Probability Tokenization

A modern approach for automatic prediction market

catslovefish.eth and **Edward Lee**

January 27, 2025

Abstract. We introduce a *probabilistic tokenization* protocol that designates each event outcome as a fungible ERC-20 token, underpinned by a single bonding curve reserve. A reserve-to-tokens mapping $f^{-1} : \mathbb{R} \to \mathbb{R}^N$ (and $f : \mathbb{R}^N \to \mathbb{R}$ vise versa) ensures that each USD deposit co-mints an identical supply of *all* outcome tokens, anchoring their amounts to the reserve and preserving a conservation law from both a *local derivative* (price) and a *global fraction* (probability) perspective. Consequently, the total of these outcome-token prices in USD remains exactly one, establishing a robust "price–probability" duality and simplifying final settlement—only the correctly realized outcome can be redeemed at its designated value. Moreover, this property persists under external trading, since the bonding curve fixes each outcome's total supply. By mapping traditional binary option contracts into interoperable ERC-20 tokens, the framework provides a scalable, transparent, and composable protocol for on-chain probability estimation and information aggregation.

1 Introduction

From the moment our distant ancestors first tumbled from branches to forage and hunt, **prediction** has been integral to survival and progress. Whether deciding when to harvest crops or where to invest resources, the ability to forecast future outcomes underpins nearly every human endeavor. Historically, people turned to oracles, shamans, or witches for mystical guidance. In modern times, we rely on science and powerful computers to analyze data—yet no single machine can capture all the variables of complex phenomena, like multi-body physics or quantum-scale biological processes. We also look to experts, such as those who poll and predict U.S. elections, but these specialists are neither omniscient nor always incentivized to produce the most accurate estimates.

Such limitations motivate a search for alternative ways to gather and synthesize information. One promising approach, rooted in *game theory*, leverages carefully designed incentive mechanisms to harness the collective wisdom of diverse participants. In practice, this often takes the form of prediction markets, which aggregate individual bets on future events into probabilistic forecasts. By tapping into the combined knowledge of participants, prediction markets can complement—and sometimes outperform—traditional expert analysis and brute-force computational methods.

Historically, the notion of wagering on future events dates back at least to the early 1500s, when people placed bets on papal succession; by the late 19th century, Wall Street was already hosting election betting for U.S. presidential races. A key theoretical milestone came in 1907, when Francis Galton showed that a crowd's median estimate could outperform individual experts, presaging the "wisdom of the crowd" idea. Economists like Friedrich Hayek and Ludwig von Mises later argued that markets are adept at aggregating dispersed information, ushering in modern prediction markets. Platforms such as the Iowa Electronic Markets (1988) and more recent on-chain protocols like Polymarket illustrate the ongoing evolution of these markets into powerful tools for capturing crowd-sourced intelligence.

This paper proposes how ***bonding curves***—can power new on-chain prediction markets with continuous liquidity management, automatic result settlement. The goal is to merge crowd-sourced insights with robust incentive designs to achieve more transparent, scalable, and composition-friendly solutions than conventional order-book systems. The remainder of this paper is organized as follows:

- **Section 2: Mathematical Model** – Presents the foundational mathematics behind our proposal.
- **Section 3: Comparison with Other Systems** – Evaluates key differences and benefits.
- **Section 4: Open Questions** – Discusses potential challenges and future directions.
- **Section 5: Acknowledgments**

2 Mathematical model

Notation	Definition
Y, N, D	fungible tokens representing the yes, no and draw outcome
$\{s_Y, s_N, s_D\}$	quantity of token set supplied by bonding curve
$f : \mathbb{R}^N \to \mathbb{R}$	mapping (function) from a set to a number
r_U	quantity of token USD[1] deposited (locked) into the bonding curve contract
r_X	quantity of token X reserved (locked) in the constant-formula-based liquidity pool
$p_{XY} = \frac{\partial Y}{\partial X}$	the relative price p_{XY} of token X in terms of token Y[2]
p_i	the probability of outcome i

Table 1: Notation

2.1 Initial Setup for Token Set

In the context of prediction markets, we can consider a mint-and-burn mechanism for a *token set*[3] that defines the relationship between the USD token reserve and the supplies of each outcome token via a *mapping function*:

$$r_U = f(S),$$

where:

- r_U is the reserve of the USD token within the bonding curve,
- $S = \{s_{outcome_1}, s_{outcome_2}, \dots, s_{outcome_N}\}$ is the set of token supplies for the respective outcomes, and
- $f : \mathbb{R}^N \to \mathbb{R}$ maps a set of outcome token supplies to the corresponding USD reserve, encapsulating the mint-and-burn dynamics.

Conversely, the inverse mapping function f^{-1} translates the USD reserve back into the required outcome-token supplies:

$$S = f^{-1}(r_U).$$

Ensuring One-to-One Final Settlement

A crucial design goal in prediction markets is that upon outcome revelation, exactly one token (the winning outcome) redeems for 1 USD each, while other outcome tokens become worthless. To achieve this in an *automatic* fashion (i.e., no extra redistribution steps), the supplies of all tokens must *match* the USD reserve:

$$r_U = s_{outcome_1} = s_{outcome_2} = \dots = s_{outcome_N}.$$

Why This Matters:

- If outcome i is prevailed, its supply $s_{outcome_i}$ *alone* equals r_U. Hence each token of outcome i redeems for exactly 1 USD, *automatically* ensuring a one-to-one liquidation price.
- No complex fractionation or reallocation is needed at settlement: once the real-world event is known, holders of the winning token can directly claim their 1 USD redemption from the bonding curve, matching token supply to the entire USD reserve.
- This design cements a clear equivalence: "$s_{outcome_i} = r_U$" means the entire USD reserve backs the single winning token supply, seamlessly facilitating final payout.

[1] In principle, this base token can be any other fungible token like DAI, ETH or SOL. We use USD here for reading fluency.

[2] This represents the amount of token Y required to mint an additional unit of token X. Additionally, it should be noted that $p_{YX} = \frac{\partial X}{\partial Y} = \frac{1}{p_{XY}}$.

[3] This pertains to a token set mapping rather than a total differential, so notation like $dr_U = ds_{outcome_1} + \cdots + ds_{outcome_N}$ is unclear. See Section 2.5 for a more intuitive illustration.

Illustrative Example

Suppose there are three possible outcomes: Yes (Y), No (N), and Draw (D). The mapping function and its inverse can be expressed as:

$$r_U = f\big(\{s_Y, s_N, s_D\}\big), \quad \{s_Y, s_N, s_D\} = f^{-1}\big(r_U\big).$$

Current State If the current supply is 10 for each outcome, we have:

$$10 = f\big(\{10, 10, 10\}\big), \quad \{10, 10, 10\} = f^{-1}\big(10\big).$$

Deposit (Mint) When a user try to mint a token set, the contract will accept 1 USD in an exchange of a complete token set {1, 1, 1}. The reserve of USD will increase to 11 and the total supply will increase to {11, 11, 11}.

Withdraw (Burn) When a user try to redeem 1 USD, the contract will accept and only accept a*complete* token pair {1, 1, 1} in an exchange of 1 USD ; submitting a single token like {1, 0, 0} is not sufficient for redemption. The reserve of USD will decrease to 9 and the total supply will increase to {9, 9, 9}.

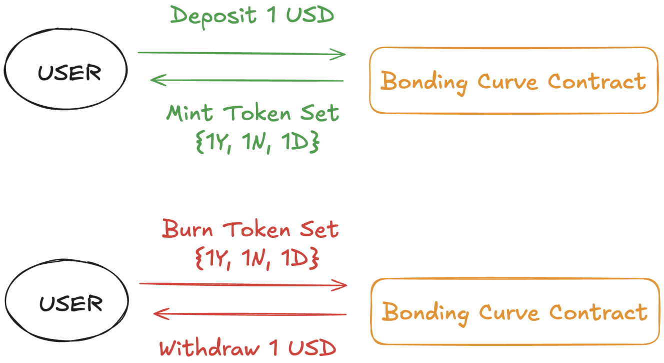

Figure 1: Illustration of the Mint-and-Burn Mechanism

2.2 Price–Probability Duality

In our mint-and-burn mechanism, each one USD deposit mints exactly one token for each outcome, and conversely, burning one complete token set from each outcome returns one USD from the contract. Under this setup, we have

$$r_U = s_1 = s_2 = \cdots = s_N,$$

meaning the reserve of USD, r_U, matches each outcome's token supply. This yields a notable *duality* between *price* and *probability*:

2.2.1 Local Viewpoint (Price)

Marginal Cost Definition

We define the local (marginal) cost of outcome i by the derivative:

$$\frac{\partial r_U}{\partial s_i} = 1.$$

This indicates that an infinitesimal increase in the supply s_i of outcome i requires exactly 1 additional USD in the reserve. In other words, each outcome token is minted at a cost of 1 USD *in isolation.*

Naïve Summation

One might reason naïvely:

$$\underbrace{1 + 1 + \cdots + 1}_{N \text{ times}} = N.$$

In other words, if each outcome token costs 1 USD, then buying one token for each of the N outcomes would total N USD.

However, this summation is *misleading* under the system rules, because you cannot mint each outcome token *individually* for 1 USD. Instead, you always mint an *entire set* of outcome tokens simultaneously with a single 1 USD deposit.

Correct Interpretation

When a user deposits 1 USD into the bonding curve, the mechanism mints *all N outcome tokens at once* without charging 1 USD each. The marginal cost equation

$$\frac{\partial r_U}{\partial s_i} = 1$$

simply tells us that, *if we were only minting token i alone*, it would cost 1 USD. But in practice, the contract lumps the outcomes together in a *set.*

Hence, the *actual* total cost to obtain one token of each outcome $\{s_1, s_2, \ldots, s_N\}$ is still one USD in *one go*, not N USD. The derivative-based perspective ($\partial r_U/\partial s_i = 1$) does not add up across outcomes, because we do *not* purchase them separately.

$$(\text{Cost in USD per extra token}) = 1, \quad \forall\, i \in \{1, 2, \ldots, N\}.$$

Thus, from a purely derivative-based viewpoint, *every* outcome token's marginal cost is indeed one USD, yet we do *not* pay $1 \times N$ in total for all outcomes combined. Instead, we pay only *one* USD *in total*, because we always mint or burn them collectively.

Global Viewpoint (Fraction / Probability)

Share of the Entire System:

$$p_i = \frac{s_i}{\sum_{j=1}^{N} s_j}.$$

Since $s_i = r_U$ and $\sum_{j=1}^{N} s_j = N \times r_U$, we obtain

$$p_i = \frac{r_U}{N\, r_U} = \frac{1}{N}.$$

In other words, from a global ratio standpoint, each outcome holds exactly $\frac{1}{N}$ of the total "sample space." This fraction can be seen as a *probability*-like measure, suggesting equal weighting across all N outcomes.

2.2.2 Why This Duality Matters

- **Local vs. Global:** A *local* (derivative-based) lens reveals how the USD reserve changes if we tweak a single outcome token's supply. Here, every extra token costs precisely 1 USD. Meanwhile, a *global* (fraction-based) lens shows each outcome taking an equal share of the system, $\frac{1}{N}$. Both viewpoints describe the same mechanism from different angles.

- **One-Time Payment vs. Naïve Summation:** If each token's marginal cost is 1 USD, a naive addition might suggest N USD for N outcomes. In reality, the mechanism mints *all* outcome tokens at once for just 1 USD total, preventing double counting. This is consistent with the fraction viewpoint, where each token is only one piece of a unified 1 USD "pie."

- **Uniform Scenario:** Because all outcome supplies are identical ($s_1 = s_2 = \cdots = s_N = r_U$), both local cost (1 USD/token) and global fraction ($\frac{1}{N}$ each) coincide neatly. The entire system is perfectly symmetrical across outcomes, namely $p_{iU} = p_i = \frac{1}{N}$.

- **Relevance to Prediction Markets:** In prediction markets, "price" can reflect marginal cost or derivative logic (*how much USD to mint one more token*), while "probability" can reflect a ratio-of-supplies (*which outcome fraction of the total is allocated*). By showing these sums align, it reassures that one cannot "overpay" or "double count" the cost. Each user sees both a fair local cost and a global fraction that reliably sum to 1.

2.3 Introducing External Exchange

The previously described setup for mint-and-burn within the bonding curve can seem rather *static*, as each outcome token initially appears to have a fixed 1:1 cost in USD. However, once users are allowed to *freely trade* these outcome tokens on an external market (e.g., a Uniswap-style liquidity pool or a centralized exchange), the "price" becomes more dynamic. Crucially, we can still show that the sum of outcome prices in terms of USD remains 1, unifying the local (derivative) and global (fraction) perspectives.

Two-Outcome Analysis

For simplicity, consider two outcomes, Y (Yes) and N (No):

Step 1: Deposit USD into the Bonding Curve A user deposits Δ USD into the bonding curve, receiving Δ Y tokens and Δ N tokens. In the strict local sense, each outcome token can be seen as costing 1 USD, but this is part of a *combined* (minted-at-once) mechanism.

Step 2: Swap Token N for Token Y The user then *swaps* Δ N tokens to acquire additional Y tokens in an external market (e.g., Uniswap or Binance) with some exchange rate p_{YN}. Symbolically:

$$\Delta N \quad \longrightarrow \quad p_{YN}\,\Delta Y.$$

Step 3: Total Amount of Y Tokens Including the Δ Y tokens originally minted from the bonding curve, the user ends up with:

$$\Delta + p_{YN}\,\Delta$$

Y tokens in total. At a small scale ($\Delta \to 0$), we capture this trade's effect through partial derivatives.

Step 4: Relative Price of Y in Terms of USD Taking $\Delta \to 0$, we see:

$$p_{YU} = \lim_{\Delta \to 0} \frac{\Delta}{\Delta + p_{YN}\,\Delta} = \frac{1}{1 + p_{YN}}.$$

This fraction-based expression can be viewed as a "local derivative" approach to the cost ratio for outcome Y.

Step 5: Relative Price of N in Terms of USD By symmetry,

$$p_{NU} = \frac{1}{1 + p_{NY}} = \frac{p_{YN}}{1 + p_{YN}} \quad \text{(using } p_{NY} = \tfrac{1}{p_{YN}}\text{)}.$$

Step 6: Sum of Outcome Prices (or Probabilities) Combining both:

$$p_{YU} + p_{NU} = \frac{1}{1 + p_{YN}} + \frac{p_{YN}}{1 + p_{YN}} = 1.$$

Therefore, even with external token swaps, the total "value share" across outcomes remains fully accounted for, adding to 1.

Key Insight: Internal Bonding Preserves the Sum

No matter how users exchange tokens among themselves (Y for N, or vice versa), the **overall system supply** for each outcome is still governed by the mint-and-burn mechanism. The external market merely re-distributes existing Y and N among participants; it does not create or destroy any additional outcome tokens.

- **Local/Derivative Angle:** Each incremental trade can be broken down via partial derivatives, showing the cost for "a bit more Y" in place of "a bit less N." Summing these local prices always yields a full "1 unit of value."
- **Global/Fraction Angle:** Globally, Y and N still partition a single "sample space," so Y's fraction+N's fraction = 1. By maintaining a one-to-one match between token supply and reserve at *all* times, the system mirrors a probability-like partition of outcomes, guaranteeing $\sum$(probability) = 1.

- **Protocol Neutrality vs. Market Beliefs:** Internally, the bonding curve itself is "innocent", initially treating both outcomes as equally likely ($p = \frac{1}{2}$ each, or $\frac{1}{N}$ in the N-outcome case). In quantum physics language, the bonding curve maintains the outcome set a as a superposition within a black box and collapse to one result when settlement. However, once outcome tokens are free to trade on external markets, their prices adjust based on participants' knowledge and incentives, reflecting the evolving *market-driven* estimate of each outcome's probability.

- **What Restriction Do We Impose?**
  In an efficient market, the relative prices satisfy

  $$\frac{p_{YU}}{p_{NU}} = p_{YN},$$

  where

  $$p_{YU} = \frac{\partial U}{\partial Y} \quad , \quad p_{NU} = \frac{\partial U}{\partial N} \quad \text{and} \quad p_{YN} = \frac{\partial N}{\partial Y}.$$

  However, $p_{YU} + p_{NU}$ can freely vary and is not fixed by this ratio condition alone; therefore, our key design choice is to *restrict* that sum to be exactly one, effectively forcing

  $$p_{YU} + p_{NU} = 1,$$

  and thereby enforcing a "probability-like" partition of value across the two outcomes.

Thus, the **dual perspective** of local derivative (price) and global fraction (probability) endures even in the presence of an external market, thanks to the *internal bonding mechanism* that fixes the total supply for each outcome. Users may shuffle tokens among themselves, but the sum of outcome-token prices in USD remains one, maintaining a coherent *price–probability duality* across the system.

2.4 Detailed Token Flow

Also, we can show the process combined with a Uniswap-style mechanism.

Roles Defined

- **User:** An individual who interacts with the system by depositing USD to mint tokens or burn tokens to retrieve USD.

- **Bonding Curve Contract:** The smart contract that manages the mint and burn of token pairs Y and N, serving as the counterpart of users.

2.4.1 Deposit (Mint)

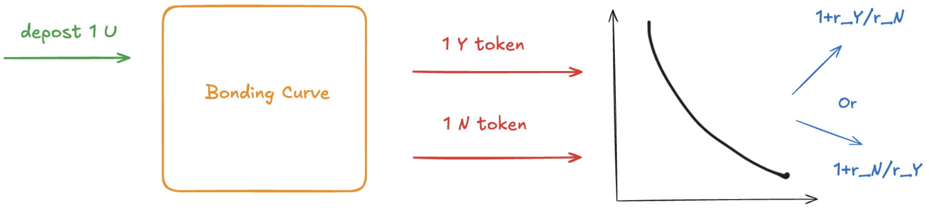

Figure 2: Illustration of Token Flow

Step 1: Deposit into the Bonding Curve We deposit Δ USD tokens into the bonding curve, receiving Δ Y tokens and Δ N tokens in return.

Step 2: Swap N Tokens for Y Tokens We swap Δ N tokens for Y tokens from the liquidity pool, which operates under the constant product formula:

$$r_Y \cdot r_N = k_{Y\&N}.$$

The change in the reserve of Y tokens is:

$$r_Y - r'_Y = \frac{k}{r_N} - \frac{k}{r_N + \Delta}, \tag{1}$$

reflecting how the Y-reserve adjusts when Δ N tokens are traded.

Step 3: Total Amount of Y Tokens Including the initial Δ Y tokens minted from the bonding curve, the total amount of Y tokens is:

$$\Delta + \frac{k}{r_N} - \frac{k}{r_N + \Delta}. \tag{2}$$

Step 4: Calculating the Relative Price of Token Y in Terms of USD To determine the relative price $p_{Y \to U}$ of Token Y in terms of USD, we take the limit as Δ approaches zero:

$$\begin{aligned}
p_{Y \to U} &= \lim_{\Delta \to 0} \frac{\Delta}{\Delta + \frac{k}{r_N} - \frac{k}{r_N + \Delta}} && (3)\\
&= \lim_{\Delta \to 0} \frac{\Delta}{\Delta + \frac{k\,\Delta}{r_N^2 + \Delta\, r_N}} && (4)\\
&= \frac{r_N^2}{k + r_N^2} && (5)\\
&= \frac{r_N}{r_Y + r_N}. && (6)
\end{aligned}$$

Step 5: Relative Price of Token N in Terms of USD By symmetry, the relative price p_{NU} of Token N in terms of USD is:

$$p_{NU} = \frac{r_Y}{r_N + r_Y}. \tag{7}$$

Step 6: Verification of the Sum of Relative Prices Finally, we verify that the sum of the relative probabilities equals one:

$$p_{YU} + p_{NU} = 1. \tag{8}$$

2.4.2 Overall and Individual State

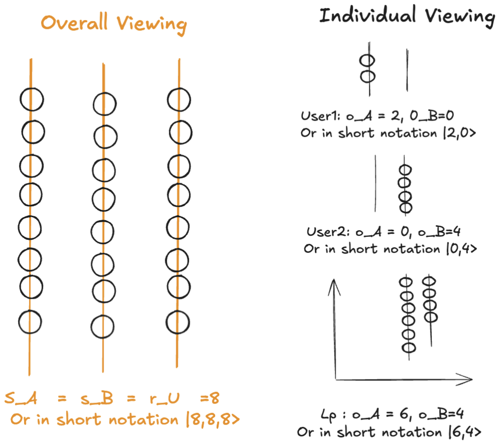

Figure 3: Overall and Individual State

2.4.3 Withdraw (Burn)

It is the reversed process of aforementioned Deposit (Mint) process, we omit the details.

2.4.4 Result Settlement

When the result of a certain outcome i prevailed and updated within the bonding curve contract, all the holder of this specific outcome can redeem 1 USD per $outcome_i$ token.

Remember, now the contract only takes $\{0, \cdots, \Delta outcome_i, ..., 0\}$ for input.

2.5 In the Language of Physics

This section models the creation and annihilation of token pairs A and B using creation operator ($\hat{c}^\dagger$) and annihilation operator ($\hat{c}$) operators, inspired [4] by the concept of Cooper pairs from quantum physics. The model integrates the role of USD as a stablecoin, defining the interactions between users and the bonding curve contract.

- **Cooper Pair (Physics):** Two electrons bound together at low temperatures, enabling superconductivity.
- **Token Pair A and B (Cryptocurrency):** Analogous to Cooper pairs, these tokens are intrinsically linked, representing a combined asset backed by USD.

Creation of a Token Pair

When a **user** deposits USD into the **bonding curve contract**, tokens A and B are minted. This process is analogous to the creation of a Cooper pair in physics.

$$\hat{c}^\dagger_A \hat{c}^\dagger_B |Vacuum\rangle = \hat{c}^\dagger_U |Vacuum\rangle = |A, B\rangle$$

Explanation:

- $\hat{c}^\dagger_A$: Creation operator for token A.
- $\hat{c}^\dagger_B$: Creation operator for token B.
- $|USD\rangle$: State representing USDT held by the bonding curve contract.
- $|A, B\rangle$: State representing the minted token pair A and B.

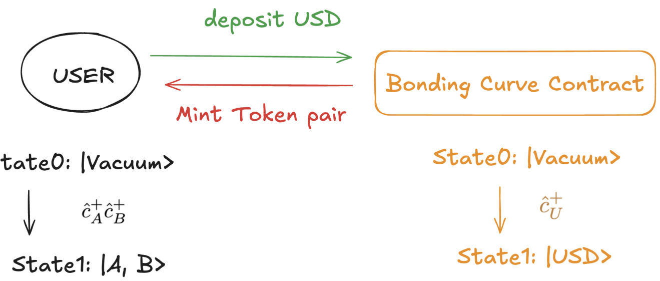

Figure 4: Token Pair Creation (Mint) Mechanism

[4] For a record where this insight comes from, see *Introduction to Many-Body Physics* by Piers Coleman. I found this book deeply inspiring as a student and still cherish my physical copy—despite an online version being available—as a tribute to both the theoretical physics and lovely memories of learning. Well, when I am typing this article, I feel I am on slippery slope away from something, maybe one day I should go back...

Annihilation of a Token Pair

When a **user** decides to burn tokens A and B, the **bonding curve contract** destroys these tokens to release the equivalent amount of USD.

$$\hat{c}_A\hat{c}_B|A,B\rangle = \hat{c}_U|A,B\rangle = |Vacuum\rangle$$

Explanation:

- $\hat{c}_A$: Annihilation operator for token A.
- $\hat{c}_B$: Annihilation operator for token B.
- $|A,B\rangle$: State representing the token pair held by the user.
- $|USD\rangle$: State representing the retrieved USD sent back to the user.

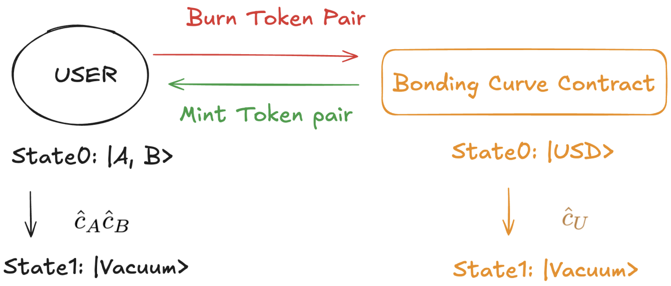

Figure 5: Token Pair Annihilation (Burn) Mechanism

Key Insight: Token Pair Mechanism

When a user mints tokens, the bonding curve contract accepts USD and issues a *pair* of tokens {A, B}. Conversely, upon burning, the contract will only redeem USD in exchange for the *complete* token pair {A, B}; submitting a single token A *or* B alone is not sufficient for redemption. This requirement ensures that the system always maintains balanced reserves relative to token supplies.

2.6 N-outcome proof

Previously, we have already shown that the sum of prices from a two-outcome token set in the existence of external market, now we are going to prove the general case.

Condition Given:

1. $p_{iU} = \frac{1}{\sum_{j=1}^{N} p_{ij}}$.
2. $\frac{p_{ij}}{p_{ik}} = p_{jk}$.

We want to show:

$$\sum_{i=1}^{N} p_{iU} = 1.$$

Step 1: Fix a Reference Outcome k Pick any outcome k as a reference. The ratio condition

$$\frac{p_{ij}}{p_{ik}} = p_{jk}$$

implies:

$$p_{ij} = \frac{p_{ik}}{p_{jk}}.$$

Step 2: Rewrite $\sum_{j=1}^{N} p_{ij}$ Using $p_{ij} = \frac{p_{ik}}{p_{jk}}$, we obtain:

$$\sum_{j=1}^{N} p_{ij} = \sum_{j=1}^{N} \frac{p_{ik}}{p_{jk}} = p_{ik} \sum_{j=1}^{N} \frac{1}{p_{jk}}.$$

Step 3: Express p_{iU} From condition (1),

$$p_{iU} = \frac{1}{\sum_{j=1}^{N} p_{ij}} = \frac{1}{p_{ik} \sum_{j=1}^{N} \frac{1}{p_{jk}}} = \frac{1}{p_{ik}} \cdot \frac{1}{\sum_{j=1}^{N} \frac{1}{p_{jk}}}.$$

Step 4: Summation Over i Consider

$$\sum_{i=1}^{N} p_{iU} = \sum_{i=1}^{N} \left(\frac{1}{p_{ik}} \cdot \frac{1}{\sum_{j=1}^{N} \frac{1}{p_{jk}}} \right).$$

Factor out $\dfrac{1}{\sum_{j=1}^{N} \frac{1}{p_{jk}}}$:

$$\sum_{i=1}^{N} p_{iU} = \frac{1}{\sum_{j=1}^{N} \frac{1}{p_{jk}}} \sum_{i=1}^{N} \frac{1}{p_{ik}}.$$

Step 5: Show $\sum_{i=1}^{N} \frac{1}{p_{ik}} = \sum_{j=1}^{N} \frac{1}{p_{jk}}$ By symmetry of the ratio condition and picking outcome k as a reference, all such sums match. Thus:

$$\sum_{i=1}^{N} \frac{1}{p_{ik}} = \sum_{j=1}^{N} \frac{1}{p_{jk}}.$$

Step 6: Conclude the Sum is 1 Therefore,

$$\sum_{i=1}^{N} p_{iU} = \frac{\sum_{j=1}^{N} \frac{1}{p_{jk}}}{\sum_{j=1}^{N} \frac{1}{p_{jk}}} = 1.$$

Hence we conclude:

$$\sum_{i=1}^{N} p_{iU} = 1.$$

Intuitive Explanation The ratio condition $\frac{p_{ij}}{p_{ik}} = p_{jk}$ ensures consistent pairwise exchange rates. By choosing a single "reference" outcome k, every outcome's price in USD is tied back to how it relates to outcome k. The function

$$p_{iU} = \frac{1}{\sum_{j=1}^{N} p_{ij}}$$

then partitions a single unit of value across the N outcomes. Summing these partitions recovers the entire unit, yielding $\sum_{i=1}^{N} p_{iU} = 1$.

2.7 Add n+1 outcome after n-outcome setup

This is very simple, all we need to do is sending extra outcome tokens to all the previous participants proportionally. The detail process is left to readers.

2.8 Is There a Room for Arbitrage?

This question can be answered both *yes* and *no*, depending on which part of the system we examine:

Yes, Arbitrage is Possible If external markets (e.g., centralized exchanges, or multiple liquidity pools) quote different exchange rates among Y–U, N–U, and D–U, there may be *triangular arbitrage* opportunities. For instance, if the combined trades across these three pools do not maintain consistent ratios, a user could profit by cycling through $(Y \to N \to D \to Y)$ or similar paths. Disparities between CEX prices and bonding-curve quotes can also create arbitrage if one market lags behind the other.

No, Not Within the Mint-and-Burn Process By contrast, the mint-and-burn mechanism itself does *not* introduce direct arbitrage inside the bonding curve logic. Once an outcome token is minted or burned, the internal bonding curve sets its supply relative to the USD reserve with no "price mismatch" for an arbitrageur to exploit. Any perceived arbitrage relies on the *external* markets' fluctuations rather than on the **local** creation or destruction of tokens.

Insight:

- *Triangular Arbitrage in External Pools*: Occurs if Y–U, N–U, and D–U exchange rates drift out of equilibrium relative to each other. A savvy trader can cycle through tokens to capitalize on small pricing inconsistencies.
- *Bonding Curve's Immunity*: The bond between token supplies and the USD reserve in the mint-and-burn scheme does not, by itself, offer a direct arbitrage path. Token supply changes happen at a fixed ratio, unaffected by external trades.

Move investigation needed here Still, this question is left open.

2.9 Tax

Taxation in this system is straightforward: a small fee is *pre-charged* whenever a user deposits or withdraws USD. For example, if a user wishes to mint one set of tokens $\{Y, N, D\}$, they might pay 1.01 USD instead of exactly 1.00 USD, reflecting a small tax. Conversely, upon burning that same token set and redeeming USD, the user could receive 0.99 USD, again accounting for the tax overhead.

In this way, the system levies a symmetrical charge on *both* depositing and withdrawing, ensuring the fee is applied consistently across all user interactions with the bonding curve.

2.10 Integration with Other DeFi

Because each outcome token is minted as a standard fungible asset, users can freely incorporate these tokens into any compatible DeFi protocol. For instance, they may provide liquidity on Uniswap, stake on Governance or lend tokens on Aave, or otherwise leverage their holdings in various yield strategies. This flexibility ensures outcome tokens retain the composability and portability characteristic of other widely adopted ERC-20 (or similar) tokens in the DeFi ecosystem.

2.11 Duality between Impermanent Loss and Deviation Loss

Because we have mapped outcome *probabilities* directly into *prices*, the same mechanism that causes "impermanent loss" in systems like Uniswap also induces what we can call "deviation loss" from a probability perspective. Essentially, if a participant bets on multiple outcomes and provides liquidity (e.g., on an AMM or external exchange), they seek to earn trading fees but must also handle the risk that the real event probabilities diverge from current market estimates.

Concept: When a participant provides liquidity in both (or all) outcomes, they face:

- **Tax/Fee Accumulation**: Earning fees from trades in the pool.
- **Deviation (Impermanent) Loss**: A potential shortfall arising if actual event probabilities (*true* prices) deviate from the participant's holdings in each outcome.Even with active management or hedging, the participant cannot fully eliminate this loss, because it stems from inherent uncertainty[5] about the final outcome.

[5] A loose analogy to Heisenberg's uncertainty principle reminds us that certain outcomes cannot be measured or predicted with complete precision:

$$\sigma_x \, \sigma_p \;\geq\; \frac{\hbar}{2},$$

indicating that fundamental limits of knowledge introduce irreducible uncertainty. In prediction markets, no strategy can perfectly hedge against every possible future shift in outcome probabilities.

- **Probability Angle:** Since the real-world probabilities may differ from perceived or posted probabilities, any misalignment can erode the value of a participant's liquidity. This phenomenon is analogous to "impermanent loss," reinterpreted in probability terms as "deviation loss."

Thus, from both a *price* and a *probability* vantage, liquidity providers in a multi-outcome system confront a trade-off between collecting fees and enduring some level of deviation loss whenever the market's probabilities evolve in ways they did not anticipate.

3 Comparison with Other Systems

To evaluate the advantages and disadvantages of the bonding curve approach, it is essential to compare it with other systems. This section clarifies key terminologies and examines how different market mechanisms operate, highlighting their respective strengths and limitations.

3.1 Traditional Bookmaker Systems

Traditional bookmaker systems rely heavily on centralized brokers who set static odds based on historical data and proprietary algorithms. These bookmakers act as counterparts for all users, facilitating bets by taking the opposite side of each wager. Key characteristics include:

- **Centralized Control**: A single entity manages the order matching process, setting odds and ensuring liquidity. This centralization can lead to potential conflicts of interest and reduced transparency.
- **Static Odds**: Bookmakers establish fixed odds prior to events, which may not accurately reflect real-time market sentiments or emerging information.
- **Liquidity Provider**: The bookmaker serves as the primary liquidity provider, absorbing all bets and managing risk internally.

While traditional bookmakers offer simplicity and ease of use, their centralized nature can result in inefficiencies such as wider spreads and slower price adjustments, potentially limiting the accuracy and reliability of predictions.

3.2 Central Limit Order Book (CLOB)

Central Limit Order Books (CLOB) are a prevalent implementation in modern prediction markets, exemplified by platforms like Polymarket. CLOB systems facilitate decentralized trading by allowing users to place buy and sell orders, which are then matched based on price and time priority. Key features include:

- **Order Matching**: Orders are matched based on bid (highest buyer price) and ask (lowest seller price), with the spread indicating liquidity—the smaller the spread, the higher the liquidity.
- **Bid and Ask**: The bid price represents the highest price a buyer is willing to pay for a security, while the ask price is the lowest price a seller is willing to accept. The difference between them, known as the spread, is a key indicator of the asset's liquidity.
- **CFT**: The Conditional Tokens Framework (CTF) is a protocol for creating tokenized logic; tokens redeemable for underlying collateral when a specific condition is true.

Polymarket Example

Polymarket operates a CLOB where all event outcomes are tokenized and traded on the Polygon network in a non-custodial manner. Specifically, Polymarket's outcome shares are binary—Yes/No—and utilize Gnosis's Conditional Tokens Framework (CTF) to represent these outcomes. The CTF allows for *splitting* and *merging* of complete outcome sets. This means that for a given condition (market), any user can split one binary outcome token (e.g., Yes) into an equivalent unit of collateral (USDC) or merge tokens back into the collateral. Consequently, in a well-structured market, the prices of these tokens should always sum to one, maintaining a probability-like partition.

Order Matching Examples

To better understand the splitting and merging mechanisms, consider the following examples:

Example 1: Non-Complementary Bid/Ask

- Jack places a limit order to buy 10 shares of the ”Yes” outcome at \$0.35 each.
- Jill places a limit order to sell 10 shares of the ”Yes” outcome at \$0.35 each.

In this case, the orders directly match, resulting in a transfer of \$3.50 from Jack to Jill in exchange for the 10 ”Yes” shares.

Example 2: Complementary Bid/Ask

- Jack places a limit order to buy 10 shares of the ”Yes” outcome at \$0.35 each.
- Jill places a limit order to buy 10 shares of the ”No” outcome at \$0.65 each.

At first glance, these orders appear unrelated. However, since one USD can be split into one ”Yes” and one ”No” token, these orders can be matched by splitting the collateral. Jack’s \$3.50 and Jill’s \$6.50 (totaling \$10.00) are used to mint 10 ”Yes” and 10 ”No” tokens, fulfilling both orders through the split operation.

Example 3: Complementary Ask

- Jack places a limit order to sell 20 shares of the ”Yes” outcome at \$0.75 each.
- Jill places a limit order to sell 20 shares of the ”No” outcome at \$0.25 each.

Here, Jack’s 20 ”Yes” shares and Jill’s 20 ”No” shares can be merged into \$20.00. Jack receives \$15.00 (20 shares × \$0.75) and Jill receives \$5.00 (20 shares × \$0.25) through the merge operation.

3.3 Robin Hanson’s Bonding Curve

For simplicity[6], we use a two-outcome scenario for illustration:

$$r_U = b \cdot \ln\left(e^{\frac{s_Y}{b}} + e^{\frac{s_N}{b}}\right),$$

$$p_{YU} = \frac{\partial r_U}{\partial s_Y} = \frac{e^{\frac{s_Y}{b}}}{e^{\frac{s_Y}{b}} + e^{\frac{s_N}{b}}},$$
$$p_{NU} = \frac{\partial r_U}{\partial s_N} = \frac{e^{\frac{s_N}{b}}}{e^{\frac{s_Y}{b}} + e^{\frac{s_N}{b}}}.$$

At first glance, this bonding curve preserves the property:

$$p_{YU} + p_{NU} = 1. \tag{9}$$

However, this curve can lead to disproportionate outcomes during the settlement stage. For example, suppose $b = 1$, $s_Y = 10$, and $s_N = 100$, then:

$$r_U = b \cdot \ln\left(e^{10} + e^{100}\right) \approx 100.$$

In this scenario, the liquidation price for ”Yes” would be 10 and for ”No” it would be 1, resulting in an unintuitive and skewed outcome.

[6] The reader can verify the general case:

$$r_U = b \cdot \ln\left(e^{\frac{s_{\text{outcome}_1}}{b}} + \ldots + e^{\frac{s_{\text{outcome}_n}}{b}} + \ldots + e^{\frac{s_{\text{outcome}_N}}{b}}\right),$$

where N is the total number of outcomes.

3.4 Bonding Curve-Based Automated Prediction Market Makers (BAPMM)

In contrast to traditional systems, Bonding Curve-Based Automated Prediction Market Makers (BAPMM) offer a decentralized and automated approach to supply management and price discovery. Key characteristics include:

- **Decentralized Supply**: The token set is provided by a bonding curve smart contract that maintains reserves of a base asset (e.g., USD) and outcome tokens, eliminating the need for centralized brokers.
- **Automatic Pricing**: Prices are determined algorithmically based on the bonding curve, which adjusts the token supply in response to user deposits and withdrawals, ensuring dynamic and real-time price adjustments.
- **Price–Probability Duality**: The bonding curve ensures that the sum of outcome-token prices remains one, maintaining a coherent probability distribution across all outcomes.
- **Seamless Result Settlement**: The bonding curve ensures one correct result to one USD redemption automatically.

Advantages Over CLOB

- **Continuous Liquidity**: Unlike CLOBs, where liquidity can be fragmented and dependent on active market makers, BAPMM ensures continuous liquidity. The supply is always available since users can mint directly from the bonding curve even when matching orders in the CLOB are sparse or nonexistent.
- **Automatic Pricing**: Prices are determined algorithmically based on the bonding curve, which adjusts the token supply in response to user deposits and withdrawals, ensuring dynamic and real-time price adjustments.
- **Automatic Handling of Imbalances**: In cases where CLOB systems fail to match orders (e.g., if the ”Yes” order is priced at $0.60 and the ”No” order at $0.20, leaving no matching), BAPMM can still facilitate transactions based on algorithmic pricing.
- **Scalability**: BAPMM can handle a large number of outcomes without the complexity of managing numerous order books, making them more scalable for events with multiple possible outcomes.
- **No Need for Operator**: BAPMM operates in an AMM style, enhancing transparency and reducing single points of operator failure compared to CLOB systems.

Use Case: Long-Tail Scenarios

While CLOB systems like Polymarket are effective for events with clear and high-interest outcomes, they may struggle in long-tail scenarios where outcomes have low probability or limited liquidity. Bonding curve-based AMMs excel in these situations by:

- **Direct Outcome Pricing**: The bonding curve can automatically adjust prices based on overall supply and demand, ensuring that even low-probability outcomes have meaningful pricing without relying on active order matching.
- **Automatic Handling of Imbalances**: In scenarios where the CLOB fails to match orders due to imbalanced prices, the bonding curve can still facilitate transactions by adjusting the token supply according to the algorithm, ensuring that the market remains functional and accurately represents collective probabilities.

3.5 An Art of Subtraction

Philosophically speaking, BAPMM presents an art of subtraction compared with CLOB (See polymarket doc for more details). They converge to the same when there is high trading volume, similar to the comparison between Uniswap-like protocols and order books [7]. Still, our protocol further removes the role of the order operator in CLOB.

4 Open Questions

Below, we highlight potential challenges and opportunities for broader applications and future research.

[7] Uniswap-like liquidity pools are essentially a promise to sell or buy within a certain price range as a market maker

4.1 Oracle

Because a blockchain has no inherent knowledge of real-world events, an external oracle must provide definitive outcome data for settlement. Among possible solutions, a zero-knowledge trustless (ZK-TLS) approach or a community-driven voting mechanism represent two major design choices. Each has trade-offs in terms of trust assumptions, incentive alignment, and decentralization, and selecting one requires careful consideration of security, reliability, and governance.

4.2 Abnormal Situations

One potential concern arises if a user irreversibly destroys winning tokens—for instance, by sending them to a blackhole address. This action locks the base asset (USD or other) in the bonding curve contract indefinitely, as no one can redeem those tokens. Should the protocol forcibly redistribute or otherwise unlock these assets, or should it remain immutable in line with blockchain principles? Another scenario is oracle failure or an indeterminate real-world outcome, such as an event that does not unfold clearly (or at all). If the oracle system chooses an incorrect answer or cannot settle at all, the question becomes how (and by whom) to initiate a fallback procedure. Extreme edge cases, like catastrophic blockchain failure or existential events, suggest the need for contingency plans that can override or amend the original settlement rules.

4.3 Governance

While the protocol automates liquidity management, it still requires a governance layer for upgrades, policy decisions, and emergency measures. True decentralization calls for community-based discussions and voting, so no single entity can unilaterally impose changes. Moreover, unknown unknowns—unforeseen issues beyond the scope of standard fail-safes—may demand human intervention or protocol alterations. Robust governance structures help maintain the system's integrity when confronting such exceptional challenges.

4.4 Generalization to Options and Futures

Although this paper focuses on probability tokenization for discrete event outcomes, the same principles can be extended to more complex financial instruments such as options and futures. Designing suitable bonding curves for continuous payoffs or time-dependent contracts opens further avenues for research. This broader scope could integrate price discovery, hedging, and settlement across a range of traditional and novel derivatives markets, all within a composable, on-chain framework.

5 Acknowledgments

We gratefully acknowledge the joint efforts of catslovefish.eth, who developed the theoretical foundations of the price–probability duality and served as the primary typist for this manuscript, and Edward Lee, whose insightful direction guided the project's scope and focus. Cat also wishes to extend personal thanks to two former physics supervisors, whose mentorship fostered both a deep understanding of many-body physics and a genuine joy in exploring elegant theoretical models, as well as to an Internet artist who introduced me to the film *Solaris*, offering a fresh emotional perspective on the "wisdom of the crowd".