LEMMA2

Jan Mikulik

September 2025

[S2 — Cross-term attenuation under bounded degree] Fix instance–independent constants $\epsilon_{\text{lock}}, \rho_{\text{lock}}, \zeta_0 \in (0,1)$ and an integer $d = O(1)$. Let $L = 3$, $R = \Theta(\log C)$, $T = RL$, and $m = \lfloor \rho_{\text{lock}} T \rfloor$. For each clause $j \in [C]$ construct a deterministic phase schedule $\{\phi_{t,j}\}_{t=1}^{T}$ as follows:

1. **De-aliased offsets.** Choose a prime $p \geq C$ and nonzero $a \in \{1, \ldots, p-1\}$ and $b \in \{0, \ldots, p-1\}$. Set

$$o_j \;=\; \Big\lfloor T \cdot \frac{(aj+b) \bmod p}{p} \Big\rfloor \in \{0, \ldots, T-1\}.$$

   The lock window for clause j is the interval of length m starting at o_j (mod T).

2. **Low-correlation lock masks.** Let $m_{\text{eff}} = 2^{\lceil \log_2 m \rceil}$. Assign to each clause a distinct row $s_j \in \{\pm 1\}^{m_{\text{eff}}}$ of the Walsh–Hadamard matrix (or an explicit small-bias code) and restrict to the first m slots. Enforce the mis-phase fraction ζ_0 by flipping at most $\delta_\zeta m$ positions so that exactly $k = \lfloor \zeta_0 m \rfloor$ slots in the lock are at phase π (entries -1).

3. **Phases.** Inside the lock of clause j, set $e^{i\phi_{t,j}} = +1$ where $s_j(t) = +1$ and $e^{i\phi_{t,j}} = -1$ where $s_j(t) = -1$. Outside the lock, set $e^{i\phi_{t,j}} = -1$.

Let $Z \in C^{T \times C}$ collect $e^{i\phi_{t,j}}$ and let the clause Gram block be $G_{\text{clause}} = 1T\,|Z^* Z| \in R^{C \times C}$. Let H_C be a fixed d-regular wiring on clauses (circulant/expander). Then for every row i,

$$\sum_{j \in \mathcal{N}(i)} \big|G_{\text{clause}}(i,j)\big| \;\leq\; d\,\kappa, \qquad \kappa \;=\; (1-2\zeta_0)^2 \;+\; \varepsilon(m) \;+\; \frac{1}{T}, \tag{1}$$

where the mask–correlation slack satisfies

$$\varepsilon(m) \;\leq\; \frac{1}{\sqrt{m_{\text{eff}}}} \;+\; \frac{2}{m} \;=\; 2^{-\lceil \log_2 m \rceil/2} \;+\; \frac{2}{m}. \tag{2}$$

[Proof sketch] Write the off–diagonal entry as a time average:

$$G_{\text{clause}}(i,j) \;=\; \frac{1}{T}\Big|\sum_{t=1}^{T} \overline{z_{t,i}}\, z_{t,j}\Big|, \qquad z_{t,\ell} = e^{i\phi_{t,\ell}} \in \{\pm 1\}.$$

Partition time into three disjoint regions: lock–lock overlap, one–in–lock, and outside–outside. *(i) Lock–lock:* On the overlap, the product $\overline{z_{t,i}}z_{t,j}$ equals $+1$ on positions where both masks agree, and -1 where they disagree. The Hadamard (or small-bias) property implies the empirical correlation between two distinct rows deviates by at most $1/\sqrt{m_{\text{eff}}}$; enforcing the exact mis-phase $k = \lfloor \zeta_0 m \rfloor$ adds at most $\delta_\zeta \le 1/m$ change per clause, hence a total slack $\le 2/m$. Therefore the lock–lock contribution is upper bounded by $(1-2\zeta_0)^2 + \varepsilon(m)$. *(ii) One–in–lock:* When exactly one clause is in its lock, the other is at -1; the contribution is bounded by the same $\varepsilon(m)$ since the code is balanced and the rounding flips are controlled by δ_ζ. *(iii) Outside–outside:* Both clauses are at -1; over a length-T period the discretization and wrap effects introduce at most $1/T$ slack (absorbed into the last term). Summing the three parts and using $|\mathcal{N}(i)| = d$ yields eq:S2-rowsum.

Numerical instantiation (your run). With $\rho_{\text{lock}} = 0.50$, $\zeta_0 = 0.40$, $C = 1000$, $R = 104$, $T = 312$, we have $m = \lfloor 0.5 \cdot 312 \rfloor = 156$, hence $m_{\text{eff}} = 256$ and

$$(1-2\zeta_0)^2 = (1-0.8)^2 = 0.04, \qquad \varepsilon(m) \le 1\sqrt{256} + 2156 = 0.0625 + 0.01282 \approx 0.07532, \qquad 1T \approx 0.00321.$$

Thus

$$\kappa \;\le\; 0.04 + 0.07532 + 0.00321 \;\approx\; 0.11853.$$

For $d = 4$: $\sum_{j \in \mathcal{N}(i)} |G_{\text{clause}}(i,j)| \le d\kappa \approx 0.4741$. Pro $d = 6$: $d\kappa \approx 0.7112$. V obou případech jsme hluboko pod dřívějšími empirickými $\sim 3.93/5.9$ a s velkou rezervou pro Gershgorinův/row-sum krok (S3).

Remarks on $\varepsilon(m)$ and δ_ζ.

- *Exact mis-phase.* Nastavení $k = \lfloor \zeta_0 m \rfloor$ vyvolá relativní odchylku $\delta_\zeta \le 1/m$ na každé masce; efekt v korelaci dvou masek je $\le 2/m$, proto je zahrnut v eq:epsm.
- *Code choice.* Hadamard dává $\langle s_i, s_j \rangle / m_{\text{eff}} = 0$ (ideálně); po zkrácení na m a přesné ζ_0 dostáváme $\varepsilon(m)$ výše. Alternativně lze použít Legendre/Petrank–Shamir malé-bias kódy se stejným tvarem $\varepsilon(m) = O(1/\sqrt{m})$.
- *Design knob.* Větší m (tj. větší R) snižuje $\varepsilon(m)$ i $1/T$; současně Matrix-Bernstein (A4) zmenší σ a zvětší mezeru Δ_{spec}.

Corollary (A3 in one line). Under the construction above and any fixed $d = O(1)$,

$$\max_i \sum_{j \in \mathcal{N}(i)} |G_{\text{clause}}(i,j)| \;\le\; d\,\kappa \quad \textit{with} \quad \kappa = (1-2\zeta_0)^2 + \varepsilon(m) + 1T,$$

hence the cross-term (row-sum) control required by A3 holds with instance-independent constants.