Appendix

In the appendices we present the proofs, and additional lemmas that are used in the proofs.

A. Lemma 1

Lemma 1 proves that if (5) is satisfied for some action \( a \in A(I) \) on iteration \( T \), then the value of action \( a \) and all its descendants on every iteration played so far can be set to the \( T \)-near counterfactual best response value. The same lemma holds if one replaces the \( T \)-near counterfactual best response values with exact counterfactual best response values. The proof for Lemma 1 draws from recent work on warm starting CFR using only an average strategy profile (Brown & Sandholm, 2016).

**Lemma 1.** Assume \( T \) iterations of CFR with RM have been played in a two-player zero-sum game. If \( T(\psi^T, T(I, a)) \leq T_{\psi^T, T(I, a)} \) and one sets \( v_\sigma^T(I, a) = \psi^{T', T}(I, a) \) for each \( t \leq T \) and for each \( I \in D(I, a) \) sets \( v_\sigma^T(I', a') = \psi^{T', T}(T(I, a)) \) and \( v_\sigma^T(I') = \psi^{T', T}(I') \) then after \( T' \) additional iterations of CFR with RM, the bound on exploitability of \( \sigma^T + T' \) is no worse than having played \( T + T' \) iterations of CFR with RM unaltered.

**Proof.** The proof builds upon Theorem 2 in (Brown & Sandholm, 2016). Assume \( T(\psi^T, T(I, a)) \leq T_{\psi^T, T(I, a)} \). We wish to warm start to \( T \) iterations. For each \( I \in D(I, a) \) set \( v_\sigma^T(I', a') = \psi^{T', T}(I, a') \) and \( v_\sigma^T(I') = \psi^{T', T}(I') \) and set \( v_\sigma^T(I, a) = \psi^{T', T}(I, a) \) for all \( t \leq T \). For every other action, leave regret unchanged. For each \( I' \in D(I, a) \) we know by construction that \( \Phi(R^T(I')) \) is within the CFR bound \( y^T \), after changing regret. By assumption \( T(\psi^T, T(I, a)) \leq T_{\psi^T, T(I, a)} \), so \( R^T(I, a) \) \leq 0 and therefore \( \Phi(R^T(I)) \) is unchanged. Finally, since the \( T \) iterations were played according to CFR with RM and regret is unchanged for every other information set \( I' \), so the conditions for Theorem 2 in (Brown & Sandholm, 2016) hold for every information set, and therefore we can warm start to \( T \) iterations of CFR with RM with no penalty to the convergence bound. \( \square \)

B. Proof of Theorem 1

**Proof.** From Lemma 1 we can immediately set regret for \( a \in A(I) \) to \( v_\sigma^T(I, a) = \psi^{T', T}(I, a) \). By construction of \( T' \), \( R^T(I, a) \) is guaranteed to be nonpositive for \( t \leq T + T' \) and therefore \( \sigma^T(I, a) = 0 \). Thus, \( \sigma^T + T'(I') \) for \( I' \in D(I, a) \) is identical regardless of what is played in \( D(I, a) \) during \( T \leq t \leq T + T' \).

Since \( (T + T')(\psi^{T', T'+T'}(I, a)) \leq T(\psi^{T', T}(I, a)) + T'(U(I, a)) \) and \( \sum_{t=1}^{T'} v_\sigma^T(I) \geq \sum_{t=1}^{T'} v_\sigma^T(I) + T'(I(I)) \), so by the definition of \( T', (T + T')(\psi^{T', T'+T'}(I, a)) \leq \sum_{t=1}^{T'+T'} v_\sigma^T(I) \). So if regrets in \( D(I, a) \) and \( R^T + T'(I, a) \) are set according to Lemma 1, then after \( T'' \) additional iterations of CFR with RM, the bound on exploitability of \( \sigma^T + T' + T'' \) is no worse than having played \( T + T' + T'' \) iterations of CFR with RM from scratch. \( \square \)

C. Proof of Theorem 2

**Proof.** Consider an information set \( I \) and action \( a \in A(I) \) where for every opponent Nash equilibrium strategy \( \sigma^*_P(I) \), \( CBV^\sigma^*_P(I, a) < CBV^\sigma^*_P(I) \). Let \( i = P(I) \). Let \( \delta = \min_{\sigma^*_i} \left( CBV^\sigma^*_i(I) - CBV^\sigma^*_i(I, a) \right) \) where \( \Sigma^i \) is the set of Nash equilibria. Let \( \sigma^i \leq 1 - \arg \max_{\sigma^i, \in \Sigma^i} \left( CBV^\sigma^*_i(I) - CBV^\sigma^*_i(I, a) \right) \). Since \( \sigma^i \) is not a Nash equilibrium strategy and CFR converges to a Nash equilibrium strategy for both players, so there exists a \( T \) such that for all \( T \geq T \), \( CBV^\sigma^*_i(I) - CBV^\sigma^*_i(I, a) \geq \frac{3}{4} \). Let \( T_{i,a} = \frac{4\delta^2 A_i |A|}{\delta^2} \). For \( T \geq T_{i,a} \) since \( R^T \leq \sum_{i \in I} R^T(I) \), so \( CBV^\sigma^*_i(I, a) - \sum_{t=1}^{T} v_\sigma^T(I) \leq \frac{\delta^2}{4} \). Let \( T_{i,a} = \max(T_{i,a}, T_{i,a}) \) and \( \delta_{i,a} = \frac{\delta}{4} \). Then for \( T \geq T_{i,a} \), \( CBV^\sigma^*_i(I, a) - \sum_{t=1}^{T} v_\sigma^T(I) \leq -\delta_{i,a} \). \( \square \)

D. Proof of Corollary 1

**Proof.** Let \( I \notin I_S \). Then \( I \in D(I', a') \) for some \( I' \) and \( a' \in A(I') \) such that for every opponent Nash equilibrium strategy \( \sigma^*_P(I') \), \( CBV^\sigma^*_P(I', a') < CBV^\sigma^*_P(I') \). Applying Theorem 2, this means there exists a \( T_{i,a'} \) and \( \delta_{i,a'} \) such that for \( T \geq T_{i,a'} \), \( CBV^\sigma^*_i(I, a') - \sum_{t=1}^{T} v_\sigma^T(I') \leq -\delta_{i,a'} \). So (5) always applies for \( T \geq T_{i,a'} \) for \( I' \) and \( a' \) and \( I \) will always be pruned. Since (8) does not require knowledge of regret, it need not be stored for \( I \).

Since \( D(I', a') \) will always be pruned for \( T \geq T_{i,a'} \), for any \( T \geq (T_{i,a'}^2)^2 \) iterations for some constant \( C > 0 \), \( \pi^T(I) \leq \frac{C}{\sqrt{T}} \), which satisfies the threshold of the average strategy. Thus, the average strategy in \( D(I, a) \) can be discarded. \( \square \)

E. Lemma 2

**Lemma 2.** If for all \( T \geq T' \) iterations of CFR with BRP, \( T(CBV^\sigma^*(I, a)) - \sum_{t=1}^{T} v^\sigma(I) \leq -\epsilon T \) for some \( x > 0 \), then any history \( h' \) such that \( h \cdot a \subseteq h' \) for some \( h \in I \) need only be traversed at most \( O(\ln(T)) \) times. \( \square \)
Proof. Let \( a \in A(I) \) be an action such that for all \( T \geq T' \), \( T(CBV^{\sigma_T^x}(I,a)) - \sum_{t=1}^{T} v^{\sigma_t^x}(I) \leq -xT \) for some \( x > 0 \). \( \psi^{\sigma_{t+1}}_{T_{t+1}}(I,a) \leq CBV^{\bar{\sigma}_{T_{t+1}}^x}, \) so from Theorem 1, \( D(I,a) \) can be pruned for \( m \geq \lceil \frac{x}{U(I,a) - L(I)} \rceil \) iterations on iteration \( T \). Thus, over iterations \( T \leq t \leq T + m \), only a constant number of traversals must be done. So each iteration requires only \( \frac{C}{m} \) work when amortized, where \( C \) is a constant. Since \( x, U(I,a), \) and \( L(I) \) are constants, so on each iteration \( t \geq T' \), only an average of \( \frac{C}{T'} \) traversals of \( D(I,a) \) is required. Summing over all \( t \leq T \) for \( T \geq T' \), and recognizing that \( T' \) is a constant, we get that action \( a \) is only taken \( O(\ln(T)) \) over \( T \) iterations. Thus, any history \( h' \) such that \( h \cdot a \sqsubseteq h' \) for some \( h \in I \) need only be traversed at most \( O(\ln(T)) \) times.

F. Proof of Theorem 3

Proof. Consider an \( h^* \not\in S \). Then there exists some \( h \cdot a \sqsubseteq h^* \) such that \( h \in S \) but \( h \cdot a \not\in S \). Let \( I = I(h) \) and \( i = P(I) \). Since \( h \cdot a \not\in S \) but \( h \in S \), so for every Nash equilibrium \( \sigma^* \), \( CBV^{\sigma^*}(I,a) < CBV^{\sigma^*}(I) \). From Theorem 2, there exists a \( T_{I,a} \) and \( \delta_{I,a} > 0 \) such that after \( T \geq T_{I,a} \) iterations of CFR, \( CBV^{\bar{\sigma}_{T_{I,a}}^x}(I,a) - \sum_{t=1}^{T} v^{\sigma_t^x}(I) \leq -\delta_{I,a}. \) Thus from Lemma 2, \( h^* \) need only be traversed at most \( O(\ln(T)) \) times.