On the Soundness of Interactions via LogUp

OpenVM Authors

March 5, 2025

1 Preliminaries and Definitions

Throughout, let $\mathbb{F}_q$ be a finite field of order q.

Definition 1.1 (Polynomial Hash Family)**.** For each $\beta \in \mathbb{F}_q$, define the function

$$h_\beta \colon \mathbb{F}_q^\ell \to \mathbb{F}_q \quad \text{by} \quad h_\beta(\sigma_1, \ldots, \sigma_\ell) = \sum_{i=1}^{\ell} \beta^{i-1} \sigma_i.$$

We call $\{h_\beta : \beta \in \mathbb{F}_q\}$ the *polynomial hash family* of degree $\ell - 1$.

Definition 1.2 ($\mathbb{F}_q$-valued multiset)**.** An *$\mathbb{F}_q$-multiset* M over a set S is a function $M : S \to \mathbb{F}_q$ that assigns a *multiplicity* to each element in S. The *support* of M, denoted $\mathrm{supp}(M)$, is equal to $\{x \in S : M(x) \neq 0\}$.

2 Key Lemmas

The following lemma is for the case $\ell = 1$. We will use this lemma to prove things about the general case below.

Lemma 2.1 (Partial Fractions)**.** *Let M be an $\mathbb{F}_q$-valued multiset over $\mathbb{F}_q$ and let $k = |\mathrm{supp}(M)|$. Suppose there exists some $\sigma^* \in \mathbb{F}_q$ such that $M(\sigma^*) \neq 0$. Then for a uniformly random $\alpha \in \mathbb{F}_q$,*

$$\Pr\left(\sum_{\sigma \in \mathbb{F}_q} \frac{M(\sigma)}{\alpha + \sigma} = 0 \right) \leq \frac{k-1}{q}.$$

Proof. Consider the rational function

$$R(X) = \sum_{\sigma \in \mathbb{F}_q} \frac{M(\sigma)}{X + \sigma}.$$

Let $S = \mathrm{supp}(M)$ and express $R(X)$ as a rational function

$$R(X) = \frac{P(X)}{Q(X)},$$

with

$$P(X) = \sum_{\sigma \in S} M_\sigma \prod_{\tau \in S \setminus \{\sigma\}} (X + \tau),$$

and

$$Q(X) = \prod_{\sigma \in S} (X + \sigma).$$

Note that $P(X)$ and $Q(X)$ share no roots, so the zeroes of $R(X)$ are the same as those of $P(X)$. Moreover, since $\sigma^* \in S$, we have $P(-\sigma^*) = M_{\sigma^*} \prod_{\tau \in S \setminus \sigma^*} (-\sigma^* + \tau)$, which is nonzero since none of the terms of the product are zero. Hence P is not identically zero. Since a nonzero polynomial of degree at most $k-1$ has at most $k-1$ roots in $\mathbb{F}_q$, we obtain

$$\Pr(R(\alpha) = 0) = \Pr(P(\alpha) = 0) \leq \frac{k-1}{q}. \qquad \square$$

The following lemma guarantees that, for a random β, *some* message in Σ is mapped to a unique hash value (it does *not* require full injectivity—this allows us to achieve much smaller soundness error).

Lemma 2.2 (Hashing Lemma: At-Least-One Isolation)**.** *Let $\Sigma \subseteq \mathbb{F}_q^\ell$ with $|\Sigma| = k$ and let $\beta \sim \mathbb{F}_q$ be uniformly random. With probability at least $1 - \frac{(k-1)(\ell-1)}{q}$, there exists some $\sigma \in \Sigma$ such that $h_\beta(\sigma)$ differs from $h_\beta(\tau)$ for all $\tau \in \Sigma \setminus \{\sigma\}$.*

Proof. Fix an arbitrary $\sigma \in \Sigma$. For each $\tau \in \Sigma \setminus \{\sigma\}$, consider the event $E_\tau = \{h_\beta(\sigma) = h_\beta(\tau)\}$. Since

$$h_\beta(\sigma) = h_\beta(\tau) \iff \sum_{i=1}^{\ell} \beta^{i-1}(\sigma_i - \tau_i) = 0,$$

the event E_τ occurs iff a nonzero polynomial (because $\sigma \neq \tau$) in β evaluates to zero. A nonzero polynomial of degree at most $\ell - 1$ over $\mathbb{F}_q$ has at most $\ell - 1$ roots, so

$$\Pr\big(h_\beta(\sigma) = h_\beta(\tau)\big) \leq \frac{\ell - 1}{q}.$$

A union bound over all events E_τ for $\tau \in \Sigma \setminus \{\sigma\}$ shows

$$\Pr\Big(\exists\, \tau \in \Sigma \setminus \{\sigma\} :\, h_\beta(\sigma) = h_\beta(\tau)\Big) \;\leq\; (|\Sigma| - 1) \cdot \frac{\ell - 1}{q} \;=\; \frac{(k-1)(\ell-1)}{q}. \qquad \square$$

Remark 2.3. The bound above shows that the failure probability is $\lesssim \frac{\ell k}{q}$. We can improve this to $\lesssim \frac{\ell k}{2q}$ with a (very) slightly more complicated argument. Let $Z = |h_\beta(\Sigma \setminus \{\sigma\})|$. Consider the event $A = \{Z \leq k/2\}$. If A does not occur, then the average number of collisions per non-empty bin is less than two, so some non-empty bin will have exactly one element. Otherwise, we only have to union bound over a set of at most $k/2$ bins.

More formally, Let E be the event that there is at least one isolation. Since $\Pr(E^c \mid A^c) = 0$, we have

$$\Pr(E^c) = \Pr(E^c \mid A)\Pr(A) \leq \Pr(E^c \mid A) \leq \frac{k}{2} \cdot \frac{\ell - 1}{q},$$

where the last inequality follows from a union bound, similar to the proof above.

Lemma 2.4. *Let M be an $\mathbb{F}_q$-valued multiset over $\mathbb{F}_q^\ell$ with support size k. Suppose there exists some σ^* such that $M(\sigma^*) \neq 0$. Let α and β be sampled independently and uniformly from $\mathbb{F}_q$. Then*

$$\Pr\left(\sum_{\sigma \in \mathbb{F}_q^\ell} \frac{M(\sigma)}{\alpha + h_\beta(\sigma)} = 0\right) \;\leq\; \frac{\ell\,(k-1)}{q}.$$

Proof. Define $\tau_i = h_\beta(\sigma_i)$ for $i \in [N]$. Let A be the event that *no* message is isolated under h_β. By Lemma 2.2,

$$\Pr(A) \leq \frac{(k-1)(\ell-1)}{q}.$$

Let N be the $\mathbb{F}_q$-valued multiset over $\mathbb{F}_q$ defined by

$$N(\tau) = \sum_{\sigma \in \mathbb{F}_q^\ell} M(\sigma) \cdot \mathbf{1}\{h_\beta(\sigma) = \tau\}.$$

Define the event B as the event that the sum of interest evaluates to zero, i.e.,

$$B = \left\{\sum_{\tau \in \mathbb{F}_q^\ell} \frac{N(\tau)}{\alpha + \tau_i} = 0\right\}.$$

Applying the law of total probability,

$$\Pr(B) = \Pr(B \mid A)\Pr(A) + \Pr(B \mid A^c)\Pr(A^c) \leq \Pr(A) + \Pr(B \mid A^c).$$

Note that if A^c occurs, then there is at least one isolation, which implies that $N(\tau^*) \neq 0$ for some $\tau^* \in \mathbb{F}_q$. By Lemma 2.1, it then follows that $\Pr(B \mid A^c) \leq \frac{k-1}{q}$. Hence

$$\Pr(B) \leq \frac{(k-1)(\ell-1)}{q} + \frac{k-1}{q} = \frac{\ell(k-1)}{q}. \qquad \square$$

Remark 2.5. It may be unlikely that the bounds in the above lemma can be improved much. Indeed, if we interpret the sum $\sum_{\sigma \in \mathbb{F}_q^\ell} \frac{M(\sigma)}{\alpha + h_\beta(\sigma)}$ as a function of β, then we can rewrite it as $P(\beta)/Q(\beta)$ in which the polynomial P has degree $(\ell-1)(k-1)$. On the other hand, it is not immediately clear to us how to construct such a lower bound.

3 Interactions

Let $\mathbb{F}_q^+ = \bigcup_{t \geq 1} \mathbb{F}_q^t$. An *interaction message* (over $\mathbb{F}_q$) is a triple $(\sigma, m, b) \in \mathbb{F}_q^+ \times \mathbb{F}_q \times (\mathbb{F}_q \setminus \{0\})$. We call σ the *message*, m the *multiplicity*, and b the *bus index*.

Definition 3.1 (LogUp Sum)**.** Given a collection of *message-multiplicity* pairs $\{(\sigma_i, m_i)\}_{i=1}^N$, where $\sigma_i \in \mathbb{F}_q^\ell$ and $m_i \in \mathbb{F}_q$, and given parameters $\alpha, \beta \in \mathbb{F}_q$, define the *LogUp sum*:

$$\mathrm{LogUp}_{\alpha,\beta}\big(\{\sigma_i, m_i\}\big) \;=\; \sum_{i=1}^{N} \frac{m_i}{\alpha \;+\; h_\beta(\sigma_i)}.$$

Definition 3.2 (Concatenation)**.** Let $\sigma \in \mathbb{F}_q^\ell$ and let $\tau \in \mathbb{F}_q^t$. The *concatenation* of σ and τ, denoted $\sigma \circ \tau$, is the element of $\mathbb{F}_q^{\ell+t}$ given by

$$(\sigma \circ \tau)_i = \begin{cases} \sigma_i & \text{if } i \leq \ell, \\ \tau_{i-\ell} & \text{otherwise} \end{cases}$$

for $i \in [\ell + t]$.

For the rest of this note, let $(\sigma_1, m_1, b_1), \ldots, (\sigma_N, m_N, b_N)$ be a sequence of interaction messages. Set $\ell = \max_i |\sigma_i|$. For each $i \in [N]$, letting $\ell_i = |\sigma_i|$, define $\sigma_i' = \sigma_i \circ b_i \circ 0^{\ell - \ell_i}$. Note that $\sigma_i' \in \mathbb{F}_q^{\ell+1}$.

Definition 3.3 (Balanced bus)**.** A bus b is *balanced* if

$$\sum_{i=1}^{N} m_i \cdot \mathbf{1}\{b = b_i \wedge \sigma = \sigma_i\} = 0.$$

Theorem 3.4 (LogUp Theorem for $\mathbb{F}$-multiplicities)**.** *Let $\alpha, \beta \in \mathbb{F}_q$ be independent and uniformly random.*

- *Completeness: If every bus is balanced, then* $\mathrm{LogUp}_{\alpha,\beta}(\{(\sigma_i', m_i)\}) = 0$.
- *Soundness: If some bus is not balanced, then*

$$\Pr(\mathrm{LogUp}_{\alpha,\beta}(\{(\sigma_i', m_i)\}) \neq 0) \geq 1 - \frac{(\ell+1)(k-1)}{q}.$$

Proof. Completeness is obvious. For soundness, first note that since $b_i \neq 0$ for all i, if $b_i \neq b_j$, then $\sigma_i' \neq \sigma_j'$. Indeed, suppose without loss of generality that $|\sigma_i| \geq |\sigma_j|$. Then the two messages σ_i and σ_j differ in position $|\sigma_i| + 1$. In other words, $(b_i, \sigma_i) \neq (b_j, \sigma_j)$ if and only if $\sigma_i' \neq \sigma_j'$.

Define the $\mathbb{F}_q$-valued multiset M over $\mathbb{F}_q^{\ell+1}$ by

$$M(\sigma') = \sum_{i=1}^{N} m_i \cdot \mathbf{1}\{\sigma' = \sigma_i'\}.$$

Observe then that

$$\mathrm{LogUp}_{\alpha,\beta}(\{(\sigma_i', m_i)\}) = \sum_{i=1}^{N} \frac{m_i}{\alpha + h_\beta(\sigma_i')} = \sum_{\sigma' \in \mathbb{F}_q^{\ell+1}} \frac{M(\sigma')}{\alpha + h_\beta(\sigma')}.$$

Hence by Lemma 2.4, it follows that

$$\Pr(\mathrm{LogUp}_{\alpha,\beta}(\{(\sigma_i', m_i)\}) = 0) \leq \frac{(\ell+1)(k-1)}{q}. \qquad \square$$

We now show that a balanced bus, together with some other simple conditions that prevent field characteristic overflow, allow to treat the multiplicities as integers.

Lemma 3.5. *Let $n_1, \ldots, n_k \in \mathbb{Z}$ and suppose $\sum_{i=1}^{k} n_i \equiv 0 \pmod{p}$. Let $I = \{i \in [k] : n_i \geq 0\}$ be the indices corresponding to nonnegative representations. If*

$$\sum_{i \in I} n_i < p,$$

then

$$\sum_{i \in I} n_i \leq -\sum_{i \notin I} n_i.$$

Proof. We have

$$\sum_{i \in I} n_i + \sum_{i \notin I} n_i \leq \sum_{i \in I} n_i < p.$$

Since p divides the sum on the left, it follows that

$$\sum_{i \in I} n_i + \sum_{i \notin I} n_i \leq 0,$$

as desired. □

If we identify the mulitiplicities m_i with integers (negative or positive), then this partitions the messages into two sets: those associated with positive values and those with negative. If the bus is balanced, we can conclude that these two sets are equal as $\mathbb{F}_q$-valued multisets. The above lemma says that if the, say, positive multiplicities do not overflow the field, then the positive set is a (multi-)subset of the negative set. By applying the same argument, if the negative set also do not overflow the field characteristic, then the two multisets are equal.

In the theorem below, let $\varphi : \mathbb{F}_p \to \mathbb{Z}$ be the canonical mapping that sends each element in $\mathbb{F}_p$ to its smallest nonnegative integer representation.

Corollary 3.6. *Suppose $m_{i_1}, \ldots, m_{i_t} \in \mathbb{F}_p$ are multiplicities of a given balanced bus b, and let $I \subseteq [t]$ and let $J = [t] \setminus I$. For $k \in [t]$, define*

$$n_k = \begin{cases} \varphi(m_{i_k}) & \text{if } k \in I, \\ \varphi(m_{i_k}) - p & \text{if } k \in J. \end{cases}$$

Define two (integer-valued) multisets M_I and M_J by $M_I(\sigma) = \sum_{k \in I} n_k \cdot \mathbf{1}\{\sigma = \sigma_{i_k}\}$ and $M_J(\sigma) = -\sum_{k \in J} n_k \cdot \mathbf{1}\{\sigma = \sigma_{i_k}\}$. If $\sum_{k \in I} n_k < p$, then $M_I(\sigma) \leq M_J(\sigma)$ for all σ. Moreover, if we also have $\sum_{k \in J} n_k > -p$, then $M_J(\sigma) = M_I(\sigma)$.

Proof. Fix a message σ and let K be the indices such that $\sigma_k = \sigma$ for all $k \in K$. Note that since $n_k \equiv m_{i_k} \pmod{p}$ and the bus is balanced, we have that $\sum_{k \in K} n_k \equiv 0 \bmod p$. If $\sum_{k \in I} n_k < 0$, then clearly $\sum_{k \in K \cap I} n_k < 0$ and by Lemma 3.5, we have

$$\sum_{i \in K \cap I} n_k \leq -\sum_{k \in K \setminus I} n_k.$$

Note that the left-hand side is exactly the integer multiplicity of the message σ over the bus b in the multiset M_I, and the right-hand side is the integer multiplicity in M_J.

If also $\sum_{k \in J} n_k > -p$, then we can apply Lemma 3.5 again on $\{-n_k\}$ to obtain that

$$\sum_{i \in K \cap I} n_k \geq -\sum_{k \in K \setminus I} n_k.$$

This establishes the multiset equality. □

Lookup tables. One typical use is to treat I as the set of interaction messages not coming from a fixed "lookup table" AIR. Assign negative integer representatives to the lookup-table multiplicities and constrain the other multiplicities to be boolean (0 or 1). If the total number of lookups is less than p, the corollary guarantees that any message found in the sub-multiset given by I must also appear (with at least the same multiplicity) in the lookup-table multiset.

Permutation checks. Another use is to enforce exact equality of two multisets. Suppose the interaction-message multiplicities lie in $\{0, 1, p-1\}$. We identify these field elements respectively with the integers $\{0, 1, -1\}$. If the total number of interaction messages is at most p and the bus is balanced, then we can conclude that every interaction message appears the same amount of times with multiplicity 1 as it does with multiplicity $p-1$.

4 LogUp Soundness over BabyBear Quartic Extension

Let $p_{\text{bb}} = 15 \cdot 2^{27} + 1$. Suppose we have $\mathbb{F}_q$ with $q = p_{\text{bb}}^4$. The above theorems show we can bound the soundness error by

$$\frac{(\ell+1)(k-1)}{q},$$

where ℓ is the length of the longest interaction message and k is the number of distinct messages. This yields at least

$$123.6 - \lg(\ell+1) - \lg(k-1)$$

bits of security.

4.1 Grinding for 100 bits of security

To obtain 100 bits of security (in the random oracle model), we can use grinding. This requires the prover to, after some point in the protocol with transcript τ, find a witness $w \in \mathbb{F}_p$ that satisfies some easy-to-verify predicate P with $\Pr_{w \sim \mathbb{F}_p}(P(g(\tau \,\|\, w))) \leq 2^{-t}$, where g is the random oracle and t is a security parameter defined by the protocol. Following plonky3, in our protocol, we use the predicate $\{g(x) \bmod 2^t = 0\}$, which approximately satisfies the above property. (Another sensible candidate seems to be $\{g(x) < 2^{30-t}\}$, which does satisfy the above property.)

For LogUp, before sampling α and β (but after committing to the main trace data), we demand that the prover provides a witness w satisfying the above predicate with $t \geq \lg(k) + \lg(\ell+1) - 23.6$. Then, for each transcript up until the proof-of-work step, the probability that the then prover samples a valid witness and passes the LogUp check with a commitment of trace data that does not satisfy the LogUp constraint is at most

$$2^{-t} \cdot 2^{-(122 - \log(\ell+1) - \lg(k))} \leq 2^{-100}.$$

In practice, we may guarantee in the verifier that $k \leq 2^{30}$ and $\ell + 1 \leq 2^6$, in which case we need to set the number of proof-of-work bits for the LogUp round to 17.