Designing online learning algorithms using Potentials

Yoav Freund

September 3, 2022

Abstract

Abstractly speaking.

1 Introduction

Online prediction with expert advise has been studied extensively over the years and the number of publications in the area is vast (see e.g. [**?**, **?**, **?**, **?**, **?**].

Here we focus on a simple variant of online prediction with expert advice called *the decision-theoretic online learning game* (DTOL) [**?**], we consider the *signed* version of the online game.

DTOL is a repeated zero sum game between a *learner* and an *adversary*. The adversary controls the losses of N experts, or actions, while the learner controls a distribution over the actions. Iteration $t = 1, \ldots, T$ of the game consists of the following steps:

1. The learner chooses a distribution P_j^t over the actions $j \in \{1, \ldots, N\}$.
2. The adversary chooses an *instantaneous loss* for each of the N actions: $l_j^t \in [-1, +1]$ for $j \in \{1, \ldots, N\}$.
3. The learner incurs an *instantanous expected loss* defined as $\ell^t = \sum_{j=1}^N P_j^t l_j^t$

Using standard notation we denote the *cumulative loss* of action j for times $1, \ldots, T$ by $L_j^T = \sum_{t=1}^T l_j^t$. Similarly, we denote the *cumulative loss* of the learner by $L_\ell^T = \sum_{t=1}^T \ell^T$. The *cumulative regret* of the learner with respect to action j is $R_j^T = \sum_{t=1}^T r_j^t = L_\ell^T - L_j^T$. Finally, for any $0 \leq \epsilon \leq 1$ we define the ϵ regret to R_ϵ^T to be value v such that the fraction of values of R_j^T that are smaller than v is at most ϵN.

The goal of the learner is to minimize the maximal regret at time T $R_\epsilon^T \doteq \max_j R_j^T = L_\ell^T - \min_j L_j^T$. The goal of the adversary is to maximize R_ϵ^T.

Our goal is to identify algorithms that improve over known bounds of the form $O(\sqrt{T \ln \frac{1}{\epsilon}})$.

We are interested in bounds the hold simultanously for all values of ϵ. Denote the distribution over regrets at some fixed iteration by μ. We say that the distribution μ satisfies the regret function B if

Definition 1 (Simultanous regret bound) *Let $B : \mathbb{R} \to [0, 1]$ be a non-increasing function which maps regret bounds to probabilities.*

A distribution over regrets μ is simultanously bounded by B if

$$\forall r \in \mathbb{R} \;\; \mathbf{P}_{R\sim\mu}[R \geq r] \leq B(r)$$

Where $B : \mathbb{R} \to [0, 1]$ is a non-increasing function.

Our main tool for designing algorithms and proving bounds is potential functions.

We say that A distribution over regrets μ satisfies the potential function ϕ if

Definition 2 (Average potential bound) *A distribution over he reals μ satisfies the average potential function ϕ if*

$$\mathbf{E}_{R\sim\mu}[\phi(R)] \leq 1$$

Where $\phi : \mathbb{R} \to \mathbb{R}^+$ is a non decreasing function.

Theorem 1 *A distribution μ is simultanously bounded by B if and only if it satisfies the average potential bound with $\phi(R) = B(R)^{-1}$*

Unlike uniform regret bounds, that are rather opaque, one can easily write a single step backward recursion for potentials. In other words, given a potential function at iteration t, one can construct potential functions for iteration $t - 1$. This povides a way for optimizing the potential function.

The original DTOL game does not seem to have min-max strategies. However, in an extended version of the game we can characterize min-max strategies. The extended game provides the adversary with additional possibilities, but not the learner. In other words, the standard game is more restrictive to the adversary, which means that upper bounds on the regret for the extended game also hold for the restricted game.

- min-max for known T
- Almost min/max for unknown T
- Simultanously for all ϵ
- Replacing T with V_T
- Interpretation of the limit game.

Potential Based Algorithms Many of the known algorithms for online learning are based on a *potential function*. Roughly speaking, a potential function $\phi(r,t)$ associates an action whose cumulative regret r on iteration t a non negative potential. The weight that the learner associates with the action is (approximately) the gradient of the potential function.

The standard technique for proving regret bounds is to show an upper bound on the average potential which can then be translated into an upper bound on the regret relative to the best ϵ percentile.

While potential-based online learning algorithms have a long history, other approaches such as *follow the leader* and it's variants provide similar bounds and are often much more efficient.

Surprisingly, we prove that there is a one-to-one correspondence between a simultanuous (deterministic) regret guarantees and a potential function that corresponds to the reciprocal of the bound. We thus focus on potential based learning algorithms.

Given a potential function for iteration t we can naturally calculate potential functions for iterations $t-1, t-2, \ldots$. Using this backwards recursion we can solve games with a known horizon.

2 Extending the game

Our goal is to describe a potential-based algorithm which is min-max optimal. To achieve this we give the adversary more options than it has in the standard game.

Specifically, we arbitrary division and choosing the step size.

3 The algorithm

We now describe the learning algorithm for the extended game.

We define the tail probability of a normal distribution with mean zero and variance t to be

$$\epsilon(R,t) = \sqrt{2\pi t} \int_R^\infty e^{-\frac{\rho^2}{2t}} d\rho \tag{1}$$

As the tail probability is monotonically decreasing from 1 to 0, the inverse function $R(\epsilon,t)$ is well defined so that $\epsilon(R(\epsilon_0,t),t) = \epsilon_0$ and $R(\epsilon(R_0,t),t) = R_0$

We define a potential that is the reciprocal of the tail:

$$\phi(R,t) = \frac{1}{\epsilon(R,t)}$$

On iteration t our algorithm assigns to expert j, whose regret on iteration t is R_j^t the weight

$$w_j^t = \phi(R_j + 2, t+1) - \phi(R_j - 2, t+1)$$

And the normalized weights are $P_j^t = w_j^t / Z_t$ where $Z_t = \sum_{j=1}^N w_j^t$

Our main result is

Theorem 2 *There exists a parameter-free DTOL algorithm and a constant $c > 0$ such that for any N loss sequences of length t, for any $1 > \epsilon > 0$ we have that*

$$\#\left\{1 \leq i \leq N \,|R_i \geq R(\epsilon, t + c\log t)\right\} \leq \epsilon N$$

4 Discrete time game

Our analysis is based on defining an expanded version of DTOL game which allows for a tighter analysis. The extended game provides more options to the adversary, but not to the learner. In particular, it allows the adversary to play the standard DTOL game. As a consequence, regret bounds that hold for the extended game also hold for standard DTOL in which the adversary is more restricted.

In the extended game we game we use separate between the notion of iteration $i = 1, 2, 3, \ldots$ and the notion of time $0 = t_0 \leq t_1 \leq t_2 \leq \cdots$. The gap $t_{i+1} - t_i \leq 1$ and therefor $t_i \leq i$. Our regret bounds will involve t_i instead of i and will be at most as large as the bounds on standard DTOL.

The first extension of DTOL is to allow the adversary to choose the range of the losses at each iteration. In standard DTOL this range is fixed to $[-1, +1]$. In the extended game we allow the adversary to choose the range $[-s_i, +s_i]$ for some $s_i \in [0, 1]$. It is not hard to see that if this was the only change then the adversary would always choose $s_i = 1$. To make the choice $s_i < 1$ viable

To make it worthwhile for the adversary to choose $s_i < 1$ we replace the "number of iterations" in our regret bounds with a notion of "time" that is defined as follows. Let t_i be a real number that defines the time at iteration i, $t_0 = 0$ and $t_{i+1} = t_i + s_i^2$ we assume there exists a finite n such that $t_n = T$.

We need to put a bound on $B(t_i, R)$ because without a bound the bias can be s_i and the cumulative bias after unit time would be $s_i/s_i^2 = 1/s_i$ which goes to infinity as $s_i \to 0$. Infinite biases make that total loss undefined.

We will later give some particular potential functions for which no a-priori knowledge of the termination condition is needed. The associated bounds will hold for any iteration of the game.

On iteration $i = 1, 2, \ldots$

1. If $t_{i-1} = T$ the game terminates.
2. The adversary chooses a *step size* $0 < s_i \leq 1$, which advances time by $t_i = t_{i-1} + s_i^2$.
3. Given s_i, the learner chooses a distribution $P(i)$ over $\mathbb{R}$.
4. The adversary chooses a mapping from $\mathbb{R}$ to distributions over $[-s_i, +s_i]$: $Q : \mathbb{R}^2 \to \Delta^{[-s_i,+s_i]}$
   such that $B(t_i, R) \doteq \mathbf{E}_{y\sim Q(t_i,R)}\left[y\right] \leq cs_i^2$ for all $R \in \mathbb{R}$.

5. The aggregate loss is calculated:

$$\ell(t_i) = \mathbf{E}_{R\sim\mathbf{\Psi}(t_i)}\left[P(t_i, R)B(t_i, R)\right] \tag{2}$$

6. The state is updated. The expectation below is over distributions. and the notation $G \oplus R$ means that distribution G over the reals is shifted by the amount defined by the scalar R:

$$\mathbf{\Psi}(t_i) = \mathbf{E}_{R\sim\mathbf{\Psi}(t_{i-1})}\left[Q(t_i)(R) \oplus (R - \ell(t_i))\right]$$

When $t_i = T$ the game is terminated, and the final value is calculated:

$$\Phi(T) = \mathbf{E}_{R\sim\mathbf{\Psi}(T)}\left[\phi(T, R)\right]$$

5 Equivalence between simultanous regret bounds and potentials

A standard technique in the analysis online learning is to prove an upper bound on the average potential and use that upper bound to prove an upper bound on the regret. Here we show a sense in which the converse is also true: if we have a *unifom* upper bound on the regret, then there is an upper bound on the average potential.

We fix the time step and consider an arbitrary probability measure μ over the real line that defines the regret distribution. We give two characterizations for μ, one bounding the regret, the other bounding the potential. We then show a one to one relationship between the two characterizations.

Definition 3 (Simultanous regret bound) *A distribution over the reals μ is simultanously bounded by B if*

$$\forall a \in \mathbb{R} \;\; \mathbf{P}_{R\sim\mu}\left[R \geq a\right] \leq B(a)$$

Where $B : \mathbb{R} \to [0, 1]$ is a non-increasing function.

Next we define the average potential bound.

Definition 4 (Average potential bound) *A distribution over he reals μ satisfies the average potential function ϕ if*

$$\mathbf{E}_{R\sim\mu}\left[\phi(R)\right] \leq 1$$

Where $\phi : \mathbb{R} \to \mathbb{R}^+$ is a non decreasing function.

We now show a one-to-one relationship between these two types of bounds

Theorem 3 *A distribution μ is simultanously bounded by B if and only if it satisfies the average potential bound with $\phi(R) = B(R)^{-1}$*

- **μ satisfies a simultanous bound for B if it satisfies an average potential bound for $\phi = B^{-1}$**
  Assume by contradiction that μ does not satisfy the simultanous bound. In other words there exists $a \in \mathbb{R}$ such that $\mathbf{P}_{R\sim\mu}[R > a] > B(a)$. From Markov inequality and the fact that ϕ is non decreasing we get

$$\mathbf{E}_{R\sim\mu}[\phi(R)] \geq \phi(a)\mathbf{P}_{R\sim\mu}[R > a] > \phi(a)B(a) = \frac{B(a)}{B(a)} = 1$$

  but $\mathbf{E}_{R\sim\mu}[\phi(R)] > 1$ contradicts the average potential assumption for the potential $\phi(R) = B(R)^{-1}$

- **μ satisfies an average potential bound for $\phi = B^{-1}$ if it satisfies a simultanous bound for B**
  As ϕ is a non-decreasing function, and assuming R, R' are drawn independently at random according to μ:

$$\begin{aligned}
\mathbf{E}_{R\sim\mu}[\phi(R)] &= \mathbf{E}_{R\sim\mu}[\phi(R)\mathbf{P}_{R'\sim\mu}[\phi(R') \geq \phi(R)]] && (3)\\
&\leq \mathbf{E}_{R\sim\mu}[\phi(R)\mathbf{P}_{R'\sim\mu}[R' \geq R]] && (4)\\
&< \mathbf{E}_{R\sim\mu}[\phi(R)B(R)] && (5)\\
&= \mathbf{E}_{R\sim\mu}\left[\frac{B(R)}{B(R)}\right] = \mathbf{E}_{R\sim\mu}[1] = 1 && (6)
\end{aligned}$$

6 Multiply convex potential functions

7 Backwards recursion of potentials functions

7.1 the rest

We can now state the main claim of this paper.

Online learning can be described as and iterative game []

We define the *state* of the online prediction game as

We define a potential function as ...

The tail bound can be written as a potential function...

We can work our way backwards.

A common approach to designing online learning algorithms is to define a *potential function*. The potential function $\phi : \mathbb{R} \to \mathbb{R}$ is a positive, continuous and non-decreasing function of the regret. One popular potential function is the exponential function which has the form $\phi(R) = \exp(\eta R)$ where $\eta > 0$ is the *learning rate*. See **[?]** for an extensive review of the use of potential functions for online learning.

A central quantity in the design and analysis of potential based algorithms is the *average potential* which we refer to here as the *score*. The score at time k is defined as:

$$\Phi^k = \frac{1}{N}\sum_{j=1}^{N}\phi(R_j^k) \tag{7}$$

The fact that $\phi(R)$ is positive and non-decreasing implies an upper bound on the regret w.r.t. any expert. Suppose that the score at time t is upper bounded by A and that the learner suffers regret B with respect to at least one expert: $B = \max_j(R_j^T)$, then $\phi(B) \leq N\Phi_T \leq NA$.

In most formulations of this technique, the potential function ϕ is fixed a-priori, and the learner's strategy is designed based on this function. This raises a natural question:

Question 1 *Is there an optimal potential function ϕ for DTOL?*

Without further constraints, this question is ill-defined. We take several steps to better define the question.

1. We allow the potential function to depend on the iteration number i, i.e. we study potential functions of the form $\phi(i, R)$. [1]

2. We fix the number of iterations T. This assumption will later be removed.

3. We fix the final potential $\phi(T, \cdot)$. This function is required to have defined and strictly positive derivatives of degrees zero to four.[2] This assumption is satisfied by most commonly used potential functions, including the exponential potential, the NormalHedge potential and polynomial potentials of degree higher than four.

This leads to a more specific question:

Question 2 *Given the length of the game T and the final potential function $\phi(T, R)$, can we define the best potentials $\phi(i, R)$ for $0 \leq i < T$*

We give a qualified answer to this question. Rather than finding a *single* potential function, we associate an *upper potential* with learner strategy P and a *lower potential* with adversarial strategy Q.

We start by fixing both Q and P. As we show in Definition 6 and Lemma 6, for any pair of strategies Q and P and for any iteration $0 \leq i \leq T$ there exists a potential function $\phi_{Q,P}(i, \cdot)$ such that if we define the score $\Phi_{Q,P}(i, \mathbf{R}^i)$ to be

$$\Phi_{Q,P}(i, \mathbf{R}^i) = \frac{1}{N}\sum_{j=1}^{N}\phi_{Q,P}\left(T-1, R_j^i\right) \tag{8}$$

then $\Phi_{Q,P}(0, \mathbf{R}^0) = \cdots = \Phi_{Q,P}(T, \mathbf{R}^T)$.

[1] Potentials that depend on time were used for NormalHedge and related algorithms [**?**, **?**]

[2] The zeroth derivative is the function itself. Therefor positivity of derivatives 0 and 1 implies the standard assumption that $\phi(T, R)$ is positive and increasing.

Using Equation 15 we define an *upper potential* and a *lower potential* for each iteration $0 \leq i \leq T$

$$\phi_P^{\downarrow}(i, R^i) = \max_Q \phi_{Q,P}(i, R^i) \text{ and } \phi_Q^{\uparrow}(i, R^i) = \min_P \phi_{Q,P}(i, R^i) \tag{9}$$

Which are given given inductive definitions in Equations (16,17)

Clearly, for any Q, P, i, R we have that $\phi_P^{\uparrow}(i, R) \leq \phi_Q^{\downarrow}(i, R)$. Our goal is to find a pair of strategies Q, P such that $\phi_P^{\uparrow}(i, R) = \phi_Q^{\downarrow}(i, R)$. That would imply that Q and P are min-max optimal strategies.

It is unclear whether min-max strategies exist for standard DTOL. In order to close the apparent gap between the upper and lower potentials and identify the min-max strategies we *expand* the game. Here expansion means giving the *adversary* more choices while keeping the learner's choices unchanged. As a consequence, any upper bounds guarantees that hold for the expanded game also hold for the standard DTOL (but the lower bounds might not).

We make two expansion steps. In the first, we allow the adversary to arbitrarily divide experts. In the second we allow the adversary to choose the range of allowed instantanous losses. Using both expansions we identify the min/max strategies and the min/max potential function thereby answering question 2.

The first expansion gives the adversary the ability to arbitrarily split experts. Intuitively, rather than associating a single loss with each expert, The adversary can arbitrarily divide the the expert into sub-experts and assign a different loss to each. This is similar to [**?**] where the number of experts is not known in advance

We define the *even split* adversarial strategy $Q^{1/2}$ as one which, at each iteration splits *each current expert* into two equal parts, one incurs loss 1 and the other loss -1. The result is that at iteration n we have 2^n experts, each of weight 2^{-n}, each corresponding to a loss sequence in $\{-1, +1\}^n$. The cumulative loss of the experts has a binomial distribution $\mathbb{B}(n, 1)$ [3]. The symmetry of $Q^{1/2}$ implies that the loss of the learner against this adversary is zero independently of the learner choices. Therefor the cumulative regret is also Binomially distributed according to $\mathbb{B}(n, 1)$. The final score is equal to the expected value of the final potential $\mathbf{E}_{R\sim\mathbb{B}(T,1)}[\phi(T, R)]$. As we will show there is a simple strategy for the learner that guarantees an upper bound of $\mathbf{E}_{R\sim\mathbb{B}(T,2)}[\phi(T, R)]$. The upper and lower bounds are both based on binomial distribution. However, as the step sizes are $1 < 2$, these strategies are not a min/max pair.

Intuitively, in order to achieve min/max optimality we need to make the step size of the adversary equal the step size of the learner. We next explain how the step size for the learner is reduced to the step size of the adversary.

We achieve that by taking the second expansion step. In this expansion we allow the the adversary to set the step ranage in iteration i to $[-s_i, +s_i]$ for any $0 \leq s_i \leq 1$. In Section 4 we give a full description of the game. For now, suppose that the adversary chooses $s_i = 1/m$ for some large natural number m and for all i. Scaling $Q^{1/2}$ to be $\pm 1/m$ with equal probabilities we find that the lower score is equal to $\mathbf{E}_{R\sim\mathbb{B}(T,1/m)}[\phi(T, R)]$. It is not hard to see that if $mto\infty$ while T remains constant then the binomial distribution degenerates to the delta function. In order to avoid

[3] We use $\mathbb{B}(n, s)$ to denote the binomial distribution that assigns probability $\binom{n}{j}2^{-n}$ to the point $(n - 2j)s$ for $j = 0, \ldots, n$

this degeneracy we set $T = m^2$. The binomial distribution $\mathbb{B}(m^2, 1/m)$ has variance one for all m and converges to the standard normal as $m \to \infty$. In Section 4 we show that in this limit the difference between the upper and lower potentials converges to zero.

To show that this limit is the min/max, we need to show that, in fact, the optimal choice for the adversary is to set m arbitrarily high. If the adversary prefers a finite m, then this is not the min/max. This is where the strictly positive fourth derivative is important. In Section **??** we show that, if the final potential function is in $\mathcal{P}^4$ then the lower potential strictly increases with m. Establishing the min/max properties of the limit $m \to \infty$.

One benefit of working with the weiner process is that the backward equations that define the potential transform, in the limit, into the Kolmogorov Backwards equations (KBE) The potential function $\phi^{\uparrow}(t, R) = \phi^{\downarrow}(t, R) = \phi(t, R)$ is the solution of the Kolmogorov backwards equation (KBE) with the boundary condition defined by $\phi(T, R)$. We have thus answered Question 2 in the affirmative.

Having identified the min/max strategies for a given final potential function $\phi(T, R)$ we next turn to choosing $\phi(T, R)$. We say that a parametrized potential function $\phi(\vec{\theta}_T, R)$ is *compatible* with KBE if the solution to KBE with boundary condition $\phi(T, R) = \phi(\vec{\theta}_T, R)$ can be written in the form $\phi(t, R) = \phi(\vec{\theta}_t, R)$. In Section 11 we identify two KBE compatible functions, one corresponds to the exponential potential, the other to the NormalHedge potential. An additional benefit of characterizing the solution as a solution of KBE is that the solution can be extended to $t > T$. This removes the need to know T a-priori and yields an *any time* algorithm.

Finally, we come back to Question 1. Intuitively, we want $\phi(T, R)$ (properly normalized) to increase as fast as possible, without causiing the score to increase with t. As we have identified the worst case adversary to be Brownian motion, we can ask what is the fastest increasing potential for which Brownian motion will not increase the score. In section 12 we show that an appropriate limit of NormalHedge has this distinction.

8 Preliminaries

The integer time game takes place on the set $(i, R) \in \{0, 1, \ldots, T\} \times \mathbb{R}$, where i corresponds to the iteration number , and R corresponds to the (cumulative) regret.

As in the standard DTOL setting, an expert corresponds to a sequence of cumulative regrets. However, unlike the standard DTOL the number of experts is allowed to be infinite. In DTOL the *state* of the learning process is defined by the by the regret of each of the N experts: $\langle R_1(i), \ldots, R_N(i) \rangle$. In order to represent the regret of a potentially uncountable set of experts we define the state as a distribution over possible regret values. Thus the *state* of the game on iteration i is a distribution (probability measure) over the real line, denoted $\mathbf{\Psi}(i)$. We denote by $R \sim \mathbf{\Psi}(i)$ a random regret R that is chosen according to the distribution corresponding to the state at iteration i. Given a measureable function $f : \mathbb{R} \to \mathbb{R}$ we define $\mathbf{D}_{R \sim \mathbf{\Psi}(i)}[f(R)]$ to be the distribution of $f(R)$ when $R \sim \mathbf{\Psi}(i)$. We similarly define the expectd value of $f(R)$ by $\mathbf{E}_{R \sim \mathbf{\Psi}(i)}[f(R)]$, and the probability of an event $\mathbf{e}$ defined using $f(R)$ by $\mathbf{P}_{R \sim \mathbf{\Psi}(i)}[\mathbf{e}(f(R))]$.

The initial state $\mathbf{\Psi}(0)$ is a point mass at $R = 0$. The state $\mathbf{\Psi}(t)$ is defined by $\mathbf{\Psi}(t_{i-1})$ and the choices made by the two players as described in the next section.

The *final potential function* $\phi(T, \cdot)$ is predefined and known to both sides. The final *score* is defined to be

$$\Phi(T) = \mathbf{E}_{R \sim \mathbf{\Psi}(T)} [\phi(T, R)] \tag{10}$$

The goal of the learner is to minimize $\Phi(T)$ and the goal of the adversary is to maximize it.

We assume that the final poptential is *strictly positive of degree* k, which is defined as follows:

Definition 5 (Strict Positivity of degree k) *A function $f : \mathbb{R} \to \mathbb{R}$ is strictly positive of degree k, denoted $f \in \mathcal{P}^k$ if the derivatives of orders 0 to k: $f(x), \frac{d}{dx} f(x), \ldots, \frac{d^k}{dx^k} f(x)$ exist and are strictly positive.*

A simple lemma connects an upper bound on any score function in $\mathcal{P}^1$ with a bound on the regret.

Lemma 4 *Let $\phi(T, R) \in \mathcal{P}^1$, $\Phi(T) \leq U$ and $\epsilon = \mathbf{P}_{R \sim \mathbf{\Psi} t} [R > R']$ be the probability of the set of experts with respect to which the regret is larger than R' Then*

then the regret of the algorithm relative to the top ϵ of the experts is upper bounded by

$$\phi(T, R) \leq U/\epsilon$$

9 Integer time game

We start with a setup in which time and iteration number are the same, i.e. $t_i = i$. In this section we suppress t_i and instead use the iteration number $i = 0, 1, 2, \ldots, T$.

We define the *state of the game* on iteration i: $\mathbf{\Psi}(i)$ as the distribution of regret over the experts. Note that experts is allowed to be be un-countably infinite. In particular the adversary can assign to the experts with regret x at iteration t an arbitrary distribution of losses in the range $[-1, +1]$.

The game is defined by three parameters:

- T : The number of iterations
- $\mathbf{\Psi}(0) = \delta(0)$ is the initial state of the game which is a point mass distribution at 0.
- $\phi(T, R)$: The function that is in $\mathcal{P}^2$.

The transition from $\mathbf{\Psi}(i)$ to $\mathbf{\Psi}(i + 1)$ is defined by the choices made by the adversary and the learner.

1. The learner chooses weights. Formally, $P(i, \cdot)$ is a density over $\mathbb{R}$:
2. The adversary chooses the losses of the experts. Formally this is a mapping from $\mathbb{R}$ to distributions over $[-1, +1]$: $Q(i) : \mathbb{R} \to \Delta^{[-1,+1]}$. We use $l \sim Q(i, R)$ to denote the distribution over the instantaneous loss associated with iteration i and regret R.

3. The aggregate loss (also called "the loss of the master") is calculated:

$$\ell_P(i) = \mathbf{E}_{R\sim\mathbf{\Psi}(i)}\left[P(i,R)\mathbf{E}_{l\sim Q(i,R)}\left[l\right]\right] \tag{11}$$

We define the *bias* at (i,R) to be $B(i,R) \doteq \mathbf{E}_{l\sim Q(i,R)}\left[l\right]$ which allows us to rewrite Eqn (11) as

$$\ell_P(i) = \mathbf{E}_{R\sim\mathbf{\Psi}(i)}\left[P(i,R)B(i,R)\right] \tag{12}$$

note that $B(i,R)$ is in $[-1,1]$ and that $\ell_P(i)$ is the mean of $P(i,R)B(i,\cdot)$. Note also that $-2 \leq y - \ell_P(i) \leq 2$ corresponds to the instantaneous regret.

4. The state is updated.

$$\mathbf{\Psi}(i+1) = \mathbf{D}_{R\sim\mathbf{\Psi}(i),l\sim Q(i,R)}\left[R + l - \ell_P(i)\right] \tag{13}$$

Which is the distribution of $R + l - \ell_P(i)$ where R is the cumulative regret at iteration i, whose distribution is to $\mathbf{\Psi}(i)$, l is the instantanous loss chosen according to the adversarial distribution $Q(i,R)$ and $\ell_P(i)$ is the average loss as defined above.

The final score is the mean of the potential according to the final state, as given in Equation (10). The goal of the learner is to minimize the final score and the goal of the adversary is to maximize it. Equations (18,20) define simple strategies for the adversary and the learner. For these strategies we prove our main result regarding the integer game.

Theorem 5 *There exists a strategy for the adversary such that for any strategy of the learner,*

$$\Phi(T) \geq \mathbf{E}_{R\sim\mathbb{B}(T,1)}\left[\phi(T,R)\right]$$

There exists a strategy for the learner such that for any strategy of the adversary,

$$\Phi(T) \leq \mathbf{E}_{R\sim\mathbb{B}(T,2)}\left[\phi(T,R)\right]$$

$\mathbb{B}(n,s)$ is defined in Footnote 3, and the proof is given in Appendix A. The potential This proof is based on the concept of upper and lower potentials which is our main tool in this paper and is explained in the following section.

10 Potentials

The definition of an integer time game specifies the The final potential $\phi(T,R)$ is set in the definition of a game. In this section we show a natural way to extend the definition to all game iterations.

Definition 6 (potential backwards recursion) *Let T be the number of iterations, $\phi(T,R)$ be the final potential, P be a learner strategy and Q be an adversarial strategy. We define the* intermediate potential functions *of $i = T-1, T-2, \ldots, 0$ using the following backwards recursion:* [4]

$$\phi_{P,Q}(i,R) = \mathbf{E}_{l\sim Q(i,R)}\left[\phi_{P,Q}(i+1, R+l-\ell_P(i))\right] \tag{14}$$

Where $\ell_P(i)$ is defined in Eqn (11).

[4] To bootstrap the recursion we set $\phi_{P,Q}(T,R) = \phi(T,R)$

The following lemma guarantees that, for this definition of the potential function, the score function does not change from iteration to iteration.

Lemma 6 *Define the score at iteration i to be*

$$\Phi_{P,Q}(i) = \mathbf{E}_{R\sim\mathbf{\Psi}(i)}\left[\phi_{P,Q}\left(i+1,R\right)\right] \tag{15}$$

then

$$\phi_{P,Q}(0,0) = \Phi_{P,Q}(0) = \Phi_{P,Q}(1) = \cdots = \Phi_{P,Q}(T)$$

The proof is given in Appendix B

Definition 6 and Lemma 6 correspond to a fixed pair of strategies. Using those, there is a natural way to define an *upper potential* $\phi_P^{\downarrow}$ we characterizes the least upper bound on the potential that is guaranteed by the learner strategy P.

$$\phi_P^{\downarrow}(i,R) = \begin{cases} \phi(T,R) & \text{if } i = T \\ \sup_Q \mathbf{E}_{l\sim Q(i,R)}\left[\phi_P^{\downarrow}(i+1,R+l-\ell_P(i))\right] & \text{if } 0 \le i < T \end{cases} \tag{16}$$

Similarly, the lower potential $\phi_Q^{\uparrow}$ charaecterizes the highest lower bound on the potentail guaranteed bby the adversarial strategy Q.

$$\phi_Q^{\uparrow}(i,R) = \begin{cases} \phi(T,R) & \text{if } i = T \\ \inf_P \mathbf{E}_{l\sim Q(i,R)}\left[\phi_Q^{\uparrow}(i+1,R+l-\ell_P(i))\right] & \text{if } 0 \le i < T \end{cases} \tag{17}$$

10.1 Strategies for the integer time game

We assign the adversary the *even split* strategy, described in the introduction. Formally, even split is defined as

$$Q^{1/2}(i,R) = \begin{cases} -1 & \text{w.p. } 1/2 \\ +1 & \text{w.p. } 1/2 \end{cases} \tag{18}$$

It is easy to see that, when $Q^{1/2}$ is used, the expected loss $\ell_P = 0$ regardless of P. In other words, the learner has no influence on the lower potential which is simply:

$$\phi_{Q^{1/2}}^{\uparrow}(i-1,R) = \frac{\phi_{Q^{1/2}}^{\uparrow}(i,R+1) + \phi_{Q^{1/2}}^{\uparrow}(i,R-1)}{2} \tag{19}$$

The learner strategy:

$$P^d(i-1,R) = \frac{1}{Z}\frac{\phi(i,R+2) - \phi(i,R-2)}{2} \tag{20}$$

Where Z is a normalization factor

$$Z = \mathbf{E}_{R\sim\mathbf{\Psi}(i)}\left[\frac{\phi(i,R+2) - \phi(i,R-2)}{2}\right]$$

guarantees the upper potential

$$\phi^{\downarrow}_{Pd}(i-1,R) = \frac{\phi^{\downarrow}_{Pd}(i,R+2) + \phi^{\downarrow}_{Pd}(i,R-2)}{2} \tag{21}$$

These strategies satisfy Theorem 5, the proof is given in Appendix A.

We find that the lower bound corresponds to an unbiased random walk with step size ± 1. The upper bound also corresponds to a an unbiased random walk with step size $\pm(1+c)$. The natural setting in the natural game is $c = 1$, which means that there is a significant difference between the upper and lower bounds. As we show in the next section, this gap converges to zero in the continuous time setting, and the upper and lower bounds match, making the strategies for both sides min-max optimal.

Note also that the adversarial strategy the aggregate loss $\ell(t)$ is always zero, regardless of the strategy of the learner, and state progression is independent of the learner's choices.

In the discrete time game the adversary has an additional choice, the choice of s_i. Thus the adversary's strategy includes that choice. There are two constraints on this choice: $s_i \geq 0$ and $\sum_{i=1}^{n} s_i^2 = T$. Note that even that by setting s_i arbitrarily small, the adversary can make the number of steps - n - arbitrarily large. We will therefor not identify a single adversarial strategy but instead consider the supremum over an infinite sequence of strategies.

We use $N(0,\sigma)$ to denote the normal distribution with mean 0 and std σ.

Theorem 7

let $A = \mathbf{E}_{R\sim N(0,\sqrt{T})}\left[\phi(T,R)\right]$

- *For any* $\epsilon > 0$ *there exists a strategy for the adversary such that for any strategy of the learner* $\Phi(T) \geq A - \epsilon$
- *There exists a strategy for the learner that guarantees, against any adversary* $\Phi(T) \leq A$.

10.2 The adversary prefers smaller steps

As noted before, if the adversary chooses $s_i = 1$ for all i the game reduces the the integer time game. The question is whether the adversary would prefer to stick with $s_i = 1$ or instead prefer to use $s_i < 1$. In this section we give a surprising answer to this question – the adversary always prefers a smaller value of s_i to a larger one. This leads to a preference for $s_i \to 0$, as it turns out, this limit is well defined and corresponds to Brownian motion, also known as Wiener process.

Consider a sequence of adversarial strategies S_k indexed by $k = 0, 1, 2,$. The adversarial strategy S_k is corresponds to always choosing $s_i = 2^{-k}$, and repeating $Q^{1/2}_{\pm 2^{-k}}$ for $T2^{2k}$ iterations. This corresponds to the distribution created by a random walk with $T2^{2k}$ time steps, each step equal to $+2^{-k}$ or -2^{-k} with probabilities $1/2, 1/2$. Note that in order to preserve the variance, halving the step size requires increasing the number of iterations by a factor of four.

Let $\phi(S_k, t, R)$ be the value associated with adversarial strategy S_k, time t (divisible by 2^{-2k}) and location R. We are ready to state our main theorem.

Theorem 8 *If the final value function has a strictly positive fourth derivative:*

$$\frac{d^4}{dR^4}\phi(T,R) > 0, \forall R$$

then for any integer $k > 0$ and any $0 \leq t \leq T$, such that t is divisible by 2^{-2k} and any R,

$$\phi(S_{k+1}, t, R)) > \phi(S_k, t, R)$$

Proofs for Theorem 8 and Lemma 9 are given in Appendix C Before proving the theorem, we describe it's consequence for the online learning problem. We can restrict Theorem 8 for the case $t = 0$,$R = 0$ in which case we get an increasing sequence:

$$\phi(S_1, 0, 0) < \phi(S_2, 0, 0) < \cdots < \phi(S_k, 0, 0) <$$

The limit of the strategies S_k as $k \to \infty$ is the well studied Brownian or Wiener process. The backwards recursion that defines the value function is the celebrated Backwards Kolmogorov Equation with zero drift and unit variance

$$\frac{\partial}{\partial t}\phi(t,R) + \frac{1}{2}\frac{\partial^2}{\partial R^2}\phi(t,R) = 0 \tag{22}$$

Given a final value function with a strictly positive fourth derivative we can use Equation (22) to compute the value function for all $0 \leq t \leq T$. We will do so in he next section.

We now go back to proving Theorem 8. The core of the proof is a lemma which compares, essentially, the value recursion when taking one step of size 1 to four steps of size 1/2.

Consider the adversarial strategies S_k and S_{k+1} at a particular time point $0 \leq t \leq T$ such that t is divisible by $\Delta t = 2^{-2k}$ and at a particular location R. Let $t' = t + \Delta t$, and fix a value function for time , $\phi(t', R)$ and compare between two values at R, t. The first value denoted $\phi_k(t, R)$ corresponds to S_k, and consists of a single random step of $\pm 2^{-k}$. The other value $\phi_{k+1}(t, R)$ corresponds to S_{k+1} and consists of four random steps of size $\pm 1/2$.

Lemma 9 *If $\phi(t', R)$ is, as a function of R continuous, strictly convex and with a strictly positive fourth derivative. Then*

- $\phi_k(t, R) < \phi_{k+1}(t, R)$
- *Both $\phi_k(t, R)$ and $\phi_{k+1}(t, R)$ are continuous, strictly convex and with a strictly positive fourth derivative.*

10.3 Strategies for the Learner in the discrete time game

The strategies we propose for the learner in the continuous time game are an adaptation of the strategies P^1, P^2 to the case where $s_i < 1$.

We start with the high-level idea. Consider iteration i of the continuous time game. We know that the adversary prefers s_i to be as small as possible. On the other hand, the adversary has to

choose some $s_i > 0$. This means that the adversary always plays sub-optimally. Based on s_i the learner makes a choice and the adversary makes a choice. As a result the current state $\mathbf{\Psi}(t_{i-1})$ is transformed to $\mathbf{\Psi}(t_i)$. To choose it's strategy, the learner needs to assign value possible states $\mathbf{\Psi}(t_i)$. How can she do that? By assuming that in the future the adversary will play optimally, i.e. setting s_i arbitrarily small. While the adversary cannot be optimal, it can get arbitrarily close to optimal, which is Brownian motion.

Solving the backwards Kolmogorov equation with the boundary condition $\phi(T, R)$ yields $\phi(t, R)$ for any $R \in \mathbb{R}$ and $t \in [0, T]$. We now explain how using this potential function we derive strategies for the the learner.

Note that the learner chooses a distribution *after* the adversary set the value of s_i. The discrete time version of P^1 (Eqn 20) is

$$P^{1d}(t_{i-1}, R) = \frac{1}{Z^{1d}} \frac{\phi(t_i, R + s_{i-1} + cs_{i-1}^2) - \phi(t_i, R - s_{i-1} - cs_{i-1}^2)}{2} \tag{23}$$
$$\text{where } Z^{1d} = \mathbf{E}_{R \sim \mathbf{\Psi}(t_i)} \left[\frac{\phi(t_i, R + s_{i-1} + cs_{i-1}^2) - \phi(t_i, R - s_{i-1} - cs_{i-1}^2)}{2} \right]$$

Next, we consider the discrete time version of P^2: (Eqn **??**)

$$P^{2d}(t_{i-1}, R) = \frac{1}{Z^{2d}} \left. \frac{\partial}{\partial r} \right|_{r=R} \phi(t_{i-1} + s_{i-1}^2, r) \tag{24}$$
$$\text{where } Z^{2d} = \mathbf{E}_{R \sim \mathbf{\Psi}(t_i)} \left[\left. \frac{\partial}{\partial r} \right|_{r=R} \phi(t_{i-1} + s_{i-1}^2, r) \right]$$

11 Two KBE compatible potential functions

The potential functions, $\phi(t, R)$ is a solution of PDE (22):

$$\frac{\partial}{\partial t}\phi(t, R) + \frac{1}{2}\frac{\partial^2}{\partial r^2}\phi(t, R) = 0 \tag{25}$$

under a boundary condition $\phi(T, R) = \phi(T, R)$, which we assume is in $\mathcal{P}^4$

So far, we assumed that the game horizon T is known in advance. We now show two value functions where knowledge of the horizon is not required. Specifically, we call a value function $\phi(t, R)$ *self consistent* if it is defined for all $t > 0$ and if for any $0 < t < T$, setting $\phi(T, R)$ as the final potential and solving for the Kolmogorov Backward Equation yields $\phi(t, R)$ regardless of the time horizon T.

We consider two solutions to the PDE, the exponential potential and the NormalHedge potential. We give the form of the potential function that satisfies Kolmogorov Equation 22, and derive the regret bound corresponding to it.

The exponential potential function which corresponds to exponential weights algorithm corresponds to the following equation

$$\phi_{\text{exp}}(R, t) = e^{\sqrt{2}\eta R - \eta^2 t}$$

Where $\eta > 0$ is the learning rate parameter.

Given ϵ we choose $\eta = \sqrt{\frac{\ln(1/\epsilon)}{t}}$ we get the regret bound that holds for any $t > 0$

$$R_\epsilon \leq \sqrt{2t \ln \frac{1}{\epsilon}} \tag{26}$$

Note that the algorithm depends on the choice of ϵ, in other words, the bound does *not* hold for all values of ϵ at the same time.

The NormalHedge value is

$$\phi_{\mathrm{NH}}(R,t) = \begin{cases} \frac{1}{\sqrt{t+\nu}} \exp\left(\frac{R^2}{2(t+\nu)}\right) & \text{if } R \geq 0 \\ \frac{1}{\sqrt{t+\nu}} & \text{if } R < 0 \end{cases} \tag{27}$$

Where $\nu > 0$ is a small constant. The function $\phi_{\mathrm{NH}}(R,t)$, restricted to $R \geq 0$ is in $\mathcal{P}^4$ and is a constant for $R \leq 0$.

The regret bound we get is:

$$R_\epsilon \leq \sqrt{(t+\nu)\left(\ln(t+\nu) + 2\ln\frac{1}{\epsilon}\right)} \tag{28}$$

This bound is slightly larger than the bound for exponential weights, however, the NormalHedge bound holds simultaneously for all $\epsilon > 0$ and the algorithm requires no tuning.

12 NormalHedge yields the fastest increasing potential

Up to this point, we considered any continuous value function with strictly positive derivatives 1-4. We characterized the min-max strategies for any such function. It is time to ask whether value functions can be compared and whether there is a "best" value function. In this section we give an informal argument that NormalHedge is the best function. We hope this argument can be formalized.

We make two observations. First, the min-max strategy for the adversary does not depend on the potential function! (as long as it has strictly positive derivatives). That strategy corresponds to the Brownian process.

Second, the argument used to show that the regret relative to ϵ-expert of the expert is based on two arguments

- The average value function does not increase with time.
- The (final) value function increases rapidly as a function of R

The first item is true by construction. The second argument suggests the following partial order on value functions. Let $\phi_1(t,R), \phi_2(t,R)$ be two value functions such that

$$\lim_{R\to\infty} \frac{\phi_1(t,R)}{\phi_2(t,R)} = \infty$$

then ϕ_1 *dominates* ϕ_2, which we denote by, $\phi_1 > \phi_2$.

On the other hand, if the value function increases too quickly, then, when playing against Brownian motion, the average value will increase without bound. Recall that the distribution of the Brownian process at time t is the standard normal with mean 0 and variance t. The question becomes what is the fastest the value function can grow, as a function of R and still have a finite expected value with respect to the normal distribution.

The answer seems to be NormalHedge (Eqn. 27). More precisely, if $\epsilon > 0$, the mean value is finite, but if $\epsilon = 0$ the mean value becomes infinite.

A Proof of Theorem 5

1. By symmetry adversarial strategy (18) guarantees that the aggregate loss (12) is zero regardless of the choice of the learner: $\ell(t) = 0$. Therefor the state update (13) is equivalent to the symmetric random walk:

$$\mathbf{\Psi}(i) = \frac{1}{2}((\mathbf{\Psi}(i-1) \oplus 1) + (\mathbf{\Psi}(i-1) \ominus 1))$$

Which in turn implies that if the adversary plays Q^* and the learner plays an arbitrary strategy P

$$\phi^{\uparrow}(i-1, R) = \frac{1}{2}(\phi(i, R-1) + \phi(i, R+1)) \tag{29}$$

As this adversarial strategy is oblivious to the strategy, it guarantees that the average value at iteration i is *equal* to the average of the lower value at iteration $i-1$.

2. Plugging learner's strategy (20) into equation (12) we find that

$$\ell(i-1) = \frac{1}{Z_{i-1}} \mathbf{E}_{R \sim \mathbf{\Psi}(i-1)} \left[(\phi(i, R+1+c) - \phi(i, R-1-c)) B(i-1, R) \right] \tag{30}$$

Consider the average value at iteration $i-1$ when the learner's strategy is P^* and the adversarial strategy is arbitrary Q:

$$\Phi_{P^*,Q}(i-1, R) = \mathbf{E}_{R \sim \mathbf{\Psi}(i-1)} \left[\mathbf{E}_{y \sim Q(i-1)(R)} \left[\phi(i, R + y - \ell(i-1)) \right] \right] \tag{31}$$

As $\phi(i, \cdot)$ is convex and as $(y - \ell(i-1)) \in [-1-c, 1+c]$,

$$\phi(i, R+y) \leq \frac{\phi(i, R+1+c) + \phi(i, R-1-c)}{2} + (y - \ell_P(i)) \frac{\phi(i, R+1+c) - \phi(i, R-1-c)}{2} \tag{32}$$

Combining the equations (30) and (31) we find that

$$\begin{aligned}
\Phi_{P^*,Q}(i-1, R) &= \mathbf{E}_{R \sim \mathbf{\Psi}(i-1)} \left[\mathbf{E}_{y \sim Q(i-1)(R)} \left[\phi(i, R + y - \ell(i-1)) \right] \right] && (33) \\
&\leq \mathbf{E}_{R \sim \mathbf{\Psi}(i-1)} \left[\frac{\phi(i, R+1+c) + \phi(i, R-1-c)}{2} \right] && (34) \\
&+ \mathbf{E}_{R \sim \mathbf{\Psi}(i-1)} \left[\mathbf{E}_{y \sim Q(i-1)(R)} \left[(y - \ell(i-1)) \frac{\phi(i, R+1+c) - \phi(i, R-1-c)}{2} \right] \right] && (35)
\end{aligned}$$

The final step is to show that the term (35) is equal to zero. As $\ell(i-1)$ is a constant with respect to R and y the term (35) can be written as:

$$\begin{aligned}
& \mathbf{E}_{R\sim\mathbf{\Psi}(i-1)}\left[\mathbf{E}_{y\sim Q(i-1)(R)}\left[(y-\ell(i-1))\frac{\phi(i+1,R+1)-\phi(i+1,R-1)}{2}\right]\right] && (36)\\
= \; & \mathbf{E}_{R\sim\mathbf{\Psi}(i-1)}\left[B(i-1,R)\frac{\phi(i,R+1+c)-\phi(i,R-1-c)}{2}\right] && (37)\\
- \; & \ell_P(i)\mathbf{E}_{R\sim\mathbf{\Psi}(i-1)}\left[\frac{\phi(i,R+1+c)-\phi(i,R-1-c)}{2}\right] && (38)\\
= \; & 0 && (39)
\end{aligned}$$

B Proof of Lemma 6

We prove the lemma by showing that $\Phi_{P,Q}(i) = \Phi_{P,Q}(i+1)$ for all $i \in \{T-1, T-2, \ldots, 0\}$

$$\begin{aligned}
\Phi_{P,Q}(i+1) &= \mathbf{E}_{R\sim\mathbf{\Psi}(i+1)}\left[\phi_{P,Q}\left(i+1,R\right)\right] && (40)\\
&= \mathbf{E}_{R\sim\mathbf{\Psi}(i),\, l\sim Q(i,R)}\left[\phi_{P,Q}\left(i+1,R+l-\ell_P(i)\right)\right] && (41)\\
&= \mathbf{E}_{R\sim\mathbf{\Psi}(i)}\left[\mathbf{E}_{l\sim Q(i,R)}\left[\phi_{P,Q}\left(i+1,R+l-\ell_P(i)\right)\right]\right] && (42)\\
&= \mathbf{E}_{R\sim\mathbf{\Psi}(i)}\left[\phi_{P,Q}(i,R)\right] && (43)\\
&= \Phi_{P,Q}(i) && (44)
\end{aligned}$$

C Proofs of Lemma 9 and Theorem 8

of Lemma 9 Recall the notations $\Delta t = 2^{-2k}$ $t' = t + \Delta t$ and $s = 2^{-k}$. We can write out explicit expressions for the two values:

- For strategy S_0 the value is

$$\phi_k(t,R) = \frac{\phi(t',R+s)+\phi(t',R-s)}{2}$$

.

- For strategy S_1 the value is

$$\phi_{k+1}(t,R) = \frac{1}{16}(\phi(t',R+2s)+4\phi(t',R+s)+6\phi(t',R)+4\phi(t',R-s)+\phi(t',R-2s))$$

.

We want to show that $\phi_1(T-1,R) > \phi_0(T-1,R)$ for all R, in other words we want to characterize the properties of $\phi(T,R)$ the would guarantee that

$$\phi_1(t,R)-\phi_0(t,R) = \frac{1}{16}(\phi(t',R+2) - 4\phi(t',R+1) + 6\phi(t',R) - 4\phi(t',R-1) + \phi(t',R-2)) > 0 \tag{45}$$

Inequalities of this form have been studied extensively under the name "divided differences" [**?**, **?**, **?**]. A function $\phi(T,R)$ that satisfies inequality 45 is said to be *4'th order convex* (see details in in [**?**]).

n-convex functions have a very simple characterization:

Theorem 10 *Let f be a function with is differentiable up to order n, and let $f^{(n)}$ denote the n'th derivative, then f is n-convex (n-strictly convex) if and only if $f^{(n)} \geq 0$ ($f^{(n)} > 0$).*

We conclude that if $\phi(t',R)$ has a strictly positive fourth derivative then $\phi_{k+1}(t,R) > \phi_k(t,R)$ for all R, proving the first part of the lemma.

The second part of the lemma follows from the fact that both $\phi_{k+1}(t,R)$ and $\phi_k(t,R)$ are convex combinations of $\phi(t,R)$ and therefor retain their continuity and convexity properties.

of Theorem 8

The proof is by double induction over k and over t. For a fixed k we take a finite backward induction over $t = T - 2^{-2k}, T - 2 \times 2^{-2k}, T - 3 \times 2^{-2k}, \cdots, 0$. Our inductive claims are that $\phi_{k+1}(t,R) > \phi_k(t,R)$ and $\phi_{k+1}(t,R)$,$\phi_k(t,R)$ are continuous, strongly convex and have a strongly positive fourth derivative. That these claims carry over from $t = T - i \times 2^{-2k}$ to $t = T - (i+1) \times 2^{-2k}$ follows directly from Lemma 9.

The theorem follows by forward induction on k.