick (Hofbauer & Sandholm, 2002; Abernethy et al., 2016). That noise has an entry for each of the exponentially many actions.

Exponential weights over the actions is exact and has to be implemented structure by structure (Takimoto & Warmuth, 2003), and there is a regime in which it is the optimal algorithm: against a mixable loss, such as the log loss, the aggregating algorithm plays at the step η of the loss's mixability and its regret stays bounded at every horizon (Vovk, 1990; 1998; Haussler et al., 1998). At the $\sqrt{T}$ rate of linear losses over T rounds both routes have been analyzed at length (Kalai & Vempala, 2005; Koolen et al., 2010; Audibert et al., 2014). In the mixable regime the noise on the variables has never been priced, because mixability is a property of the output distribution and that distribution has no closed form. The perturbations that perturb-and-MAP can compute do not give the Gibbs distribution, and its literature has worked around the same gap: bounds on the partition function (Hazan & Jaakkola, 2012; Balog et al., 2017), exact samplers from repeated MAP calls (Hazan et al., 2013) or from a search over a Gumbel process (Maddison et al., 2014), and the perturbed distribution adopted as a model in its own right (Tarlow et al., 2012; Gane et al., 2014; Berthet et al., 2020; Paulus et al., 2020). In 2024 the statistical properties of these objects are still introduced as "under-explored" (Cohen Indelman & Hazan, 2024). What is missing is a price: how much of the guarantee that one Gumbel per action would give survives the shortcut.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

We ask: against an η-mixable loss, by what factor must the step η of exponential weights be divided for the regret of the mechanism to stay bounded, and what determines that factor? The paper calls this factor the temperature of the mechanism. Its definition is the mixability inequality itself (Definition 2). Exponential weights has temperature one, and a temperature κ divides the usable step by κ. On every product decision set the temperature is exactly the fractional certificate complexity of what the loss reads, a linear program from Boolean function complexity. On every decision set it is within a factor of two of that program, and the step that the temperature prescribes is critical: the regret stays bounded at every constant step up to it and can be made to grow linearly at every larger one.

1.1 Our Results

The setting is the one the two literatures share. A decision set V, the feasible set of a combinatorial problem in the sense of Kalai and Vempala, is a set of n-tuples, each variable i taking its values in a finite set D_i; a product decision set, every combination of values allowed, is the configuration space of a graphical model. A readout M assigns every action an output in a finite set Y, and losses are charged on outputs: the readout is the part or function of the action that the loss depends on, from the whole action to a single variable, and a Boolean readout on the cube $\{0,1\}^n$ is a Boolean function. The mechanism draws one independent standard Gumbel for every variable-value pair and returns the action whose score plus perturbation is largest, the score of an action being the sum of a base score, which can be any function of the action such as the negative energy of a configuration with its interaction terms, and a score for its output.

The learner controls the output scores. It keeps, for every output, the negative of that output's cumulative loss times a constant step α, its learning rate, and in each round it predicts from the output distribution of the mechanism at these scores, every sample of which costs one optimizer call. It is charged the mix loss of that distribution, which the aggregating algorithm controls and which bounds the loss of the prediction a substitution function makes from the distribution, with equality for the log loss (Vovk, 1998). Its regret is against the best output, and success is a regret bounded at every horizon.

The temperature $\kappa^*(M)$ is the infimum of the κ at which the potential of the mechanism, the expected perturbed maximum as a function of the output scores, obeys the mixability inequality that the log-sum-exp potential of exponential weights obeys at $\kappa = 1$ (Definition 2). The inequality is required at every base score, so the temperature depends on the decision set and the readout alone. The temperature is exactly the factor asked for. At every constant step up to $\eta/\kappa^*(M)$ the regret is at most Γ/α, where the potential range Γ is a difference of two expected maxima that depends on the base scores alone, and at every larger constant step the regret can be made to grow linearly (Proposition 1); the largest step at which the regret stays bounded at every base score, the critical step, is therefore $\eta/\kappa^*(M)$. The first half is the telescoping argument of the aggregating algorithm, and the second finds a point of excess curvature, where a two-round cycle of losses returns the output distribution to its starting value while the learner is charged more than every output.

On a product decision set the temperature is the fractional certificate complexity $\mathrm{FC}(M)$ of the readout (Theorem 1), a covering program from Boolean function complexity (Tal, 2013; Gilmer et al., 2013; Kulkarni & Tal, 2016): at an action v, weights on the variables must block every action of a different output with total weight at least one on the variables where the two differ, and $\mathrm{FC}(M)$ is the least total weight the costliest action needs. On every decision set the same program measures how fast the probability of an output can grow when its score is raised: by a factor of at most $e^{\tau\,\mathrm{FC}(M)}$ under a raise of τ, and no smaller exponent works. When the loss reads the values of k variables of a product, $\mathrm{FC}(M) = k$, so the mechanism runs at step η/k whatever the energies of the graphical model, with regret at most $k \log |Y|/\eta$ at base score 0; the temperature is n when the loss reads the whole action and 1 when it reads one variable. The bound Γ/α is attained up to a constant, and the step $\eta/\,\mathrm{FC}(M)$ is critical at base score 0, on the cube $\{0,1\}^n$ with the readout that reports the parity of the action and whether it is the zero string (Proposition 2).

Away from products the identity $\kappa^*(M) = \mathrm{FC}(M)$ becomes a bracket: on every decision set $\mathrm{FC}(M) \le \kappa^*(M) \le 2\,\mathrm{FC}(M)$, the factor two from a Hessian that is at most twice its diagonal, and $\kappa^*(M) \le k$ when the loss reads k variables (Theorem 2), so the critical step is between $\eta/(2\,\mathrm{FC}(M))$ and $\eta/\,\mathrm{FC}(M)$. The Gumbel-Matching distribution of Mena et al. (2018), an inde-

pendent Gumbel on every entry of a score matrix followed by one call to the Hungarian algorithm, is the mechanism on the permutations of $[n]$. Its temperature is exactly $n/2$, and $\min\{k, n/2\}$ when the loss reads k rows (Corollary 1), the lower bound from the exchanges of two rows and the upper bound from the certificates with weight $\frac{1}{2}$ on every row or weight 1 on the rows read.

Conditioning on part of the noise gives two more results (Theorem 3). A readout that depends only on some of the variables has at most the temperature of its projection onto them, and whenever every action has a variable-value pair of its own, as in a one-hot encoding, conditioning on the other noise leaves a softmax and the temperature is 1. On the one-hot encoding of n experts the mechanism is follow-the-perturbed-leader with independent logistic noise, one difference of two Gumbels per expert. Its output distribution is not a softmax, and it runs at the step of exponential weights with regret at most $(1 + \frac{1}{2} + \cdots + \frac{1}{n-1})/\eta$ at base score 0.

The lower bound of the identity rests on the growth rate. Shifting a winner's own Gumbels by τ times a fractional certificate carries the event that it wins after its output is raised by τ into the event that it wins before, at a cost of $e^{\tau\,\mathrm{FC}(M)}$ in probability, and the dual of the certificate program shows that no smaller exponent works. A second-order expansion of the mixability inequality along the raise shows that every κ at which the inequality holds is such an exponent, hence at least $\mathrm{FC}(M)$. The upper bound is a curvature bound. Conditioning on all noise but that of one variable writes the Hessian of the potential through conditional softmax probabilities, and a family of certificates, one per action, that overlap for every two actions with different outputs bounds this Hessian by its weight times the covariance of the output distribution, which bounds the temperature by that weight. On a product the optimal certificates overlap by themselves, since two actions with different outputs whose certificates did not would combine into a hybrid action with both outputs. Beyond products, a conditioning of the noise that ignores the scores replaces the potential by a mixture of simpler potentials, and the mixability inequality survives the averaging by Jensen's inequality. Section 2 fixes the mechanism and its temperature and shows that the temperature sets the critical step, Section 3 computes the temperature on products, Section 4 bounds it on every decision set, and Appendices A to D hold the proofs.

1.2 Related Work

The noise law of the perturbed leader has been a design parameter: bounded hazard rates (Abernethy et al., 2015), log-concave perturbations and the lower bounds they obey (Cohen & Hazan, 2015), random-walk perturbations that rarely switch (Devroye et al., 2015), dropout perturbations (van Erven et al., 2014), Fréchet-type tails, which are optimal for adversarial bandits (Honda et al., 2023; Lee et al., 2024), and curvature-adaptive perturbations (Charikar et al., 2026); the smoothing view (Abernethy et al., 2014; 2016) organizes the area and records that Gumbel perturbations reproduce exponential weights. The present question is about the constant-regret regime of mixable losses.

Exponential weights and componentwise multiplicative updates over structured classes are the other route (Takimoto & Warmuth, 2003; Helmbold & Warmuth, 2009; Koolen et al., 2010), and the minimax rates of online combinatorial optimization are known (Audibert et al., 2014). The oracle-efficiency program asks what can be learned with an optimization oracle alone, counting its calls (Kakade et al., 2009; Rakhlin et al., 2012; Hazan & Koren, 2016; Syrgkanis et al., 2016; Dudík et al., 2020; Suggala & Netrapalli, 2020). Closest to this paper is the change of variables of Wang et al. (2022), a translation of the noise that compares the probabilities with which the perturbed leader picks an action before and after one round's loss is added, and the proof of Theorem 1(a) chooses that translation by a linear program, whose value is the exact constant.

Perturb-and-MAP (Papandreou & Yuille, 2011; Hazan & Jaakkola, 2012) and the Gumbel-max trick supply the model. The computable perturbations are low-rank, one Gumbel per variable and value as in this paper, and beyond the bounds and samplers recalled above the literature has concentration (Orabona et al., 2014), learning procedures and estimators built on perturbed optimizers (Tarlow et al., 2012; Gane et al., 2014; Berthet et al., 2020; Paulus et al., 2020; Niepert et al., 2021; Cohen Indelman & Hazan, 2024), relaxations for gradients (Jang et al., 2017; Maddison et al., 2017), a review (Huijben et al., 2023) and a volume (Hazan et al., 2016). The Gumbel-Matching distribution of Mena et al. (2018) is the mechanism of Corollary 1. The question of what the low-rank perturbation costs is posed by Tomczak (2016); the identity answers it for the output distribution, in terms of the regret guarantee that the distribution gives.

Mixability is due to Vovk (1990; 1998); Haussler et al. (1998) proved the constant of the aggregating algorithm optimal, and the book of Cesa-Bianchi & Lugosi (2006) is the standard account. Exact constants have also been determined for particular games (Vovk & Zhdanov, 2009). Two characterizations of mixability are the closest to ours: van Erven et al. (2012) identify mixability as a curvature of the Bayes risk relative to the log loss, the form in which this paper proves its upper bounds by certificates, and Reid et al. (2015) and Mhammedi & Williamson (2018) settle which entropies generalize it. The temperature compares the potential of the mechanism with the log-sum-exp in the mix loss, and the identity evaluates that comparison exactly, on every product decision set.

The random utility tradition supplies the object and the noise. A random utility model gives every alternative a score plus a random term and picks the maximizer; the expected maximum is its surplus function, and the Williams–Daly–Zachary theorem says that the choice probabilities are the gradient of the surplus and that its cross partial derivatives are nonpositive (Williams, 1977; Daly & Zachary, 1978; McFadden, 1981). Luce's choice axiom yields the softmax and McFadden's conditional logit obtains it from Gumbel utilities (Luce, 1959; McFadden, 1974), Yellott's theorem makes the Gumbel canonical (Yellott, 1977), and the convex analysis of the surplus function is an active subject (Fosgerau et al., 2013; Sørensen & Fosgerau, 2022). The mechanism is such a model: the random term of an action is the sum of the Gumbels of its values, so two actions sharing a value share part of their noise, the error-components structure that applied choice modeling uses to build correlated substitution patterns (Brownstone & Train, 1998; McFadden & Train, 2000), and the temperature is the curvature of its surplus relative to that of the conditional logit at the same choice probabilities.

Fractional certificate complexity and fractional block sensitivity, its linear-programming dual, were introduced by Tal (2013) and by Gilmer et al. (2013) and developed by Kulkarni & Tal (2016); they equal Aaronson's randomized certificate complexity up to a constant (Aaronson, 2008), and the surrounding measures go back to Nisan (1991) and are surveyed by Buhrman & de Wolf (2002). On total Boolean functions the classical minimax adversary bound equals fractional block sensitivity, and on partial functions the two separate (Ambainis et al., 2021); fractional certificates are the subject of recent work as well (Chakraborty et al., 2023; Lovett & Zhang, 2023). In this paper the fractional certificate complexity of a readout is the temperature of the mechanism on every product decision set.

2 Setting

This section fixes the mechanism and its temperature, shows that the temperature sets the critical step, and defines the certificate program that the temperature is compared with.

Throughout, a *decision set* is a nonempty finite set V of n-tuples $v = (v_1, \dots, v_n)$, and its elements are *actions*. We call $i \in [n]$ a *variable*, and $D_i := \{v_i : v \in V\}$ is the finite set of *values* it takes. This is the object of combinatorial online learning: V is the feasible set of an optimization problem, presented through a maximization oracle (Kalai & Vempala, 2005). We call V a *product decision set* when $V = D_1 \times \cdots \times D_n$, so that every combination of values is an action; this is the configuration space of n discrete variables, on which perturb-and-MAP runs (Papandreou & Yuille, 2011; Hazan & Jaakkola, 2012). A *readout* is a map M from V onto a finite set Y, constant when Y has one element. The elements of Y are *outputs*, and losses will be charged on them.

Definition 1 (the mechanism)**.** Let $\xi_i(a)$, for $i \in [n]$ and $a \in D_i$, be independent standard Gumbel random variables, and write $\Xi(v) := \sum_i \xi_i(v_i)$ for the perturbation of an action. The mechanism takes *base scores* $\theta \in \mathbb{R}^V$ and *output scores* $s \in \mathbb{R}^Y$. It draws the noise once and returns the *winner*

$$v^* \;:=\; \operatorname*{arg\,max}_{v \in V} \big\{\theta_v + s_{M(v)} + \Xi(v)\big\}.$$

Its *output distribution* is $p(\theta, s) \in \mathbb{R}^Y$, $p_y(\theta, s) := \Pr[M(v^*) = y]$.

The maximization is the single call to the optimizer, and the noise reaches the $|V|$ actions through the $\sum_i |D_i|$ Gumbels $\xi_i(a)$. The base scores are any function of the action. On a product decision set they can be the negative energies of the configurations of a Markov random field, interaction terms included, and the mechanism is then perturb-and-MAP with one Gumbel per variable and value. The output scores are what the learner will control. The perturbations of two distinct actions differ by a sum of independent standard logistic random variables, one for each variable on which

the actions differ (Appendix A), so the winner is almost surely unique. Adding a constant to all base scores, to all output scores, or to all the Gumbels of one variable shifts every total score equally and leaves the output distribution unchanged. The Gumbel is the law under which the mechanism on a single variable is exactly a softmax, and by Yellott's theorem it is the only such law, up to scale and shift, among choice models with at least three alternatives and independent, identically distributed additive noise (Yellott, 1977).

The *potential* at base score θ is

$$A_\theta(s) \;:=\; \mathbb{E}\max_{v\in V}\{\theta_v + s_{M(v)} + \Xi(v)\}, \qquad s \in \mathbb{R}^Y,$$

a convex function whose gradient is the output distribution, $\nabla A_\theta = p(\theta, \cdot)$ (Appendix A).

Definition 2 (temperature)**.** The *temperature* $\kappa^*(M)$ is the infimum of the *admissible* $\kappa > 0$, those for which

$$A_\theta(s+g) - A_\theta(s) \;\le\; \frac{1}{\kappa}\log\sum_{y\in Y} p_y(\theta, s)\, e^{\kappa g_y} \tag{1}$$

holds for every $\theta \in \mathbb{R}^V$ and all $s, g \in \mathbb{R}^Y$.

Quantifying over every base score makes the temperature a property of the pair (V, M) alone. Equality holds throughout (1) for the log-sum-exp potential $\kappa^{-1}\log\sum_y e^{\kappa s_y}$, whose gradient is the softmax of κs, so Definition 2 asks for the infimum of the κ at which the potential of the mechanism grows no faster along increments than that benchmark. The right side of (1) is the logarithm of a weighted power mean of e^g, so it increases with κ, and the admissible constants form a half-line, closed at $\kappa^*(M)$ when M is nonconstant (Appendix C); a constant readout has $\kappa^*(M) = 0$.

In the mixability literature the condition (1) usually appears as a comparison between second derivatives (van Erven et al., 2012), and it takes that form here. Its two sides agree to first order in g, since $\nabla A_\theta = p(\theta, \cdot)$, and κ is admissible exactly when the second-order terms satisfy $\nabla^2 A_\theta \preceq \kappa\,\mathrm{Cov}(p)$ wherever the Hessian exists, the *curvature criterion* of Appendix A. Here $\mathrm{Cov}(p) := \mathrm{diag}(p) - pp^\top$ is the covariance of the output distribution, and $\kappa\,\mathrm{Cov}(p)$ is the Hessian of the log-sum-exp potential at the point where its gradient is p. For a nonconstant readout the temperature is therefore the supremum of the ratio of the curvature $g^\top\nabla^2 A_\theta g$ of the potential along a direction g to the variance of g under the output distribution.

With A_θ in place of the log-sum-exp potential, (1) is the mixability requirement of the aggregating algorithm (Vovk, 1990; 1998): a loss mixable at rate η is aggregated with step $\eta/\kappa^*(M)$, so $\kappa^*(M)$ is the factor by which the usable step is divided. The game in which that division becomes a regret bound is the following. Fix a base score θ and a rate $\eta > 0$. The learner starts at output scores $s_1 = 0$; in round t the adversary reveals a loss vector $\ell_t \in [0, \infty)^Y$, the learner is charged the mix loss of its current output distribution,

$$m_t \;:=\; -\frac{1}{\eta}\log\sum_{y\in Y} p_y(\theta, s_t)\, e^{-\eta\ell_t(y)},$$

and updates $s_{t+1} = s_t - \alpha\ell_t$ with a constant step $\alpha > 0$. The regret after T rounds is $R_T := \sum_{t\le T} m_t - \min_y \sum_{t\le T}\ell_t(y)$. The mix loss is the quantity that the aggregating algorithm controls: for an η-mixable loss, a substitution function turns the output distribution into a prediction whose loss is at most m_t (Vovk, 1990; 1998; Cesa-Bianchi & Lugosi, 2006). The scale of the bound is the *potential range*

$$\Gamma \;:=\; A_\theta(0) - \min_{y\in Y}\; \mathbb{E}\max_{v:\, M(v)=y}\{\theta_v + \Xi(v)\},$$

a difference of two expected maxima, in which the centering of the noise cancels.

Proposition 1 (the critical step)**.** *Let M be a nonconstant readout on a decision set V and let $\eta > 0$.*

(a) *At every base score, every constant step $\alpha \le \eta/\kappa^*(M)$ has $R_T \le \Gamma/\alpha$ for every loss sequence and every T.*
(b) *For every constant step $\alpha > \eta/\kappa^*(M)$ there are a base score and a loss sequence along which $\liminf_{T\to\infty} R_T/T > 0$.*

The largest constant step at which the regret stays bounded at every base score and every horizon, the *critical step*, is therefore exactly $\eta/\kappa^*(M)$, and at that step the bound reads $\kappa^*(M)\Gamma/\eta$. Part (a) is the telescoping argument of the aggregating algorithm with (1) in place of the log-sum-exp identity: the increment $-\alpha\ell_t$ enters (1) at $\kappa = \eta/\alpha$, which is admissible when $\alpha \le \eta/\kappa^*(M)$, so the potential falls by at least αm_t in every round, while restricting the final potential to the actions of the best output shows that it falls by at most Γ plus α times that output's cumulative loss; the difference of the two bounds is $\alpha R_T \le \Gamma$. For (b), an inadmissible $\kappa = \eta/\alpha$ means that at some base score and output scores the curvature of A_θ along some direction g exceeds κ times the variance of g under the output distribution (Appendix A). The adversary drives the output scores there and then alternates, for a small $d > 0$, the losses $\frac{1}{2}\mathbf{1} + dg$ and $\frac{1}{2}\mathbf{1} - dg$, with $\mathbf{1}$ the all-ones vector: every two rounds return the output distribution to its starting value and charge every output exactly one, while the learner pays one plus a positive multiple of d^2, the multiple being α times that curvature minus η times that variance (Appendix C).

On product decision sets the constant that answers the question of Section 1 is a linear program in the combinatorics of the readout. A *fractional certificate* at an action v is a vector $z \in \mathbb{R}^n_{\ge 0}$ with $\sum_{i \in \Delta(u,v)} z_i \ge 1$ for every $u \in V$ with $M(u) \ne M(v)$, where $\Delta(u,v) := \{i : u_i \ne v_i\}$ is the set of variables on which u and v differ: every action of a different output is blocked with total weight at least one on the variables where it differs from v.

Definition 3 (fractional certificate complexity)**.** The fractional certificate complexity of M at v is the least total weight of a fractional certificate at v,

$$\mathrm{FC}(M, v) \;:=\; \min\Big\{\sum_{i=1}^{n} z_i \,:\, z \text{ a fractional certificate at } v\Big\},$$

and $\mathrm{FC}(M) := \max_{v \in V} \mathrm{FC}(M, v)$.

Its linear-programming dual weights the actions of a different output instead of the variables, and has the same value:

$$\mathrm{FC}(M, v) \;=\; \max\Big\{ \sum_{u:\, M(u) \ne M(v)} \omega_u \;:\; \omega \ge 0, \sum_{u:\, M(u) \ne M(v),\ u_i \ne v_i} \omega_u \le 1 \ \text{ for every } i \Big\}. \quad (2)$$

For a Boolean readout on the cube $\{0,1\}^n$, $\mathrm{FC}(M)$ is the measure of the same name from Boolean function complexity (Tal, 2013; Gilmer et al., 2013; Kulkarni & Tal, 2016). A constant readout has $\mathrm{FC}(M) = 0$, and a nonconstant one has $1 \le \mathrm{FC}(M) \le n$, since every action then has an action of a different output to block, and the all-ones vector blocks all of them.

3 Product Decision Sets

This section computes the temperature on product decision sets. Its lower bound holds on every decision set, and the product structure enters only through the upper bound.

Theorem 1 (the temperature on products)**.** *Let M be a readout on a decision set V, and let e_y be the standard basis vector of $\mathbb{R}^Y$.*

(a) *For all θ, s, $y \in Y$ and $\tau \ge 0$,*

$$p_y(\theta, s + \tau e_y) \;\le\; e^{\mathrm{FC}(M)\tau}\, p_y(\theta, s),$$

and no constant smaller than $\mathrm{FC}(M)$ has this property.

(b) $\mathrm{FC}(M) \le \kappa^*(M)$.

(c) *If V is a product decision set, then $\kappa^*(M) = \mathrm{FC}(M)$.*

The inequality of (a) is a transport. Fix an output y and raise its score by τ; the probability of y rises, and the question is how fast. Since the output distribution has no closed form, the estimate is made in the noise: shift the Gumbels of a winner v of output y by τ times a fractional certificate at v. The shift carries the event that v wins after the raise into the event that v wins before it, because every action of a different output is blocked with weight at least one on the variables where it differs from v, and it costs at most a factor $e^{\tau\,\mathrm{FC}(M,v)}$ in probability, because a shift by τ costs the Gumbel law at most a factor e^{τ}. Summing over the actions of output y gives the inequality of (a) (Appendix B).

The change of variables is the one Wang et al. (2022) use to compare two perturbed programs; with the certificate program as its constant, it holds for every product noise whose density ratio under shifts is bounded, the Gumbel entering only through the value of that ratio. The constant is sharp. If the score of a costliest action's output is lowered by N while the other actions of that output are suppressed, the weights of the dual program (2) combine the conditions for that action to win into one inequality on a sum of independent terms, one per variable, and a Chernoff bound makes the probability of the output decay in N at exponential rate $\mathrm{FC}(M)$; its probability before the lowering is bounded below, so no smaller constant works.

Part (b) follows from (a): a second-order expansion of (1) along the raise of one output shows that every admissible κ bounds every slope of $\operatorname{logit} p_y(\theta, s + \tau e_y)$ in τ, where $\operatorname{logit} p = \log(p/(1-p))$, hence has the property of (a) in place of $\mathrm{FC}(M)$ and is at least $\mathrm{FC}(M)$ (Appendix B).

The upper bound comes from the curvature of the potential, which the overlap of certificates controls. By the curvature criterion of Section 2 it suffices to bound $\nabla^2 A_\theta$ by $\kappa \operatorname{Cov}(p)$. As a function of a separate score for every action, the expected perturbed maximum has a Hessian that couples the actions in pairs, and the curvature of A_θ along g is the sum of these couplings over the pairs of actions with different outputs, each times the squared difference of g between their outputs (Appendix A). Conditioning on all noise except that of variable i turns the mechanism into a softmax over the values of i, one candidate per value, so two actions that differ at i are coupled with strength equal to the expected product of their conditional probabilities, the same at every variable where they differ. With a fractional certificate at every action, the coupling of a pair can therefore be split among those variables according to the smaller of the pair's two certificate weights at each, provided these add up to at least one. At each variable a weighted variance inequality then bounds the split couplings by the variance of g in which the term of every action is multiplied by its own certificate weight.

Definition 4 (compatible certificates)**.** *Compatible certificates* for M are a family $(z^v)_{v \in V}$ of vectors in $\mathbb{R}^n_{\geq 0}$ with

$$\sum_{i \in \Delta(u,v)} \min\{z^u_i, z^v_i\} \;\geq\; 1 \qquad \text{whenever } M(u) \neq M(v).$$

Their *weight* is $W := \max_{v \in V} \sum_i z^v_i$.

Each z^v is in particular a fractional certificate at v, so $W \geq \mathrm{FC}(M)$. Summing the weighted variance bounds over the variables multiplies the term of every action by the total of its certificate weights, at most W, so the curvature of the potential is at most W times the covariance of the output distribution (Lemma B.1) and $\kappa^*(M) \leq W$ on every decision set. On a product, optimal fractional certificates are compatible by themselves. If two actions u and v of different outputs had certificates z^u and z^v overlapping with weight below one, the hybrid

$$w_i \;:=\; \begin{cases} u_i, & z^u_i > z^v_i, \\ v_i, & \text{otherwise}, \end{cases}$$

would be an action, and z^u would put weight below one on the variables where w differs from u, and z^v on those where w differs from v, so w would have both outputs at once. Hence the optimal certificates are compatible, of weight $\mathrm{FC}(M)$, which gives (c). The argument is the fractional form, over arbitrary value sets, of the fact that a 0-certificate and a 1-certificate of a Boolean function on the cube intersect, and it is the one place where the product structure is used.

The scale of the answer is set by what the loss reads. When the loss reads the values of k variables of a product, so that $M(v) = v_S := (v_i)_{i \in S}$ for a set S of k variables each taking at least two values, changing one variable of S changes the output, so every certificate puts weight at least one on each variable of S, and the weights one on S form a certificate: $\mathrm{FC}(M) = k$. The mechanism runs at temperature k, which is n when the loss reads the whole action and 1 when it reads one variable. At base score 0 the variables outside S contribute equally to the two expected maxima of the potential range, so $\Gamma = \log |Y|$ (Appendix C), and Proposition 1 bounds the regret at the step η/k by $k \log |Y|/\eta$ at every horizon, k times the bound of exponential weights over the $|Y|$ outputs.

The bound of Proposition 1(a) is attained, up to a constant, on the following readout, on which the step $\eta/\mathrm{FC}(M)$ is critical at base score 0.

Proposition 2 (the cube readout)**.** *Let* $V = \{0,1\}^n$ *with* $n \geq 3$*, let* $M(v) := (\mathrm{parity}(v), \mathbb{1}\{v = 0\})$ *with* $\mathrm{parity}(v) := \sum_i v_i \bmod 2$*, and let* $\theta = 0$*. Then* $\mathrm{FC}(M) = n$ *and* $\Gamma = n \log 2$*, and*

(i) *for every constant step* α *and every* T *an adversary can force* $R_T \geq \beta_\eta \min\{T, (\Gamma - 1.08\sqrt{n})/\alpha\}$*, where* $\beta_\eta > 0$ *depends on* η *alone;*
(ii) *for every constant step* $\alpha > \eta/n$ *an adversary can force* $\liminf_{T\to\infty} R_T/T > 0$.

Every single-variable neighbor of an action has a different parity, hence a different output, which gives $\mathrm{FC}(M) = n$, and the separate output of the zero action makes Γ the expected gap between the best action and the zero action, $n \log 2$; with three outputs the bound at the critical step is already $n^2 \log 2/\eta$. At every step $\alpha \leq \eta/n$ and every horizon past $(\Gamma - 1.08\sqrt{n})/\alpha$ the two bounds agree up to the factor $\beta_\eta(1 - 1.08\sqrt{n}/\Gamma)$, so on this readout Γ/α is the right order of the regret and $\mathrm{FC}(M)\Gamma/\eta$ is the bound at the critical step. The adversary of (i) charges the output of the zero action nothing and the other two outputs one; the zero action wins only when the sum of the positive parts of n independent logistic gaps, whose mean is Γ, is at most the raise of its output, so for about Γ/α rounds its output has probability at most $\frac{1}{2}$ and the learner pays at least β_η per round. In (ii) the two-round cycle of Proposition 1(b) starts at base score 0 without any driving, with g equal to $+1$ on the two even outputs and -1 on the odd one. The parity of the winner changes when the smallest of the n independent logistic gaps is overcome, which makes the curvature along g equal to n times its variance, and the gain per cycle is $d^2(n\alpha - \eta) + o(d^2)$ (Appendix C).

4 General Decision Sets

Away from products the hybrid is unavailable: a decision set need not contain the hybrid of two of its actions. The certificate program still bounds the temperature within a factor of two, any compatible certificates still bound it, and a conditioning argument computes it exactly when every action has a variable-value pair of its own, and bounds it by the temperature of a projection.

Theorem 2 (general decision sets)**.** *Let* M *be a readout on a decision set* V.

(a) $\kappa^*(M) \leq W$ *for compatible certificates of weight* W.
(b) $\kappa^*(M) \leq 2\,\mathrm{FC}(M)$.
(c) *If* M *depends only on a set* S *of variables, so that* $M(u) = M(v)$ *whenever* $u_S = v_S$*, then* $\kappa^*(M) \leq |S|$.

With Theorem 1(b), $\mathrm{FC}(M) \leq \kappa^*(M) \leq 2\,\mathrm{FC}(M)$ on every decision set, so the critical step of Proposition 1 is between $\eta/(2\,\mathrm{FC}(M))$ and $\eta/\mathrm{FC}(M)$. The bound in (c) is attained on a product by the readout of the variables in S of Section 3, and for $|S| = 1$ it coincides with the lower bound of Theorem 1(b): a nonconstant readout of one variable has temperature exactly 1 on every decision set. Part (a) is the curvature bound of Section 3, which never used the product structure. For (b), differentiating the inequality of Theorem 1(a) at $\tau = 0$ bounds the diagonal of the Hessian of A_θ by $\mathrm{FC}(M)$ times the output distribution. The rows of that Hessian sum to zero, since raising all output scores together leaves the output distribution unchanged, and its off-diagonal entries are nonpositive, since raising the score of one output cannot raise the probability of another. Such a matrix is at most twice its diagonal as a quadratic form, and on g centered at its mean under the output distribution this gives $2\,\mathrm{FC}(M)$ times the variance of g. For (c), two actions with different outputs differ on S, so the vector with ones on S and zeros elsewhere, taken at every action, is a family of compatible certificates of weight $|S|$ (Appendix D).

The Gumbel-Matching distribution of Mena et al. (2018) adds an independent standard Gumbel to every entry of an $n \times n$ score matrix and returns the maximum-weight matching, one call to the Hungarian algorithm. With variable i the row i, and its value the column that row is assigned to, this is the mechanism on the decision set of permutations of $[n]$, which is not a product.

Corollary 1 (assignments)**.** *Let* V *be the set of permutations* σ *of* $[n]$*,* $n \geq 2$*, with variable* i *the row* i *and value* $\sigma(i)$ *its column, and let* $S \subseteq [n]$ *be a set of* $k \geq 1$ *rows. The readout* $M(\sigma) := (\sigma(i))_{i\in S}$ *has*

$$\kappa^*(M) \;=\; \mathrm{FC}(M) \;=\; \min\{k,\, n/2\};$$

in particular the readout $\sigma \mapsto \sigma$ *has temperature* $n/2$.

Exchanging the columns of a row $i \in S$ and of another row j changes the readout and alters no other row, so every certificate at any action has $z_i + z_j \geq 1$ for $i \in S$ and $j \neq i$, and these constraints alone force a total weight of at least $\min\{k, n/2\}$. Conversely, two permutations with different readouts

differ on at least two rows, one of them in S, so the weights $\frac{1}{2}$ on every row, and the weights 1 on the rows of S, are compatible certificates of weights $n/2$ and k (Appendix D).

Beyond these bounds the temperature is controlled by a conditioning of the noise that ignores the scores and turns the instance into a mixture of simpler ones. The defining inequality survives such a mixture (Lemma D.1): its right side is the logarithm of a linear function of the output distribution, hence concave in it, so Jensen's inequality moves the conditional expectation inside. The same argument applies to every softmax model with random base scores: a mixed multinomial logit (McFadden & Train, 2000) is a softmax conditional on its random coefficients, so its potential satisfies (1) at $\kappa = 1$. Two hypotheses supply the conditioning for the mechanism. A decision set is *privately anchored* when every action has a variable-value pair that no other action uses. A readout that depends only on S, as in Theorem 2(c), is $M(v) = \bar{M}(v_S)$ for a readout $\bar{M}$ on the projection $V_S := \{v_S : v \in V\}$.

Theorem 3 (anchoring and projection)**.** *Let M be a nonconstant readout on a decision set V.*

(a) *If V is privately anchored, then $\kappa^*(M) = \mathrm{FC}(M) = 1$.*
(b) *If M depends only on S by way of the readout $\bar{M}$ on V_S, then $\kappa^*(M) \le \kappa^*(\bar{M})$.*

Private anchoring leaves each action a Gumbel of its own: conditioning on every other Gumbel makes the winner the argmax of $|V|$ fixed scores plus independent standard Gumbels, whose distribution is the softmax of those scores and whose potential satisfies (1) at $\kappa = 1$ with equality. Under (b), conditioning on the noise outside S gives every action with the same v_S the same output score, so the maximizing action among them does not depend on the output scores, and what remains is the mechanism on V_S with the readout $\bar{M}$ at random base scores, which the temperature of $\bar{M}$ covers because it quantifies over every base score (Appendix D). Part (b) refines Theorem 2(c), since $\kappa^*(\bar{M}) \le |S|$ by Theorem 2(c) on V_S, and with Theorem 1(c) a readout that depends only on S has $\kappa^*(M) \le \mathrm{FC}(\bar{M})$ whenever V_S is a product.

A one-hot encoding $V = \{e_1, \dots, e_n\} \subset \{0,1\}^n$ is privately anchored, e_j alone taking the value 1 at variable j, so every nonconstant readout of it has temperature 1. Its perturbations are $\Xi(e_j) = \sum_i \xi_i(0) + \delta_j$ with $\delta_j := \xi_j(1) - \xi_j(0)$ independent standard logistic, so with the readout that reports the chosen expert the mechanism is follow-the-perturbed-leader over $n \ge 2$ experts with independent logistic noise. Its output distribution is not a softmax of the scores, and it runs at the step η of exponential weights with regret bounded at every horizon: at base score 0 its potential range is $\Gamma = 1 + \frac{1}{2} + \cdots + \frac{1}{n-1}$ (Appendix D), so its regret at step η is at most that sum divided by η, within $0.58/\eta$ of the bound $\log n/\eta$ of exponential weights.

5 Conclusion

The mechanism that follow-the-perturbed-leader with Gumbel noise and perturb-and-MAP share, one independent Gumbel per variable-value pair and one call to the optimizer, has an exact price in the regime where exponential weights is optimal. Its temperature is the factor by which the step of exponential weights must be divided for the regret to stay bounded, and above the divided step the regret can be made to grow linearly. On a product decision set the temperature is the fractional certificate complexity of the readout, k when the loss reads the values of k variables, so a linear program fixes the step before any loss is seen. On every decision set it is between that complexity and twice it, and at most the number of variables the loss reads. It is $n/2$ for Gumbel-perturbed assignment and 1 whenever every action has a variable-value pair of its own, as for follow-the-perturbed-leader with logistic noise over experts.

Each field reads these results in its own terms. To online learning they give an exact mixability constant for an oracle-based learner, on every product decision set at once. To perturb-and-MAP they give the exact price of the low-rank perturbation on the configuration spaces of graphical models, and they locate where the perturbation keeps the step of exponential weights, on privately anchored decision sets. To Boolean function complexity they give fractional certificate complexity, and with it fractional block sensitivity and the classical minimax adversary bound of Boolean functions on the cube, an operational meaning outside query complexity: it is the factor by which one optimizer call divides the learning rate.

AI USE STATEMENT

The authors used OpenAI's ChatGPT and Anthropic's Claude to help organize and polish the exposition of this paper, based on the authors' own drafts. The same tools assisted in surveying the related literature. The authors formulated and proved all mathematical claims in this paper, and checked, revised, and finalized every statement, proof, and reference. They take full responsibility for the content of the submission.

REPRODUCIBILITY STATEMENT

All results of this paper are mathematical. Section 2 specifies the mechanism, the learner and the certificate program in full, and the appendix contains a complete proof of every result: the regularity of the potentials and the curvature criterion (Appendix A), the temperature on product decision sets (Appendix B), the critical step and the cube readout (Appendix C), and the results on general decision sets (Appendix D).

REFERENCES

Scott Aaronson. Quantum certificate complexity. *Journal of Computer and System Sciences*, 74(3): 313–322, 2008.

Jacob Abernethy, Chansoo Lee, Abhinav Sinha, and Ambuj Tewari. Online linear optimization via smoothing. In *Proceedings of the 27th Conference on Learning Theory (COLT)*, 2014. arXiv:1405.6076.

Jacob Abernethy, Chansoo Lee, and Ambuj Tewari. Fighting bandits with a new kind of smoothness. In *Advances in Neural Information Processing Systems 28 (NIPS)*, 2015.

Jacob Abernethy, Chansoo Lee, and Ambuj Tewari. Perturbation techniques in online learning and optimization. In Tamir Hazan, George Papandreou, and Daniel Tarlow (eds.), *Perturbations, Optimization, and Statistics*. MIT Press, 2016.

Andris Ambainis, Martins Kokainis, Krišjānis Prūsis, Jevgēnijs Vihrovs, and Aleksejs Zajakins. All classical adversary methods are equivalent for total functions. *ACM Transactions on Computation Theory*, 13(1), 2021.

Jean-Yves Audibert, Sébastien Bubeck, and Gábor Lugosi. Regret in online combinatorial optimization. *Mathematics of Operations Research*, 39(1):31–45, 2014.

Matej Balog, Nilesh Tripuraneni, Zoubin Ghahramani, and Adrian Weller. Lost relatives of the Gumbel trick. In *Proceedings of the 34th International Conference on Machine Learning (ICML)*, volume 70 of *Proceedings of Machine Learning Research*, 2017.

Quentin Berthet, Mathieu Blondel, Olivier Teboul, Marco Cuturi, Jean-Philippe Vert, and Francis Bach. Learning with differentiable perturbed optimizers. In *Advances in Neural Information Processing Systems 33 (NeurIPS)*, 2020.

David Brownstone and Kenneth Train. Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics*, 89(1–2):109–129, 1998.

Harry Buhrman and Ronald de Wolf. Complexity measures and decision tree complexity: A survey. *Theoretical Computer Science*, 288(1):21–43, 2002.

Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

Sourav Chakraborty, Anna Gál, Sophie Laplante, Rajat Mittal, and Anupa Sunny. Certificate games. In *14th Innovations in Theoretical Computer Science Conference (ITCS)*, volume 251 of *LIPIcs*, 2023.

Moses Charikar, Chirag Pabbaraju, and Ambuj Tewari. From non-convex to strongly convex: Curvature-adaptive FTPL for online optimization, 2026. arXiv:2606.02948.

Alon Cohen and Tamir Hazan. Following the perturbed leader for online structured learning. In *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, volume 37 of *Proceedings of Machine Learning Research*, pp. 1034–1042, 2015.

Hedda Cohen Indelman and Tamir Hazan. On the statistical representation properties of the Perturb-Softmax and the Perturb-Argmax probability distributions, 2024. arXiv:2406.02180.

Andrew Daly and Stan Zachary. Improved multiple choice models. In David A. Hensher and M. Q. Dalvi (eds.), *Determinants of Travel Choice*, pp. 335–357. Saxon House, 1978.

Luc Devroye, Gábor Lugosi, and Gergely Neu. Random-walk perturbations for online combinatorial optimization. *IEEE Transactions on Information Theory*, 61:4099–4106, 2015.

Miroslav Dudík, Nika Haghtalab, Haipeng Luo, Robert E. Schapire, Vasilis Syrgkanis, and Jennifer Wortman Vaughan. Oracle-efficient online learning and auction design. *Journal of the ACM*, 67(5), 2020.

Mogens Fosgerau, Daniel McFadden, and Michel Bierlaire. Choice probability generating functions. *Journal of Choice Modelling*, 8:1–18, 2013.

Andreea Gane, Tamir Hazan, and Tommi Jaakkola. Learning with maximum a-posteriori perturbation models. In *Proceedings of the 17th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 33 of *Proceedings of Machine Learning Research*, 2014.

Justin Gilmer, Michael Saks, and Srikanth Srinivasan. Composition limits and separating examples for some Boolean function complexity measures. In *Proceedings of the 28th IEEE Conference on Computational Complexity (CCC)*, pp. 185–196, 2013.

James Hannan. Approximation to Bayes risk in repeated play. In Melvin Dresher, Albert W. Tucker, and Philip Wolfe (eds.), *Contributions to the Theory of Games*, volume 3, pp. 97–139. Princeton University Press, 1957.

David Haussler, Jyrki Kivinen, and Manfred K. Warmuth. Sequential prediction of individual sequences under general loss functions. *IEEE Transactions on Information Theory*, 44(5):1906–1925, 1998.

Elad Hazan and Tomer Koren. The computational power of optimization in online learning. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC)*, 2016. arXiv:1504.02089.

Tamir Hazan and Tommi Jaakkola. On the partition function and random maximum a-posteriori perturbations. In *Proceedings of the 29th International Conference on Machine Learning (ICML)*, 2012.

Tamir Hazan, Subhransu Maji, and Tommi Jaakkola. On sampling from the Gibbs distribution with random maximum a-posteriori perturbations. In *Advances in Neural Information Processing Systems 26 (NIPS)*, 2013.

Tamir Hazan, George Papandreou, and Daniel Tarlow (eds.). *Perturbations, Optimization, and Statistics*. MIT Press, 2016.

David P. Helmbold and Manfred K. Warmuth. Learning permutations with exponential weights. *Journal of Machine Learning Research*, 10:1705–1736, 2009.

Josef Hofbauer and William H. Sandholm. On the global convergence of stochastic fictitious play. *Econometrica*, 70(6):2265–2294, 2002.

Junya Honda, Shinji Ito, and Taira Tsuchiya. Follow-the-Perturbed-Leader achieves best-of-both-worlds for bandit problems. In *Proceedings of the 34th International Conference on Algorithmic Learning Theory (ALT)*, volume 201 of *Proceedings of Machine Learning Research*, pp. 726–754, 2023.

Iris A. M. Huijben, Wouter Kool, Max B. Paulus, and Ruud J. G. van Sloun. A review of the Gumbel-max trick and its extensions for discrete stochasticity in machine learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2):1353–1371, 2023.

Eric Jang, Shixiang Gu, and Ben Poole. Categorical reparameterization with Gumbel-softmax. In *International Conference on Learning Representations (ICLR)*, 2017.

Sham M. Kakade, Adam Tauman Kalai, and Katrina Ligett. Playing games with approximation algorithms. *SIAM Journal on Computing*, 39(3):1088–1106, 2009.

Adam Tauman Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005.

Wouter M. Koolen, Manfred K. Warmuth, and Jyrki Kivinen. Hedging structured concepts. In *Proceedings of the 23rd Conference on Learning Theory (COLT)*, pp. 93–105, 2010.

Raghav Kulkarni and Avishay Tal. On fractional block sensitivity. *Chicago Journal of Theoretical Computer Science*, 2016(8):1–16, 2016.

Jongyeong Lee, Junya Honda, Shinji Ito, and Min-hwan Oh. Follow-the-Perturbed-Leader with Fréchet-type tail distributions: Optimality in adversarial bandits and best-of-both-worlds. In *Proceedings of the 37th Conference on Learning Theory (COLT)*, volume 247 of *Proceedings of Machine Learning Research*, pp. 3375–3430, 2024.

Shachar Lovett and Jiapeng Zhang. Fractional certificates for bounded functions. In *14th Innovations in Theoretical Computer Science Conference (ITCS)*, volume 251 of *LIPIcs*, 2023.

R. Duncan Luce. *Individual Choice Behavior: A Theoretical Analysis*. Wiley, 1959.

Chris J. Maddison, Daniel Tarlow, and Tom Minka. A* sampling. In *Advances in Neural Information Processing Systems 27 (NIPS)*, 2014.

Chris J. Maddison, Andriy Mnih, and Yee Whye Teh. The Concrete distribution: A continuous relaxation of discrete random variables. In *International Conference on Learning Representations (ICLR)*, 2017.

Daniel McFadden. Conditional logit analysis of qualitative choice behavior. In Paul Zarembka (ed.), *Frontiers in Econometrics*. Academic Press, 1974.

Daniel McFadden. Econometric models of probabilistic choice. In Charles F. Manski and Daniel McFadden (eds.), *Structural Analysis of Discrete Data and Econometric Applications*, pp. 198–272. MIT Press, 1981.

Daniel McFadden and Kenneth Train. Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15(5):447–470, 2000.

Gonzalo E. Mena, David Belanger, Scott Linderman, and Jasper Snoek. Learning latent permutations with Gumbel-Sinkhorn networks. In *International Conference on Learning Representations (ICLR)*, 2018.

Zakaria Mhammedi and Robert C. Williamson. Constant regret, generalized mixability, and mirror descent. In *Advances in Neural Information Processing Systems 31 (NeurIPS)*, 2018.

Mathias Niepert, Pasquale Minervini, and Luca Franceschi. Implicit MLE: Backpropagating through discrete exponential family distributions. In *Advances in Neural Information Processing Systems 34 (NeurIPS)*, 2021.

Noam Nisan. CREW PRAMs and decision trees. *SIAM Journal on Computing*, 20(6):999–1007, 1991.

Francesco Orabona, Tamir Hazan, Anand D. Sarwate, and Tommi Jaakkola. On measure concentration of random maximum a-posteriori perturbations. In *Proceedings of the 31st International Conference on Machine Learning (ICML)*, volume 32 of *Proceedings of Machine Learning Research*, pp. 432–440, 2014.

George Papandreou and Alan L. Yuille. Perturb-and-MAP random fields: Using discrete optimization to learn and sample from energy models. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 193–200, 2011.

Max B. Paulus, Dami Choi, Daniel Tarlow, Andreas Krause, and Chris J. Maddison. Gradient estimation with stochastic softmax tricks. In *Advances in Neural Information Processing Systems 33 (NeurIPS)*, 2020.

Alexander Rakhlin, Ohad Shamir, and Karthik Sridharan. Relax and randomize: From value to algorithms. In *Advances in Neural Information Processing Systems 25 (NIPS)*, 2012.

Mark D. Reid, Rafael M. Frongillo, Robert C. Williamson, and Nishant Mehta. Generalized mixability via entropic duality. In *Proceedings of the 28th Conference on Learning Theory (COLT)*, volume 40 of *Proceedings of Machine Learning Research*, pp. 1501–1522, 2015.

Jesper R.-V. Sørensen and Mogens Fosgerau. How McFadden met Rockafellar and learned to do more with less. *Journal of Mathematical Economics*, 100:102629, 2022.

Arun Sai Suggala and Praneeth Netrapalli. Online non-convex learning: Following the perturbed leader is optimal. In *Proceedings of the 31st International Conference on Algorithmic Learning Theory (ALT)*, volume 117 of *Proceedings of Machine Learning Research*, pp. 845–861, 2020.

Vasilis Syrgkanis, Akshay Krishnamurthy, and Robert E. Schapire. Efficient algorithms for adversarial contextual learning. In *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, volume 48 of *Proceedings of Machine Learning Research*, 2016.

Eiji Takimoto and Manfred K. Warmuth. Path kernels and multiplicative updates. *Journal of Machine Learning Research*, 4:773–818, 2003.

Avishay Tal. Properties and applications of Boolean function composition. In *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science (ITCS)*, pp. 441–454, 2013.

Daniel Tarlow, Ryan P. Adams, and Richard S. Zemel. Randomized optimum models for structured prediction. In *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 22 of *Proceedings of Machine Learning Research*, pp. 1221–1229, 2012.

Jakub M. Tomczak. On some properties of the low-dimensional Gumbel perturbations in the Perturb-and-MAP model. *Statistics & Probability Letters*, 115:8–15, 2016.

Tim van Erven, Mark D. Reid, and Robert C. Williamson. Mixability is Bayes risk curvature relative to log loss. *Journal of Machine Learning Research*, 13:1639–1663, 2012.

Tim van Erven, Wojciech Kotłowski, and Manfred K. Warmuth. Follow the leader with dropout perturbations. In *Proceedings of the 27th Conference on Learning Theory (COLT)*, volume 35 of *Proceedings of Machine Learning Research*, pp. 949–974, 2014.

Vladimir Vovk and Fedor Zhdanov. Prediction with expert advice for the Brier game. *Journal of Machine Learning Research*, 10:2445–2471, 2009.

Volodya Vovk. Aggregating strategies. In *Proceedings of the Third Annual Workshop on Computational Learning Theory (COLT)*, 1990.

Volodya Vovk. A game of prediction with expert advice. *Journal of Computer and System Sciences*, 56(2):153–173, 1998.

Guanghui Wang, Zihao Hu, Vidya Muthukumar, and Jacob Abernethy. Adaptive oracle-efficient online learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022. arXiv:2210.09385.

H. C. W. L. Williams. On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A*, 9(3):285–344, 1977.

John I. Yellott, Jr. The relationship between Luce's choice axiom, Thurstone's theory of comparative judgment, and the double exponential distribution. *Journal of Mathematical Psychology*, 15(2): 109–144, 1977.

A Regularity and the Curvature Criterion

This appendix proves the regularity of the potentials and the curvature criterion of Section 2, which turns a bound on the Hessian of A_θ into (1) and back; every upper bound by certificates passes through it. The proofs work with scores on actions. Throughout the appendix, $x \in \mathbb{R}^V$ is a score vector on actions, $J\colon \mathbb{R}^Y \to \mathbb{R}^V$ is the map $(Jg)_v := g_{M(v)}$, so that the mechanism at (θ, s) runs at $x = \theta + Js$ and $J\mathbf{1}_Y = \mathbf{1}_V$, and $F_y := M^{-1}(y)$ is the *fiber* of the output y. The *action potential* is $A(x) := \mathbb{E}\max_{v\in V}\{x_v + \Xi(v)\}$, so that $A_\theta(s) = A(\theta + Js)$. The *winner distribution* is $r(x) \in \mathbb{R}^V$, $r_v(x) := \Pr[v^* = v]$, so that $p(\theta, s) = J^\top r(\theta + Js)$. For distinct actions u and v,

$$\Xi(v) - \Xi(u) \;=\; \sum_{i\in\Delta(u,v)} \big(\xi_i(v_i) - \xi_i(u_i)\big), \tag{A.1}$$

a sum of $|\Delta(u,v)| \geq 1$ independent standard logistic random variables. The winner is therefore almost surely unique, and any such difference has a density bounded by $\frac{1}{4}$, the maximum of the logistic density.

The following lemma passes from curvature to increments, and it is where the regularity of the mechanism is used.

Lemma A.1 (from curvature to increments)**.** *Let $B\colon \mathbb{R}^Y \to \mathbb{R}$ be continuously differentiable with $B(s + b\mathbf{1}_Y) = B(s) + b$ for all $s \in \mathbb{R}^Y$ and $b \in \mathbb{R}$, and suppose its gradient $q := \nabla B$ is locally Lipschitz and takes values in the set of probability vectors on Y. If $\kappa > 0$ satisfies*

$$\nabla^2 B(s) \;\preceq\; \kappa\, \mathrm{Cov}\big(q(s)\big), \qquad \mathrm{Cov}(q) := \mathrm{diag}(q) - qq^\top,$$

at almost every s, then $B(s+g) - B(s) \leq \kappa^{-1}\log\sum_y q_y(s)e^{\kappa g_y}$ for all $s, g \in \mathbb{R}^Y$.

Proof. Written for the exponentials of the scores, the hypothesis says that a positively homogeneous function is concave, and the conclusion is its gradient inequality. Define Φ on the open positive orthant of $\mathbb{R}^Y$ by

$$\Phi(\omega) \;:=\; \exp\big(\kappa B(s(\omega))\big), \qquad s(\omega) := \kappa^{-1}\log\omega,$$

the logarithm taken entrywise. Equivariance of B gives $\Phi(\lambda\omega) = \lambda\Phi(\omega)$ for every $\lambda > 0$, so Φ is positively homogeneous of degree one and Euler's identity reads $\langle\nabla\Phi(\omega), \omega\rangle = \Phi(\omega)$. Differentiating once gives $\partial\Phi/\partial\omega_y = \Phi(\omega)q_y(s)/\omega_y$, and differentiating a second time gives

$$\nabla^2\Phi(\omega) \;=\; \frac{\Phi(\omega)}{\kappa}\,\mathrm{diag}(\omega)^{-1}\big[\nabla^2 B(s) - \kappa\,\mathrm{Cov}(q(s))\big]\,\mathrm{diag}(\omega)^{-1},$$

so the hypothesis states precisely that $\nabla^2\Phi \preceq 0$ almost everywhere. The gradient of Φ is a product of locally Lipschitz functions on the orthant and is therefore locally Lipschitz, so this pointwise Hessian is also the distributional one. Let Ω be a bounded convex open set whose closure lies in the orthant, and let ψ_ε be a mollifier. On Ω the function $\Phi * \psi_\varepsilon$ is smooth with Hessian $(\nabla^2\Phi) * \psi_\varepsilon \preceq 0$, hence concave, and it converges to Φ uniformly on compact subsets, so Φ is concave on Ω. Every pair of points of the orthant lies in such a set, so Φ is concave.

Concavity and Euler's identity give $\Phi(\omega') \leq \Phi(\omega) + \langle\nabla\Phi(\omega), \omega' - \omega\rangle = \langle\nabla\Phi(\omega), \omega'\rangle$ for all ω, ω' in the orthant. Taking $\omega = e^{\kappa s}$ and $\omega' = e^{\kappa(s+g)}$ entrywise, so that $\omega'_y/\omega_y = e^{\kappa g_y}$, turns this into

$$e^{\kappa B(s+g)} \;\leq\; e^{\kappa B(s)} \sum_{y\in Y} q_y(s)\, e^{\kappa g_y},$$

which is the claim. □

The next proposition verifies those hypotheses for the potentials of the mechanism and records the criterion they yield.

Proposition A.2 (regularity and the curvature criterion)**.** *Let M be a readout on V.*

(i) *A is differentiable on $\mathbb{R}^V$ with $\nabla A = r$, and r is Lipschitz. Consequently $\Lambda := \nabla^2 A$ exists at almost every score vector, is symmetric and positive semidefinite, and satisfies $\Lambda\mathbf{1}_V = 0$ and $\Lambda_{uv} \leq 0$ for $u \neq v$, so that its quadratic form is*

$$h^\top\Lambda h \;=\; \sum_{\{u,v\}} c_{uv}\,(h_u - h_v)^2, \qquad c_{uv} := -\Lambda_{uv} \geq 0, \quad h \in \mathbb{R}^V,$$

with the couplings c_{uv}.

(ii) *For every θ the potential A_θ is convex, satisfies $A_\theta(s+b\mathbf{1}_Y) = A_\theta(s)+b$, and is continuously differentiable with $\nabla A_\theta = p(\theta,\cdot)$; the output distribution is Lipschitz, satisfies $p_y > 0$ for every $y \in Y$, and $\tau \mapsto p_y(\theta, s+\tau e_y)$ is nondecreasing. The pairs (θ, s) at which $\Lambda(\theta+Js)$ fails to exist form a null set, and at every other pair $p(\theta,\cdot)$ is differentiable at s with derivative $\nabla^2 A_\theta(s) = J^\top \Lambda(\theta+Js)J$.*

(iii) *If $\kappa > 0$ satisfies $J^\top\Lambda(\theta+Js)J \preceq \kappa \operatorname{Cov}\big(p(\theta,s)\big)$ at almost every (θ,s), then $\kappa^*(M) \le \kappa$.*

(iv) *If κ is admissible, then $J^\top\Lambda(\theta+Js)J \preceq \kappa \operatorname{Cov}\big(p(\theta,s)\big)$ at every pair (θ,s) at which $\Lambda(\theta+Js)$ exists. Hence, for a nonconstant readout, $\kappa^*(M)$ is the supremum of $g^\top J^\top\Lambda(\theta+Js)Jg/\operatorname{Var}_{p(\theta,s)}(g)$ over those pairs and over the nonconstant $g \in \mathbb{R}^Y$.*

Proof. (i) The function $x \mapsto \max_v\{x_v + \Xi(v)\}$ is convex and 1-Lipschitz for the supremum norm, and by (A.1) its maximizer is almost surely unique. For almost every realization that maximizer is unchanged under a small enough perturbation of x, so dominated convergence applied to difference quotients gives $\nabla A = r$. Raising the scores of the actions in a set can only move the winner into that set. Applied to a single action this bounds the Lipschitz constant: for $x' = x + \tau e_u$ with $\tau > 0$ the winner either is unchanged or becomes u, so

$$\begin{aligned}\big\|r(x') - r(x)\big\|_1 &= 2\Pr\big[v^*(x) \neq u,\ v^*(x') = u\big] \\ &\le 2\sum_{v\neq u}\Pr\big[0 \le x_v + \Xi(v) - x_u - \Xi(u) \le \tau\big] \ \le\ \tfrac{1}{2}(|V|-1)\,\tau\end{aligned}$$

by the density bound after (A.1). The same estimate applies when a component is lowered, and any two score vectors are joined by $|V|$ such moves, so r is Lipschitz and Rademacher's theorem produces Λ almost everywhere. Symmetry and positive semidefiniteness come from convexity, $\Lambda\mathbf{1}_V = 0$ from equivariance, and $\Lambda_{uv} \le 0$ from the same monotonicity, since r_v is nonincreasing in x_u for $v \neq u$. A symmetric matrix with vanishing row sums and nonpositive off-diagonal entries has the displayed Laplacian form.

(ii) Convexity, equivariance and $\nabla A_\theta(s) = J^\top r(\theta+Js) = p(\theta,s)$ follow from $A_\theta = A(\theta+J\,\cdot\,)$ and $J\mathbf{1}_Y = \mathbf{1}_V$, and p is Lipschitz because r is. The monotonicity along a raise of y is the same principle applied to the set F_y. Every action wins with positive probability: on the event that $\xi_i(v_i) > K$ for every i while $\xi_i(a) < 0$ for every i and every $a \neq v_i$, every other action u has $\Xi(u) < \Xi(v) - K$ by (A.1), so v wins as soon as $K \ge \max_u x_u - x_v$, and the event has positive probability. Hence $p_y > 0$ for every y. Finally, if r is differentiable at $\theta + Js$ then $p(\theta,\cdot) = J^\top r(\theta + J\,\cdot\,)$ is differentiable at s with derivative $J^\top\Lambda(\theta+Js)J$. The score vectors at which Λ fails to exist form a set contained in a Borel null set $\mathcal{N}$; for each fixed s the set $\{\theta : \theta + Js \in \mathcal{N}\}$ is null, so by Fubini the pairs (θ, s) with $\theta + Js \in \mathcal{N}$ form a null set.

(iii) By (ii) and Fubini, for almost every θ the potential A_θ satisfies the hypotheses of Lemma A.1 together with the stated Hessian bound at almost every s, so (1) holds at that θ. Both sides of (1) are continuous in θ, hence it holds at every θ.

(iv) At such a pair, (ii) makes $p(\theta,\cdot)$ differentiable at s with derivative $H := J^\top\Lambda(\theta+Js)J$. Write $p := p(\theta,s)$ and fix $g \in \mathbb{R}^Y$. As $\varepsilon \to 0$, the left side of (1) at the increment εg is $\int_0^\varepsilon \langle p(\theta, s+\zeta g), g\rangle\, d\zeta = \varepsilon\langle p, g\rangle + \frac{1}{2}\varepsilon^2 g^\top H g + o(\varepsilon^2)$, and its right side is $\varepsilon\langle p, g\rangle + \frac{1}{2}\kappa\varepsilon^2 g^\top \operatorname{Cov}(p) g + O(\varepsilon^3)$; comparing the terms of order ε^2 gives $g^\top H g \le \kappa\, g^\top \operatorname{Cov}(p) g$. For the supremum $\bar\kappa$ of the last claim, note that $g^\top \operatorname{Cov}(p) g = \operatorname{Var}_p(g)$ is positive for nonconstant g, because $p_y > 0$ for every y by (ii). Every $\kappa > \kappa^*(M)$ is admissible, so $\bar\kappa \le \kappa^*(M)$; and if $\bar\kappa$ is finite, the bound at $\bar\kappa$ holds at every pair at which Λ exists, hence at almost every pair, and (iii) gives $\kappa^*(M) \le \bar\kappa$. □

B Proof of Theorem 1

This appendix proves Theorem 1, with the notation of Appendix A. The inequality of (a) comes from moving the noise of each winner along a certificate, and the dual of the certificate program shows that its constant cannot be lowered; (b) compares that rate of growth with a second-order expansion of (1); (c) combines the curvature bound of compatible certificates (Lemma B.1), which holds on every decision set, with the hybrid argument.

Proof of Theorem 1(a). For the inequality, we shift the noise of each winner of output y along a certificate and bound the cost of the shift by the following fact: for a standard Gumbel G, a Borel set $I \subseteq \mathbb{R}$ and $\tau \geq 0$,

$$\Pr[G \in I] \;\leq\; e^{\tau} \Pr[G \in I + \tau], \tag{B.1}$$

because the density $f(\zeta) = \exp(-\zeta - e^{-\zeta})$ has $(\log f)'(\zeta) = e^{-\zeta} - 1 \geq -1$, so $f(\zeta - \tau) \leq e^{\tau} f(\zeta)$ for every ζ, and integrating over $I + \tau$ gives (B.1).

Fix $v \in F_y$ and an optimal fractional certificate z at v, so that $\sum_i z_i = \mathrm{FC}(M, v)$. Let E_v be the event that v wins at output scores $s + \tau e_y$, and let E'_v be its image under the translation that adds τz_i to $\xi_i(v_i)$ for every i and leaves all other Gumbels fixed. The translation carries E_v into the event that v wins at output scores s. It raises $\Xi(u)$ by $\tau \sum_{i : u_i = v_i} z_i$, hence raises the score of v relative to that of u by $\tau \sum_{i \in \Delta(u,v)} z_i$. On E_v, at output scores s, the score of v exceeds that of u minus τ when $M(u) \neq y$, and exceeds that of u when $M(u) = y$; in the first case z being a certificate gives $\sum_{i \in \Delta(u,v)} z_i \geq 1$, and the translation is nonnegative in the second, so v beats u after it. The translation moves n independent Gumbels by the nonnegative amounts τz_i, so applying (B.1) to one of them at a time gives $\Pr[E_v] \leq e^{\tau \, \mathrm{FC}(M,v)} \Pr[E'_v]$. The winner is almost surely unique, so p_y is the sum over F_y of the winning probabilities; summing over $v \in F_y$ and bounding $\mathrm{FC}(M, v)$ by $\mathrm{FC}(M)$ gives the inequality of (a).

For the converse we lower the score of a costliest action's output, suppress the other actions of that output, and read the decay rate off the dual of the certificate program. Fix an action v with $\mathrm{FC}(M, v) = \mathrm{FC}(M) =: \varphi$ and write $y := M(v)$. Let ω be an optimal solution of the dual program (2) at v, which weights the opponents of v, the actions of a different output, with total weight φ. For $N > 0$ take the base scores

$$\theta_v = 0, \qquad \theta_u = 0 \;\text{ for } M(u) \neq y, \qquad \theta_u = -\varphi N \;\text{ for } u \in F_y \setminus \{v\},$$

and compare the output scores $-Ne_y$ and 0.

At output scores $-Ne_y$ the action v wins only if $\Xi(v) - \Xi(u) > N$ for every u with $M(u) \neq y$. The dual weights turn that family of inequalities into a single one about independent random variables: weighting by ω_u and summing gives $\sum_i X_i \geq \varphi N$, where

$$X_i \;:=\; \sum_{u :\, u_i \neq v_i} \omega_u \big(\xi_i(v_i) - \xi_i(u_i)\big) \;=\; q_i \, \xi_i(v_i) - \sum_{a \neq v_i} \omega_a^{(i)} \, \xi_i(a), \qquad q_i := \sum_{u :\, u_i \neq v_i} \omega_u \leq 1,$$

and $\omega_a^{(i)}$ collects the weight of the opponents using value a at variable i, so that $\sum_{a \neq v_i} \omega_a^{(i)} = q_i$. The random variables $X_1, \ldots, X_n$ are independent, since X_i involves only the noise of variable i.

Both tails are controlled by one estimate. If G is a standard Gumbel then $Z := e^{-G}$ is standard exponential and $\mathbb{E}[e^{\nu G}] = \mathbb{E}[Z^{-\nu}]$, so Jensen's inequality gives $\mathbb{E}[e^{-\nu G}] \leq 1$ for $\nu \in [0, 1]$, while splitting $\int_0^{\infty} \zeta^{-\nu} e^{-\zeta} d\zeta$ at $\zeta = 1$ gives

$$\mathbb{E}\big[e^{\nu G}\big] \;\leq\; K_\nu \;:=\; \frac{1}{1 - \nu} + e^{-1} \qquad \text{for } \nu \in [0, 1),$$

with K_ν nondecreasing in ν. Fix $\nu \in (0, 1)$. In $\mathbb{E}[e^{\nu X_i}]$ the Gumbel $\xi_i(v_i)$ has the positive multiplier $\nu q_i \leq \nu$ and each $\xi_i(a)$ has a negative multiplier of magnitude at most $\nu q_i \leq 1$, so $\mathbb{E}[e^{\nu X_i}] \leq K_\nu$ and Chernoff's bound gives $\Pr[v \text{ wins}] \leq K_\nu^n e^{-\nu \varphi N}$.

An action $u \in F_y \setminus \{v\}$ wins only if it beats v, that is, only if $\Xi(u) - \Xi(v) > \varphi N$. That difference is a sum of at most n independent standard logistic random variables, each a difference of two independent Gumbels and so with moment generating function at most K_ν at ν. Adding the estimates over the fiber,

$$p_y(\theta, -Ne_y) \;\leq\; C_\nu \, e^{-\nu \varphi N}, \qquad C_\nu := |V| \, K_\nu^n.$$

At output scores 0 the suppression only helps v, so $p_y(\theta, 0) \geq \Pr[\Xi(v) > \Xi(u) \text{ for all } u \neq v] =: P_0$, a positive constant independent of N by Proposition A.2(ii). Suppose the inequality of (a) holds with some constant in place of $\mathrm{FC}(M)$. With $s = -Ne_y$ and $\tau = N$ that constant is at least

$$\frac{1}{N} \log \frac{p_y(\theta, 0)}{p_y(\theta, -Ne_y)} \;\geq\; \nu\varphi + \frac{1}{N} \log \frac{P_0}{C_\nu} \;\xrightarrow[N \to \infty]{}\; \nu\varphi,$$

and letting $\nu \uparrow 1$ shows that it is at least φ. $\square$

The Gumbel enters the inequality of (a) only through (B.1). For a product of independent copies of any noise law μ with $\mu(I) \le e^{\lambda_\mu \tau}\mu(I+\tau)$ for all Borel I and $\tau \ge 0$, the same argument gives the constant $\lambda_\mu\,\mathrm{FC}(M)$; such a law has no atoms, since an atom would force one at every translate, so the winner is again almost surely unique.

Proof of Theorem 1(b). If M is constant both sides are 0. Otherwise fix θ, s and y and follow the output distribution along the raise of y: by Proposition A.2(ii) the map $\tau \mapsto p_y(\theta, s+\tau e_y)$ is Lipschitz, nondecreasing and takes values in $(0,1)$, hence is differentiable at almost every τ. Let κ be admissible in (1). At such a τ, apply (1) at output scores $s+\tau e_y$ with the increment εe_y and expand in $\varepsilon > 0$: the left side is $\int_0^\varepsilon p_y(\theta, s+(\tau+\zeta)e_y)\,d\zeta$, whose expansion begins $p_y\varepsilon + \frac{1}{2}\frac{d}{d\tau}p_y \cdot \varepsilon^2$, while the right side begins $p_y\varepsilon + \frac{\kappa}{2}p_y(1-p_y)\varepsilon^2$. Comparing the quadratic terms gives $\frac{d}{d\tau}p_y \le \kappa\, p_y(1-p_y)$, so the logit of the curve has derivative at most κ almost everywhere. That logit is locally Lipschitz, hence the integral of its derivative, so

$$\operatorname{logit} p_y(\theta, s+(\tau+\varepsilon)e_y) - \operatorname{logit} p_y(\theta, s+\tau e_y) \;\le\; \kappa\varepsilon \qquad \text{for all } \tau, \varepsilon \ge 0.$$

The curve is nondecreasing, so $\log p_y$ obeys the same bound, and κ has the property of Theorem 1(a) in place of $\mathrm{FC}(M)$; by (a), $\kappa \ge \mathrm{FC}(M)$. As this holds for every admissible κ, $\mathrm{FC}(M) \le \kappa^*(M)$. □

The upper bound of Section 3 is a curvature bound, and compatible certificates give it.

Lemma B.1 (compatible certificates bound the curvature)**.** *Let $(z^v)_{v\in V}$ be compatible certificates (Definition 4) of weight W. Then at every score vector x at which Λ exists, and for every $g \in \mathbb{R}^Y$,*

$$(Jg)^\top \Lambda\,(Jg) \;\le\; W\, g^\top \operatorname{Cov}(p)\, g, \qquad p := J^\top r(x).$$

Proof. We write each coupling through the softmax left by conditioning on the noise outside one variable, split it along the overlap of the two certificates, and bound the split sums at each variable by a weighted variance inequality.

Fix a variable i and condition on $\mathcal{F}_i$, the noise outside variable i. For each value $a \in D_i$, the best action of the slice $\{v : v_i = a\}$ and its score $\tilde\theta_a := \max\{x_v + \sum_{j\ne i}\xi_j(v_j) : v_i = a\}$ are determined by $\mathcal{F}_i$, and the winner is the best action of the value maximizing $\tilde\theta_a + \xi_i(a)$. The $\xi_i(a)$ are independent standard Gumbels, so the winning value is the softmax of $(\tilde\theta_a)_{a\in D_i}$ (McFadden, 1974). Write $r^{(i)}$ for the resulting conditional winner distribution, so that $\mathbb{E}[r^{(i)}] = r$.

Let $u \ne v$ with $i \in \Delta(u,v)$. Raising x_u moves no slice score other than $\tilde\theta_{u_i}$, and that one only when u is the best action of its slice, and it leaves the best action of the slice of v_i untouched, so $r_v^{(i)}$ varies only through the softmax, with $\partial r_v^{(i)}/\partial x_u = -r_u^{(i)} r_v^{(i)}$. The difference quotients are bounded by $\frac{1}{4}$, so dominated convergence turns this into

$$c_{uv} \;=\; \mathbb{E}\big[r_u^{(i)}\, r_v^{(i)}\big] \qquad \text{for every } i \in \Delta(u,v). \tag{B.2}$$

The weighted variance inequality reads, for a probability vector ρ on V, a function h on V and any constant $\bar h$:

$$\sum_{\{u,v\}} \min\{z_i^u, z_i^v\}\,\rho_u\rho_v(h_u - h_v)^2 \;\le\; \sum_v \rho_v z_i^v (h_v - \bar h)^2. \tag{B.3}$$

Indeed $\min\{z_i^u, z_i^v\} = \int_0^\infty \mathbb{1}\{z_i^u > \zeta\}\mathbb{1}\{z_i^v > \zeta\}\,d\zeta$, and for any set $U \subseteq V$,

$$\sum_{\{u,v\}\subseteq U} \rho_u\rho_v(h_u - h_v)^2 \;=\; \rho(U)\sum_{v\in U}\rho_v(h_v-\bar h)^2 - \Big(\sum_{v\in U}\rho_v(h_v-\bar h)\Big)^2 \;\le\; \sum_{v\in U}\rho_v(h_v-\bar h)^2;$$

integrating this over ζ with $U = \{v : z_i^v > \zeta\}$ gives (B.3).

Now put $h := Jg$ and $\bar h := \mathbb{E}_r[h]$. By Proposition A.2 the form $(Jg)^\top\Lambda(Jg)$ is $\sum_{\{u,v\}} c_{uv}(h_u - h_v)^2$, in which only pairs with different outputs contribute. For such a pair, compatibility and (B.2)

give $c_{uv} \le \sum_{i\in\Delta(u,v)} \min\{z_i^u, z_i^v\}\mathbb{E}[r_u^{(i)} r_v^{(i)}]$. Summing over pairs, enlarging each inner sum to all pairs, and applying (B.3) with $\rho = r^{(i)}$ at every i,

$$(Jg)^\top \Lambda (Jg) \;\le\; \sum_i \mathbb{E}\Big[\sum_{\{u,v\}} \min\{z_i^u, z_i^v\} r_u^{(i)} r_v^{(i)} (h_u - h_v)^2\Big] \;\le\; \sum_i \sum_v \mathbb{E}\big[r_v^{(i)}\big] z_i^v (h_v - \bar{h})^2.$$

Since $\mathbb{E}[r^{(i)}] = r$ for every i, the right side is $\sum_v r_v \big(\sum_i z_i^v\big)(h_v - \bar{h})^2 \le W \operatorname{Var}_r(h)$, and $\operatorname{Var}_r(Jg) = g^\top \operatorname{Cov}(p) g$. □

Proof of Theorem 1(c). The hybrid of two actions whose certificates overlap too little would have both outputs at once. Let V be a product and let z^v be an optimal fractional certificate at each action. Suppose some pair with $M(u) \ne M(v)$ had $\sum_{i\in\Delta(u,v)} \min\{z_i^u, z_i^v\} < 1$, and form the hybrid w of Section 3, $w_i := u_i$ if $z_i^u > z_i^v$ and $w_i := v_i$ otherwise, which is an action because V is a product. A variable at which w differs from v has $z_i^u > z_i^v$, so it lies in $\Delta(u,v)$ and its z_i^v is the smaller of the two weights; a variable at which w differs from u has $z_i^u \le z_i^v$, so it too lies in $\Delta(u,v)$ and its z_i^u is the smaller. The two sets of variables are disjoint, so

$$\sum_{i\in\Delta(w,v)} z_i^v + \sum_{i\in\Delta(w,u)} z_i^u \;\le\; \sum_{i\in\Delta(u,v)} \min\{z_i^u, z_i^v\} \;<\; 1.$$

Both sums on the left are therefore below 1, and z^v and z^u being certificates force $M(w) = M(v)$ and $M(w) = M(u)$, which contradicts $M(u) \ne M(v)$. The optimal certificates are therefore compatible, of weight $\mathrm{FC}(M)$, so Lemma B.1 bounds $J^\top \Lambda J$ by $\mathrm{FC}(M)\operatorname{Cov}(p)$ wherever Λ exists, hence at almost every (θ, s) by Proposition A.2(ii), and Proposition A.2(iii) gives $\kappa^*(M) \le \mathrm{FC}(M)$. With (b) the two are equal. □

C Proofs of Propositions 1 and 2

This appendix proves Propositions 1 and 2 and computes the potential range quoted in Section 3. Part (a) of Proposition 1 at the critical step needs (1) at $\kappa^*(M)$ itself, so we first show that the admissible constants of Definition 2 form a half-line closed at $\kappa^*(M)$ when M is nonconstant. Indeed $\kappa^*(M) \ge \mathrm{FC}(M) \ge 1$ by Theorem 1(b), and $\kappa^*(M) \le n$ by Lemma B.1 and Proposition A.2(iii), because the all-ones vectors are compatible certificates of weight n on every decision set, two actions with different outputs differing on at least one variable. The right side of (1) is continuous and nondecreasing in $\kappa > 0$, so if (1) holds along a sequence of constants decreasing to $\kappa^*(M)$ it holds at $\kappa^*(M)$.

Proof of Proposition 1(a). We telescope (1) along the learner's output scores. It holds at $\kappa^*(M)$ and therefore at every $\kappa \ge \kappa^*(M)$, the right side being nondecreasing in κ. Take $\kappa = \eta/\alpha \ge \kappa^*(M)$ and $g = -\alpha \ell_t$ at output scores s_t:

$$A_\theta(s_{t+1}) - A_\theta(s_t) \;\le\; \frac{\alpha}{\eta} \log \sum_y p_y(\theta, s_t) e^{-\eta \ell_t(y)} \;=\; -\alpha m_t.$$

Summing over $t \le T$ and writing $L_T := \sum_{t\le T} \ell_t$ gives $\alpha \sum_{t\le T} m_t \le A_\theta(0) - A_\theta(-\alpha L_T)$. Restricting the maximum defining $A_\theta(-\alpha L_T)$ to the fiber of a minimizing output $y^\star$ of L_T,

$$A_\theta(-\alpha L_T) \;\ge\; -\alpha L_T(y^\star) + \mathbb{E} \max_{v\in F_{y^\star}} \{\theta_v + \Xi(v)\} \;\ge\; -\alpha \min_y L_T(y) + \min_y \mathbb{E} \max_{v\in F_y} \{\theta_v + \Xi(v)\},$$

so $\alpha \sum_{t\le T} m_t \le \Gamma + \alpha \min_y L_T(y)$, which is the claim. □

Part (b) uses a two-round cycle of losses that restores the output distribution and charges every output the same total loss; the following lemma computes the learner's excess mix loss over the cycle. Write $\langle p, g\rangle$ and $\langle p, g^2\rangle$ for the mean of g and of g^2 under a probability vector p.

Lemma C.1 (the two-round cycle)**.** *Fix a base score θ, output scores $s_0 \in \mathbb{R}^Y$ and a nonzero $g \in \mathbb{R}^Y$, and suppose that $\varepsilon \mapsto \langle p(\theta, s_0 - \varepsilon g), g\rangle$ is differentiable at $\varepsilon = 0$ with derivative $-q$. Let the learner, at output scores $s_0 + b\mathbf{1}$ for some $b \in \mathbb{R}$, be charged $\ell^+ := \frac{1}{2}\mathbf{1} + dg$ and then*

$\ell^- := \frac{1}{2}\mathbf{1} - dg$, *where* $0 < d \le 1/(2\max_y |g_y|)$. *After the two rounds its output scores are* $s_0 + (b-\alpha)\mathbf{1}$, *every output's cumulative loss has grown by exactly* 1, *and its two mix losses sum to* $1 + \psi(d)$, *where*

$$\psi(d) \;=\; d^2\big(\alpha q - \eta \operatorname{Var}_{p(\theta,s_0)}(g)\big) + o(d^2) \qquad (d \to 0).$$

Proof. We expand the two mix losses to second order in d. Both loss vectors lie in $[0,1]^Y$ by the bound on d. Write $p_0 := p(\theta, s_0)$ and $p_1 := p(\theta, s_0 - \alpha d g)$. Adding a constant to all output scores leaves the output distribution unchanged, so the learner predicts with p_0 in the first round and, after the update $s_0 + b\mathbf{1} - \alpha\ell^+ = s_0 + (b-\alpha/2)\mathbf{1} - \alpha d g$, with p_1 in the second; the second update returns the scores to $s_0 + (b-\alpha)\mathbf{1}$, and $\ell^+_y + \ell^-_y = 1$ for every y. The two mix losses are $\frac{1}{2} - \eta^{-1}\log\sum_y p_{0,y}e^{-\eta d g_y}$ and $\frac{1}{2} - \eta^{-1}\log\sum_y p_{1,y}e^{\eta d g_y}$, so

$$\psi(d) \;=\; -\frac{1}{\eta}\Big[\log\sum_y p_{0,y}e^{-\eta d g_y} + \log\sum_y p_{1,y}e^{\eta d g_y}\Big].$$

The first logarithm is $-\eta d\langle p_0, g\rangle + \frac{1}{2}\eta^2 d^2 \operatorname{Var}_{p_0}(g) + O(d^3)$. In the second, $\langle p_1, g\rangle = \langle p_0, g\rangle - \alpha q d + o(d)$ by hypothesis and $\langle p_1, g^2\rangle = \langle p_0, g^2\rangle + O(d)$ because p is Lipschitz (Proposition A.2), so

$$\sum_y p_{1,y}e^{\eta d g_y} \;=\; 1 + \eta d\langle p_0, g\rangle - \eta\alpha q d^2 + \tfrac{1}{2}\eta^2 d^2\langle p_0, g^2\rangle + o(d^2)$$

and its logarithm is $\eta d\langle p_0, g\rangle - \eta\alpha q d^2 + \frac{1}{2}\eta^2 d^2 \operatorname{Var}_{p_0}(g) + o(d^2)$. Adding the two logarithms leaves $\eta^2 d^2 \operatorname{Var}_{p_0}(g) - \eta\alpha q d^2 + o(d^2)$, which is the claim. □

Proof of Proposition 1(b). We locate a point of excess curvature, drive the output scores there, and repeat the cycle of Lemma C.1. Let $\alpha > \eta/\kappa^*(M)$ and $\kappa := \eta/\alpha < \kappa^*(M)$. By Proposition A.2(iv) there are a pair (θ_0, s_0) at which $\Lambda(\theta_0 + Js_0)$ exists and a $g \in \mathbb{R}^Y$ with $g^\top H g > \kappa\, g^\top \operatorname{Cov}(p_0) g = \kappa \operatorname{Var}_{p_0}(g)$, where $H := J^\top \Lambda(\theta_0 + Js_0) J$ and $p_0 := p(\theta_0, s_0)$. By Proposition A.2(ii), H is the derivative of $p(\theta_0, \cdot)$ at s_0, so $\varepsilon \mapsto \langle p(\theta_0, s_0 - \varepsilon g), g\rangle$ has derivative $-g^\top H g$ at $\varepsilon = 0$, and Lemma C.1 applies at θ_0 with $q = g^\top H g$: since $\alpha q > \alpha\kappa \operatorname{Var}_{p_0}(g) = \eta \operatorname{Var}_{p_0}(g)$, there is a d with $\psi(d) > 0$.

The adversary plays at base score θ_0. Let $b_0 := \max_y s_{0,y}$ and let T_0 be the least positive integer with $(b_0\mathbf{1} - s_0)/(\alpha T_0) \in [0,1]^Y$; charging this constant loss vector for T_0 rounds moves the output scores from 0 to $s_0 - b_0\mathbf{1}$. From then on the adversary alternates ℓ^+ and ℓ^-. After every full cycle the output scores are again s_0 up to a constant, every output's cumulative loss has grown by 1 and the learner has paid $1 + \psi(d)$, so $R_T = R_{T_0} + (T - T_0)\psi(d)/2$ whenever $T - T_0$ is even. After a half cycle the learner has paid at least $\frac{1}{2} - d\max_y |g_y|$ while the smallest cumulative loss has grown by at most $\frac{1}{2} + d\max_y |g_y|$, so $R_{T+1} \ge R_T - 1$ for such T. Hence $R_T \ge R_{T_0} + \lfloor (T - T_0)/2 \rfloor \psi(d) - 1$ for every $T \ge T_0$, and $\liminf_{T\to\infty} R_T/T \ge \psi(d)/2 > 0$. □

When M is one-to-one every base score admits such an adversary, since the score vector $\theta_0 + Js_0$ is then reached from any base score θ by driving the output scores to $J^{-1}(\theta_0 - \theta) + s_0$, up to a constant.

We compute the potential range quoted in Section 3 for the readout of k variables of a product, $M(v) = v_S$ with $|S| = k$. At $\theta = 0$ the perturbation of an action factorizes, so $A_0(0) = \sum_i \mathbb{E}\max_{a\in D_i}\xi_i(a) = \sum_i(\gamma + \log|D_i|)$, the maximum of the $|D_i|$ independent standard Gumbels of variable i being a standard Gumbel shifted by $\log|D_i|$ and γ the Euler constant; for an output $y \in Y = \prod_{i\in S} D_i$ the maximum over its fiber has expectation $k\gamma + \sum_{i\notin S}(\gamma + \log|D_i|)$, so $\Gamma = \sum_{i\in S}\log|D_i| = \log|Y|$, as stated in Section 3.

Proof of Proposition 2. We compute $\mathrm{FC}(M)$ and Γ, then give the two adversaries. The readout has the three outputs $y_0 := (0,1)$, the output of the zero action alone, and the two parity classes of the remaining actions. Every single-variable neighbor of an action has a different output, so $\mathrm{FC}(M) = n$. At $\theta = 0$ the perturbation of an action factorizes, so $A_0(0) = \sum_i \mathbb{E}\max\{\xi_i(0), \xi_i(1)\}$ while the smallest fiber maximum is $\mathbb{E}[\Xi(0)] = \sum_i \mathbb{E}[\xi_i(0)]$; the maximum of two independent standard Gumbels is a standard Gumbel shifted by $\log 2$, whence $\Gamma = n\log 2$.

(i) Let the adversary charge the output y_0 nothing and the two other outputs one in every round, giving the comparator loss 0 and $R_T = \sum_{t \le T} m_t$. Adding a constant to all output scores leaves the output distribution unchanged, so the state in round t is the raise of y_0 by $\tau_t := \alpha(t-1)$. With $\delta_i := \xi_i(1) - \xi_i(0)$, which is standard logistic, the action 0 wins at a raise τ only if τ exceeds $\max_{v \neq 0} \sum_{i : v_i = 1} \delta_i$; that maximum equals $\sum_i \delta_i^+$ unless every δ_i is negative, in which case it is negative while $\sum_i \delta_i^+ = 0$. Hence $p_{y_0} \le \Pr[\sum_i \delta_i^+ \le \tau]$. The random variable δ^+ has mean $\log 2$ and variance $\pi^2/6 - (\log 2)^2 < 1.08^2$, so Cantelli's inequality gives $\Pr[\sum_i \delta_i^+ \le n \log 2 - 1.08\sqrt{n}] < 1/2$. The learner therefore has $p_{y_0}(\theta, s_t) \le 1/2$ in every round with $\tau_t \le \Gamma - 1.08\sqrt{n}$, a quantity that is positive because $n \ge 3$. The mix loss $-\frac{1}{\eta} \log[p_{y_0} + (1 - p_{y_0})e^{-\eta}]$ decreases in p_{y_0}, so those rounds each contribute at least $\beta_\eta := -\frac{1}{\eta} \log \frac{1+e^{-\eta}}{2} > 0$, and the remaining rounds contribute at least 0. There are at least $\min\{T, (\Gamma - 1.08\sqrt{n})/\alpha\}$ of them.

(ii) Let $g \in \{\pm 1\}^Y$ be $+1$ on the two even outputs and -1 on the odd one, and let the adversary alternate $\ell^+ = \frac{1}{2}\mathbf{1} + dg$ and $\ell^- = \frac{1}{2}\mathbf{1} - dg$ from the first round on, with $0 < d \le \frac{1}{2}$. At base score 0 and constant output scores the winner maximizes Ξ alone, so its values at the n variables are independent fair bits, and the two even outputs have probability $\frac{1}{2}$ together, so $\langle p_0, g \rangle = 0$ and $\mathrm{Var}_{p_0}(g) = 1$ at $p_0 := p(0,0)$. Moving the output scores by $-\varepsilon g$ raises the odd output by 2ε relative to the even ones. For $\tau \in \mathbb{R}$, let $Q(\tau)$ be the probability that the odd output wins when its score exceeds those of the two even outputs by τ. Passing from the argmax to the best action of the opposite parity means flipping the variable with the smallest $|\delta_i|$, and the parity of the argmax is independent of those gaps, so for $\tau \ge 0$

$$Q(\tau) \;=\; \tfrac{1}{2} + \tfrac{1}{2} \Pr\big[\min_i |\delta_i| \le \tau\big] \;=\; 1 - \tfrac{1}{2} \Pr\big[|\delta| \ge \tau\big]^n \;=\; 1 - \tfrac{1}{2}\Big(\frac{2}{1+e^{\tau}}\Big)^n.$$

Flipping the value of one variable exchanges the two parity classes and preserves the law of the noise, so $Q(-\tau) = 1 - Q(\tau)$ for $\tau \ge 0$, and Q is differentiable at 0 with $Q'(0) = n/4$. Hence $\langle p(0, -\varepsilon g), g \rangle = 1 - 2Q(2\varepsilon)$ has derivative $-n$ at $\varepsilon = 0$, and Lemma C.1 applies at $s_0 = 0$ with that derivative: the gain per cycle is $\psi(d) = d^2(n\alpha - \eta) + o(d^2)$, positive for some d when $\alpha > \eta/n$. The output distribution recurs after every cycle, so $R_T = T\psi(d)/2$ for even T, a half cycle lowers the regret by at most 1 as in the proof of Proposition 1(b), and $\liminf_{T \to \infty} R_T/T \ge \psi(d)/2 > 0$. □

D Proofs for Section 4

This appendix proves the results of Section 4, with the notation of Appendix A: Theorem 2 and Corollary 1 from certificates, and Theorem 3 by conditioning on part of the noise. It ends with the potential range of the one-hot encoding.

Proof of Theorem 2. (a) Lemma B.1 bounds $J^\top \Lambda J$ by $W \,\mathrm{Cov}(p)$ wherever Λ exists, hence at almost every (θ, s) by Proposition A.2(ii), and Proposition A.2(iii) gives $\kappa^*(M) \le W$.

(b) Differentiating the inequality of Theorem 1(a) at $\tau = 0$, at every pair (θ, s) at which $\Lambda(\theta + Js)$ exists, bounds the diagonal of $H := J^\top \Lambda(\theta + Js) J$ by $H_{yy} \le \mathrm{FC}(M) p_y$. The matrix H inherits vanishing row sums and nonpositive off-diagonal entries from Λ, so with $\bar{g} := \mathbb{E}_p[g]$,

$$\begin{aligned} g^\top H g \;=\; \sum_{\{y,y'\}} (-H_{yy'})(g_y - g_{y'})^2 \;&\le\; 2 \sum_{\{y,y'\}} (-H_{yy'}) \big[(g_y - \bar{g})^2 + (g_{y'} - \bar{g})^2\big] \\ &=\; 2 \sum_y H_{yy} (g_y - \bar{g})^2 \;\le\; 2\,\mathrm{FC}(M)\,\mathrm{Var}_p(g). \end{aligned}$$

Proposition A.2(iii) concludes.

(c) Let z^v be the vector with ones on S and zeros elsewhere, for every v. If $M(u) \neq M(v)$ then $u_S \neq v_S$, so $\sum_{i \in \Delta(u,v)} \min\{z_i^u, z_i^v\} = |\Delta(u,v) \cap S| \ge 1$: the family is compatible, of weight $|S|$, and (a) gives $\kappa^*(M) \le |S|$. □

Proof of Corollary 1. Transpositions give the lower bound and two explicit families the upper bound. Write $k := |S|$. Exchanging the columns of a row $i \in S$ and of a row $j \neq i$ changes the

readout and alters no other row, so every fractional certificate z at any action satisfies $z_i + z_j \geq 1$ for all $i \in S$ and $j \neq i$. Let $z_{\min}$ be the smallest weight of z. If $z_{\min} > \frac{1}{2}$ then $\sum_i z_i > n/2$. If $z_{\min} \leq \frac{1}{2}$ is attained at a row of S, every other row has weight at least $1 - z_{\min}$, so $\sum_i z_i \geq (n-1) - (n-2)z_{\min} \geq n/2$. If it is attained at a row outside S, every row of S has weight at least $1-z_{\min}$ and every other row at least $z_{\min}$, so $\sum_i z_i \geq k(1-z_{\min})+(n-k)z_{\min} = k+(n-2k)z_{\min}$, which is at least k when $n \geq 2k$ and at least $n/2$ when $n < 2k$. Hence $\mathrm{FC}(M) \geq \min\{k, n/2\}$, and Theorem 1(b) gives $\kappa^*(M) \geq \min\{k, n/2\}$.

Two permutations with different readouts differ on at least two rows and on at least one row of S. The weights $\frac{1}{2}$ on every row therefore overlap with weight at least 1 on the rows where any two such permutations differ, and so do the weights 1 on S: both families are compatible certificates, of weights $n/2$ and k, and Theorem 2(a) gives $\kappa^*(M) \leq \min\{k, n/2\}$. Each family also consists of fractional certificates, so $\mathrm{FC}(M) \leq \min\{k, n/2\}$. □

Both parts of Theorem 3 condition on part of the noise, and the following lemma transfers the defining inequality through the conditioning.

Lemma D.1 (mixing). *Let $\mathcal{F}$ be a sub-σ-field of the σ-field generated by the Gumbel noise, and write*

$$A_\theta^{\mathcal{F}}(s) := \mathbb{E}\Big[\max_{v \in V}\{\theta_v + s_{M(v)} + \Xi(v)\} \;\Big|\; \mathcal{F}\Big], \qquad p_y^{\mathcal{F}}(\theta, s) := \Pr\big[M(v^*) = y \mid \mathcal{F}\big]$$

for the winner v^* at (θ, s). Fix $\theta \in \mathbb{R}^V$ and $\kappa > 0$. If almost surely

$$A_\theta^{\mathcal{F}}(s+g) - A_\theta^{\mathcal{F}}(s) \;\leq\; \frac{1}{\kappa} \log \sum_{y \in Y} p_y^{\mathcal{F}}(\theta, s)\, e^{\kappa g_y} \qquad \text{for all } s, g \in \mathbb{R}^Y,$$

then (1) holds at κ and θ.

Proof. Take expectations and move them inside the logarithm. The tower property gives $\mathbb{E}[A_\theta^{\mathcal{F}}(s)] = A_\theta(s)$ and $\mathbb{E}[p^{\mathcal{F}}(\theta, s)] = p(\theta, s)$, so the left side averages to $A_\theta(s+g) - A_\theta(s)$. The right side is a concave function of $p^{\mathcal{F}}(\theta, s)$, being the logarithm of a linear function of it, so Jensen's inequality bounds its average by $\kappa^{-1} \log \sum_y p_y(\theta, s) e^{\kappa g_y}$. □

Proof of Theorem 3. Each part conditions on some of the noise and appeals to Lemma D.1; what the conditioning leaves is a softmax in (a) and the projected mechanism in (b).

(a) For each action v choose a pair $(i(v), v_{i(v)})$ that no other action uses; distinct actions have distinct pairs. Condition on every Gumbel except these, so that the total score of v is $\tilde{\theta}_v + s_{M(v)} + \xi_{i(v)}(v_{i(v)})$ with $\tilde{\theta}_v$ measurable. Conditionally the winner is the argmax of $|V|$ scores perturbed by independent standard Gumbels, so its distribution is the softmax of those scores and

$$A_\theta^{\mathcal{F}}(s) \;=\; \log \sum_{v \in V} e^{\tilde{\theta}_v + s_{M(v)}} + \gamma,$$

with γ the Euler constant. Grouping that softmax by fibers gives $p^{\mathcal{F}}$, so both sides of the hypothesis of Lemma D.1 at $\kappa = 1$ equal $\log \sum_y p_y^{\mathcal{F}}(\theta, s) e^{g_y}$, and $\kappa^*(M) \leq 1$. Theorem 1(b) gives $\mathrm{FC}(M) \leq \kappa^*(M)$, and $\mathrm{FC}(M) \geq 1$ for a nonconstant readout.

(b) Condition on the noise outside S. Every action of the slice $\{v : v_S = \tilde{v}\}$ has the output score $s_{\bar{M}(\tilde{v})}$, so the maximum over that slice is

$$\tilde{\theta}_{\tilde{v}} + \sum_{i \in S} \xi_i(\tilde{v}_i) + s_{\bar{M}(\tilde{v})}, \qquad \tilde{\theta}_{\tilde{v}} := \max\Big\{\theta_v + \sum_{i \notin S} \xi_i(v_i) \;:\; v_S = \tilde{v}\Big\},$$

with $\tilde{\theta}_{\tilde{v}}$ measurable and free of s. Conditionally the model is therefore the mechanism on the decision set V_S with readout $\bar{M}$ at base score $\tilde{\theta} \in \mathbb{R}^{V_S}$, which satisfies the hypothesis of Lemma D.1 at $\kappa = \kappa^*(\bar{M})$, because the temperature quantifies over all base scores and is attained for the nonconstant readout $\bar{M}$ (Appendix C); hence $\kappa^*(M) \leq \kappa^*(\bar{M})$. □

We compute the potential range quoted in Section 4 for the one-hot encoding $V = \{e_1, \dots, e_n\}$ with the readout $M(e_j) = j$ at $\theta = 0$. The perturbations are $\Xi(e_j) = \sum_i \xi_i(0) + \delta_j$ with $\delta_j := \xi_j(1) - \xi_j(0)$ independent standard logistic by (A.1), so $A_0(0) = \sum_i \mathbb{E}[\xi_i(0)] + \mathbb{E}\max_j \delta_j$, every fiber maximum has expectation $\sum_i \mathbb{E}[\xi_i(0)]$, and $\Gamma = \mathbb{E}\max_{j\le n} \delta_j$. The logistic distribution function is $\nu \mapsto (1 + e^{-\nu})^{-1}$, and the substitution $\zeta = (1 + e^{-\nu})^{-1}$, $d\zeta = \zeta(1-\zeta)\, d\nu$, gives

$$\mathbb{E}\max_{j\le n} \delta_j \;=\; \int_0^\infty \Big(1 - \frac{1}{(1+e^{-\nu})^n}\Big) d\nu - \int_{-\infty}^0 \frac{d\nu}{(1+e^{-\nu})^n} \;=\; \int_{1/2}^1 \frac{1-\zeta^n}{\zeta(1-\zeta)}\, d\zeta - \int_0^{1/2} \frac{\zeta^{n-1}}{1-\zeta}\, d\zeta.$$

Writing $(1-\zeta^n)/(1-\zeta) = \sum_{j=1}^n \zeta^{j-1}$ in the first integrand and $\zeta^{n-1}/(1-\zeta) = (1-\zeta)^{-1} - \sum_{j=1}^{n-1} \zeta^{j-1}$ in the second, the first integral is $\log 2 + \sum_{j=1}^{n-1}(1 - 2^{-j})/j$ and the second is $\log 2 - \sum_{j=1}^{n-1} 2^{-j}/j$, so

$$\Gamma \;=\; \sum_{j=1}^{n-1} \frac{1-2^{-j}}{j} + \sum_{j=1}^{n-1} \frac{2^{-j}}{j} \;=\; \sum_{j=1}^{n-1} \frac{1}{j}.$$

Proposition 1(a) at $\kappa^*(M) = 1$ and $\alpha = \eta$ gives the regret bound of Section 4. The difference $\sum_{j<n} 1/j - \log n$ is positive at $n = 2$ and increases to Euler's constant $\gamma < 0.58$, since passing from n to $n+1$ adds $1/n - \log(1 + 1/n) > 0$, so that bound exceeds $\log n/\eta$ by less than $0.58/\eta$.