Improved Lower Bound for Frankl's Union-Closed Sets Conjecture

Ryan Alweiss[a] Brice Huang[b] Mark Sellke[c]

Submitted: Jul 18, 2023; Accepted: Jul 1, 2024; Published: Sep 20, 2024

Abstract

We verify an explicit inequality conjectured in [8], thus proving that for any nonempty union-closed family $\mathcal{F} \subseteq 2^{[n]}$, some $i \in [n]$ is contained in at least a $\frac{3-\sqrt{5}}{2} \approx 0.38$ fraction of the sets in $\mathcal{F}$. One case, an explicit one-variable inequality, is checked by computer calculation. **Mathematics Subject Classifications**: 05D05

Mathematics Subject Classifications: 05D05

1 Introduction

Let $\mathcal{M}_\phi$ be the set of probability measures $\mu \in \mathcal{P}([0,1])$ with expectation ϕ. Define

$$F(\mu) = \mathop{\mathbb{E}}_{(x,y)\sim\mu\times\mu} H(xy) - \mathop{\mathbb{E}}_{x\sim\mu} H(x) \tag{1}$$

where $H(x) = -x\log x - (1-x)\log(1-x)$ is the entropy function and log denotes the natural logarithm. Note that F is continuous in the weak topology and $\mathcal{M}_\phi$ is compact, so F has a minimizer over $\mathcal{M}_\phi$. In this note, we will show the following results.

Theorem 1. *For all $\phi \in [0,1]$, the minimum of $F(\mu)$ over $\mathcal{M}_\phi$ is attained at some μ supported on at most two points. Furthermore, if a minimizer is supported on exactly two points, then one of the points is 0.*

The case of μ supported on $\{0, x\}$ leads to the following definition:

$$S = \left\{\phi \in [0,1] : \phi H(x^2) \geqslant xH(x)\ \forall x \in [\phi, 1]\right\}, \quad \phi^* = \min(S).$$

Note that the condition defining S is monotone in ϕ and S is clearly closed, so $\min(S)$ is well defined. As in the recent breakthrough [8] by Gilmer, a bound on Frankl's union-closed conjecture follows from the above.

[a]Department of Mathematics and Mathematical Statistics, University of Cambridge, UK (ryeguy10@gmail.com)
[b]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, USA (b.huang97@gmail.com).
[c]Department of Statistics, Harvard University, USA (msellke@gmail.com).

Theorem 2. *The union-closed conjecture holds with constant $1-\phi^*$, i.e. for any non-empty union-closed family $\mathcal{F} \subseteq 2^{[n]}$, some $i \in [n]$ is contained in at least $1-\phi^*$ fraction of the sets in $\mathcal{F}$.*

Throughout this paper we set $\varphi = \frac{\sqrt{5}-1}{2}$. In the Appendix, we give a numerical verification of the following claim. We require certain computer calculations (detailed in an attached Python file) to be accurate to within margin of error 10^{-3}, which can be made completely rigorous using interval arithmetic.

Claim 3. *If $x \in [\varphi, 1]$, then $\varphi H(x^2) \geqslant xH(x)$, with equality if and only if $x \in \{\varphi, 1\}$.*

Assuming Claim 3, the following claim identifies the value of ϕ^*. Then, Theorem 2 implies that the union-closed conjecture holds with constant $1-\varphi = \frac{3-\sqrt{5}}{2}$. This is a natural barrier for the method of [8] as explained therein. Interestingly, Claim 3 has been mentioned previously in a different context by [3].

Claim 4. *We have that $\phi^* = \varphi$.*

Related Work. The union-closed conjecture has been the subject of much study, see [1, 10, 14, 2, 9] or the survey [4]. The recent breakthrough [8] by Gilmer showed that this conjecture holds with constant 0.01.

Concurrently with and independently of this work, Chase and Lovett [6], Sawin [12], and Pebody [11] also proved the union-closed conjecture with constant $\frac{3-\sqrt{5}}{2}$. [12] also outlined an argument to improve this bound by an additional small constant, which was subsequently made explicit in [15] (using Lemma 5 below) and [5]. Moreover, [12] and Ellis [7] found counterexamples to [8, Conjecture 1], which would have implied the full union-closed conjecture with constant $\frac{1}{2}$.

Acknowledgements. We thank Zachary Chase and Shachar Lovett for sharing their writeup [6] with us. We thank Mehtaab Sawhney and the anonymous referee for helpful comments. RA was supported by an NSF Mathematical Sciences Postdoctoral Research Fellowship. BH was supported by an NSF Graduate Research Fellowship, a Siebel scholarship, NSF awards CCF-1940205 and DMS-1940092, and NSF-Simons collaboration grant DMS-2031883.

2 Reduction to Two Point Masses

Lemma 5. *F is concave on $\mathcal{M}_\phi$ for any $\phi \in [0,1]$, i.e.*

$$pF(\mu_1) + (1-p)F(\mu_2) \leqslant F(p\mu_1 + (1-p)\mu_2) \quad \forall\ \mu_1, \mu_2 \in \mathcal{M}_\phi,\ p \in [0,1]\,. \tag{2}$$

Proof. Let $\gamma(x) = \mu([0,x])$ be the cumulative distribution function of μ. Thus $\gamma(1) = 1$ and

$$\phi = \int_0^1 x\mu(\mathrm{d}x) = 1 - \int_0^1 \gamma(x)\ \mathrm{d}x\,,$$

so

$$\int_0^1 \gamma(x)\ \mathrm{d}x = 1-\phi\,. \tag{3}$$

Using integration by parts,

$$\int_0^1 H(x)\mu(\mathrm{d}x) = H(x)\gamma(x)\Big|_0^1 - \int_0^1 H'(x)\gamma(x)\ \mathrm{d}x = \int_0^1 \left(\log\frac{x}{1-x}\right)\gamma(x)\ \mathrm{d}x\,.$$

Similarly,

$$\begin{aligned}
\int_0^1 H(xy)\mu(\mathrm{d}y) &= H(xy)\gamma(y)\Big|_0^1 - \int_0^1 xH'(xy)\gamma(y)\ \mathrm{d}y \\
&= H(x) + \int_0^1 \left(x\log\frac{xy}{1-xy}\right)\gamma(y)\ \mathrm{d}y\,; \\
\int_0^1 \left(x\log\frac{xy}{1-xy}\right)\mu(\mathrm{d}x) &= \left(x\log\frac{xy}{1-xy}\right)\gamma(x)\Big|_0^1 - \int_0^1 \frac{\mathrm{d}}{\mathrm{d}x}\left(x\log\frac{xy}{1-xy}\right)\gamma(x)\ \mathrm{d}x\,, \\
&= \log\frac{y}{1-y} - \int_0^1 \left(\frac{1}{1-xy} + \log\frac{xy}{1-xy}\right)\gamma(x)\ \mathrm{d}x\,; \\
\iint_{[0,1]^2} H(xy)\mu(\mathrm{d}x)\mu(\mathrm{d}y) &= \int_0^1 H(x)\mu(\mathrm{d}x) + \int_0^1 \gamma(y)\int_0^1 x\log\frac{xy}{1-xy}\mu(\mathrm{d}x)\ \mathrm{d}y\,, \\
&= 2\int_0^1 \left(\log\frac{x}{1-x}\right)\gamma(x)\ \mathrm{d}x \\
&- \iint_{[0,1]^2} \left(\frac{1}{1-xy} + \log\frac{xy}{1-xy}\right)\gamma(x)\gamma(y)\mathrm{d}x\mathrm{d}y.
\end{aligned}$$

So, letting $F(\gamma) = F(\mu)$ by slight abuse of notation, we have

$$\begin{aligned}
F(\gamma) = &\int_0^1 \left(\log\frac{x}{1-x}\right)\gamma(x)\ \mathrm{d}x \\
&- \iint_{[0,1]^2} \left(\log x + \log y + \frac{1}{1-xy} + \log\frac{1}{1-xy}\right)\gamma(x)\gamma(y)\ \mathrm{d}x\ \mathrm{d}y\,.
\end{aligned}$$

We will show this is concave in γ. The first integral is manifestly linear in γ, and the contributions of $\log x$ and $\log y$ are linear because, in light of (3),

$$\iint_{[0,1]^2} (\log x)\gamma(x)\gamma(y)\ \mathrm{d}x\ \mathrm{d}y = (1-\phi)\int_0^1 (\log x)\gamma(x)\ \mathrm{d}x\,.$$

After removing these terms, we are reduced to showing convexity of

$$\iint_{[0,1]^2} \left(\frac{1}{1-xy} + \log\frac{1}{1-xy}\right)\gamma(x)\gamma(y)\mathrm{d}x\mathrm{d}y\,.$$

Note that both $\frac{1}{1-xy}$ and $\log \frac{1}{1-xy}$ are of the form $\sum_{k\geqslant 0} a_k x^k y^k$ for constants $a_k \geqslant 0$. Hence it suffices to prove convexity of

$$\iint_{[0,1]^2} x^k y^k \gamma(x)\gamma(y) \mathrm{d}x\mathrm{d}y = \left(\int_0^1 x^k \gamma(x) \mathrm{d}x\right)^2$$

for any $k \geqslant 0$. This is the square of a linear function of γ, and hence is convex. (Note that all integrands are in L^1 and so there are no convergence issues.) $\square$

Lemma 6. $\arg\min_{\mu\in\mathcal{M}_\phi} F(\mu)$ *contains some* μ *supported on at most two points.*

Proof. This follows immediately from Lemma 5 and the Krein-Milman theorem since $\mathcal{M}_\phi$ is compact in the weak-* topology and convex, and all extreme measures in $\mathcal{M}_\phi$ are supported on 1 or 2 points (see e.g. [13] for more on the latter point).

For self-containedness, we also include an explicit and elementary version of this argument. First let $\mu \in \mathcal{M}_\phi$ be any minimizer of F and note that μ can be approximated arbitrarily well in the weak topology by $\hat{\mu}$ with finite support. In particular for any $\varepsilon > 0$, there exists $\hat{\mu} \in \mathcal{M}_\phi$ with $F(\hat{\mu}) \leqslant F(\mu) + \varepsilon$ of the form

$$\hat{\mu}(a_i) = b_i - b_{i-1}, \quad 1 \leqslant i \leqslant k$$

for constants $0 \leqslant a_1 < \cdots < a_k \leqslant 1$ and $0 = b_0 < b_1 < \cdots < b_k = 1$. We claim that for any $\varepsilon > 0$, the minimal k such that such a $\hat{\mu}$ exists is at most two. Indeed given such a $\hat{\mu}$ with $k \geqslant 3$, we may consider $\hat{\mu}_\eta$ defined by

$$\begin{aligned}
\hat{\mu}_\eta(a_1) &= b_1 - b_0 + \eta(a_3 - a_2),\\
\hat{\mu}_\eta(a_2) &= b_2 - b_1 - \eta(a_3 - a_1),\\
\hat{\mu}_\eta(a_3) &= b_3 - b_2 + \eta(a_2 - a_1),\\
\hat{\mu}_\eta(a_i) &= \hat{\mu}(a_i) = b_i - b_{i-1}, \quad \forall\, i \in \{4, 5, \ldots, k\}.
\end{aligned}$$

Then there exist $c_1, c_2 > 0$ such that $\hat{\mu}_\eta \in \mathcal{M}_\phi$ if and only if $-c_1 \leqslant \eta \leqslant c_2$; moreover the map $\eta \mapsto F(\hat{\mu}_\eta)$ is concave by Lemma 5. It is easy to see that both $\hat{\mu}_{-c_1}, \hat{\mu}_{c_2}$ have support size at most $k-1$, and at least one of $F(\hat{\mu}_{-c_1}), F(\hat{\mu}_{c_2})$ is at most $F(\hat{\mu})$ by concavity. Iterating this argument, we find a $\tilde{\mu} \in \mathcal{M}_\phi$ with support size at most 2 and with $F(\tilde{\mu}) \leqslant F(\hat{\mu}) \leqslant F(\mu) + \varepsilon$. Taking a subsequential weak limit of the resulting $\tilde{\mu}$ as $\varepsilon \to 0$ completes the proof. $\square$

3 Optimization over Two Point Masses

Lemma 7. *If* μ *is supported on exactly two points, neither of which is* 0*, then* μ *is not a minimizer of* F *over* $\mathcal{M}_\phi$.

Proof. Suppose $\mu = p\delta_x + (1-p)\delta_y$ is a minimizer for F over $\mathcal{M}_\phi$ for $0 < y < x < 1$ distinct and $0 < p < 1$. Then any $z \in [0,1]$ can be written as $z = qx + (1-q)y$ for some $q \in \mathbb{R}$ (which may be negative). We have

$$\mu + t\delta_z - tq\delta_x - t(1-q)\delta_y \in \mathcal{M}_\phi$$

for sufficiently small $t \geqslant 0$ and so

$$\lim_{t \to 0^+} \frac{F\big(\mu + t\delta_z - tq\delta_x - t(1-q)\delta_y\big) - F(\mu)}{t} \geqslant 0.$$

It is not difficult to see from the definition (1) of F that the left-hand limit equals

$$f(z) - qf(x) - (1-q)f(y) \geqslant 0\,, \tag{4}$$

for

$$f(w) := 2[pH(xw) + (1-p)H(yw)] - H(w).$$

Equation (4) implies that f lies above the line passing through $(x, f(x))$ and $(y, f(y))$. Since f is a smooth function and x, y are in the interior of $[0,1]$, we deduce that

(a) $f'(x) = f'(y) = \frac{f(x)-f(y)}{x-y}$, and

(b) $f''(x), f''(y) \geqslant 0$.

Moreover, (a) implies

(c) $f''(z) \leqslant 0$ for some $z \in [y, x]$.

However we compute using $H'(w) = \log \frac{1-w}{w}$ that:

$$f'(z) = 2\left[px \log \frac{1-xz}{xz} + (1-p)y \log \frac{1-yz}{yz}\right] - \log \frac{1-z}{z}\,,$$
$$f''(z) = -2\left[\frac{px}{z(1-xz)} + \frac{(1-p)y}{z(1-yz)}\right] + \frac{1}{z(1-z)}\,.$$

Note that $g(z) := z(1-z)(1-xz)(1-yz)f''(z)$ has the same sign as $f''(z)$ and is a quadratic function in z with leading coefficient

$$-2pxy - 2(1-p)xy + xy = -xy < 0\,.$$

Hence the inequalities $g(x), g(y) \geqslant 0$ and $g(z) \leqslant 0$ can hold only if g and hence f'' vanishes on the entire interval $[x, y]$. This is impossible since we just saw g has non-zero leading coefficient.

The case $x = 1$, $y > 0$ is very similar. While we have $f''(y) \geqslant 0$ as above, since 1 is not in the interior of $[0,1]$, we cannot immediately deduce that $f''(1) \geqslant 0$. However in this case $g(z)$ is a multiple of $1-z$, and so $g(1) = 0 \geqslant 0$. Then the same argument applies: $g(z)$ is a quadratic polynomial with negative leading coefficient $-y < 0$. Because g takes non-negative values at y and 1, it takes positive values in between. However since f is continuous on $[0,1]$ and smooth on $(0,1)$, and stays above the line segment through $(y, f(y))$ and $(1, f(1))$, it must have non-positive second derivative at some $z \in (y, 1)$. Since g and f'' have the same sign on $(0,1)$, this is a contradiction. (Note that $f''(1)$ does not actually exist if $x = 1$ and is not used in this argument.) $\square$

4 Conclusion

Proof of Theorem 1. Follows from Lemmas 6 and 7. □

Lemma 8. *We have that $\phi^* \geqslant \varphi$.*

Proof. Note that $H(\varphi^2) = H(\varphi)$. If $\phi < \varphi$, then $\phi H(\varphi^2) < \varphi H(\varphi)$, and so $\phi \notin S$. □

Corollary 9. *If $\phi \geqslant \phi^*$, then $F(\mu) \geqslant 0$ for all $\mu \in \mathcal{M}_\phi$.*

Proof. By Theorem 1, it suffices to check $F(\mu) \geqslant 0$ for $\mu = p\delta_x + (1-p)\delta_0$ with $p = \phi/x$ and $x \in [\phi, 1]$ (this includes the case $\mu = \delta_\phi$, corresponding to $x = \phi$). By monotonicity of the condition defining S, $\phi \in S$. So,

$$F(\mu) = \frac{\phi^2}{x^2}H(x^2) - \frac{\phi}{x}H(x) = \frac{\phi}{x^2}(\phi H(x^2) - xH(x)) \geqslant 0\,.$$ □

From Theorem 1, we deduce the following tight version of Gilmer's [8, Lemma 1]. Theorem 2 follows from Corollary 10 by the same argument as in [8, Proof of Theorem 1]. We recall Gilmer's ingenious insight was that given a union-closed family $\mathcal{F} \subseteq 2^{[n]}$, if A, A' are independent uniformly random samples from $\mathcal{F}$, then $A \cup A' \in \mathcal{F}$ is *not* uniformly random and thus has strictly smaller entropy. On the other hand, Corollary 10 can be applied element-by-element to show that $A \cup A'$ actually has equal or larger entropy.

Corollary 10. *Suppose $\{p_c\}_{c\in\mathcal{S}} \subset [0,1]$ is a finite sequence of real numbers and c is a random variable supported on $\mathcal{S}$ such that $\mathbb{E}_c[p_c] \leqslant 1 - \phi^*$. If c' is an independent copy of c, then*

$$\mathop{\mathbb{E}}_{c,c'}[H(p_c + p_{c'} - p_c p_{c'})] \geqslant \mathop{\mathbb{E}}_{c}[H(p_c)]\,.$$

Proof. Let μ be the distribution of $x = 1 - p_c$. Let $\phi = \mathbb{E}_{x\sim\mu}[x]$, so $\phi > \phi^*$. By Corollary 9,

$$\mathop{\mathbb{E}}_{c,c'}[H(p_c + p_{c'} - p_c p_{c'})] - \mathop{\mathbb{E}}_{c}[H(p_c)] = \mathop{\mathbb{E}}_{(x,y)\sim\mu\times\mu} H(xy) - \mathop{\mathbb{E}}_{x\sim\mu} H(x) = F(\mu) \geqslant 0\,.$$ □

Finally, we verify Claim 4 assuming Claim 3.

Proof of Claim 4. Claim 3 implies $\varphi \in S$, so $\phi^* \leqslant \varphi$ by definition of ϕ^*. On the other hand, Lemma 8 gives $\phi^* \geqslant \varphi$. □

References

[1] Polymath 11. `https://gowers.wordpress.com/2016/01/21/frankls-union-closed-conjecture-a-possible-polymath-project/`.

[2] Balla, I., Bollobás, B., and Eccles, T. Union-closed families of sets. *Journal of Combinatorial Theory, Series A 120*, 3 (2013), 531–544.

[3] Boppana, R. B. Amplification of probabilistic Boolean formulas. In *26th Annual Symposium on Foundations of Computer Science (SFCS 1985)* (1985), IEEE, pp. 20–29.

[4] Bruhn, H., and Schaudt, O. The journey of the union-closed sets conjecture. *Graphs and Combinatorics 31*, 6 (2015), 2043–2074.

[5] Cambie, S. Better bounds for the union-closed sets conjecture using the entropy approach. arXiv:2212.12500 (2022).

[6] Chase, Z., and Lovett, S. Approximate union closed conjecture. arXiv:2211.11689 (2022).

[7] Ellis, D. Note: a counterexample to a conjecture of Gilmer which would imply the union-closed conjecture. arXiv:2211.12401 (2022).

[8] Gilmer, J. A constant lower bound for the union-closed sets conjecture. arXiv:2211.09055 (2022).

[9] Karpas, I. Two results on union-closed families. arXiv:1708.01434 (2017).

[10] Knill, E. Graph generated union-closed families of sets. arXiv:math/9409215 (1994).

[11] Pebody, L. Extension of a Method of Gilmer. arXiv:2211.13139 (2022).

[12] Sawin, W. An improved lower bound for the union-closed set conjecture. arXiv:2211.11504 (2022).

[13] Winkler, G. Extreme points of moment sets. *Mathematics of Operations Research 13*, 4 (1988), 581–587.

[14] Wójcik, P. Union-closed families of sets. *Discrete Mathematics 199*, 1-3 (1999), 173–182.

[15] Yu, L. Dimension-free bounds for the union-closed sets conjecture. *Entropy 25*, 5 (2023), 767.

A Proof of Claim 3

In this appendix, we prove Claim 3. Throughout this appendix, we use Claims to indicate results requiring the correctness of computer outputs within margin of error 10^{-3} or greater. The only computations which rely on a computer are the entries in Tables 1 and 2. Figure 1 plots the function

$$G(x) = \varphi H(x^2) - xH(x),$$

from which Claim 3 can be checked visually. We show below that, assuming correctness of certain computer calculations to within margin of error 10^{-3},

$$G(x) \geqslant 0, \quad \forall x \in [\varphi, 1].$$

We verify this separately on the intervals $I_1 = [\varphi, 0.77], I_2 = [0.76, 0.98], I_3 = [0.98, 1]$.

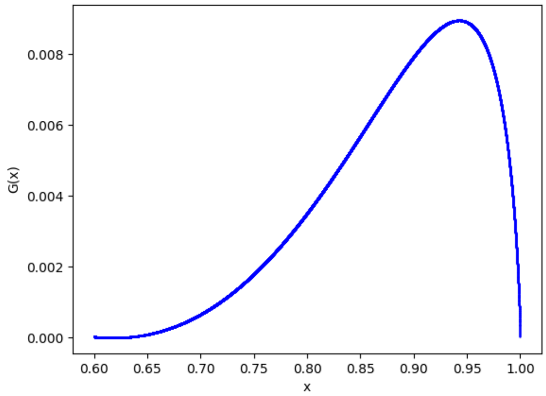

Figure 1: Plot of $G(x)$ for $x \in [0.6, 1]$. Claim 3 states the minimum value of 0 on $x \in [\varphi, 1]$ is achieved precisely at the endpoints $x \in \{\varphi, 1\}$.

A.1 Verification on I_1

We first compute the derivative of G:

$$\begin{aligned} G'(x) =& 2x\varphi \log \frac{1-x^2}{x^2} - H(x) - x \log \frac{1-x}{x} \\ &= 2x\varphi \log \frac{1-x^2}{x^2} + x\log x + (1-x)\log(1-x) + x \log x - x\log(1-x) \\ &= 2x\varphi \log \frac{1-x^2}{x^2} + 2x\log x + (1-2x)\log(1-x) \end{aligned}$$

Note that $G(\varphi) = G'(\varphi) = 0$, the latter since

$$\begin{aligned} G'(\varphi) &= 2\varphi^2 \log(1/\varphi) + 2\varphi \log(\varphi) + (1-2\varphi)\log(\varphi^2) \\ &= (-2\varphi^2 + 2\varphi + 2(1-2\varphi))\log\varphi \\ &= 2(1-\varphi-\varphi^2)\log(\varphi) = 0. \end{aligned}$$

Claim 11. *Claim 3 holds on $I_1 = [\varphi, 0.77]$.*

Proof. As $G(\varphi) = G'(\varphi) = 0$, it suffices to verify that G is convex on I_1. It is not hard to check that its second derivative equals $G''(x) = L(x)/(1-x^2)$, where

$$L(x) := 2\varphi(1-x^2)\log(x^{-2}-1) - 4\varphi - 2x^2\log x + 2(x^2-1)\log(1-x) + x + 2\log(x) + 1\,.$$

We now estimate the Lipschitz constant of each non-constant term of L on $x \in I_1$. For the first term,

$$\begin{aligned} \left|\frac{d}{dx}\big(2\varphi(1-x^2)\log(x^{-2}-1)\big)\right| &\leqslant 2\varphi \sup_{x\in I_1}\big(|2x^3| + 2|x\log(x^{-2}-1)|\big) \\ &\leqslant 2\varphi(1.1 + 1.6\cdot\log(2)) \\ &\leqslant 2\varphi \cdot 2.3 \leqslant 3 \end{aligned} \tag{5}$$

since $\log(2) \leqslant 0.75$ and $\varphi \leqslant 5/8$. Next,

$$\begin{aligned}\left|\frac{d}{dx}\big(2x^2\log(x)\big)\right| &\leqslant \sup_{x\in I_1}|4x\log(x)+2x| \\ &\leqslant 1.6\sup_{x\in I_1}|2\log(x)+1| \\ &\leqslant 1.6\end{aligned}$$

since $\log(x) \in [-1,0]$ for all $x \in I_1$. Continuing, using $\log(5) \leqslant 2$,

$$\begin{aligned}\left|\frac{d}{dx}\big(2(x^2-1)\log(1-x)\big)\right| &\leqslant 2\sup_{x\in I_1}|2x\log(1-x)-\frac{x^2-1}{1-x}| \\ &\leqslant 2\sup_{x\in I_1}|2x\log(1-x)+x+1| \\ &\leqslant 2\cdot\max(1.6\log(5),1.8) \\ &\leqslant 2\cdot 1.6\cdot 2 = 6.4.\end{aligned}$$

Finally $\frac{d}{dx}(x) = 1$ and $\frac{d}{dx}(2\log x) = 2/x \leqslant 3.5$. Combining, we find that $L(x)$ restricted to I_1 has Lipschitz constant at most

$$1.6 + 6.4 + 1 + 3 + 3.5 \leqslant 15.5.$$

Therefore to show G is convex and hence non-negative on $I_1 = [\varphi, 0.77]$ it suffices to exhibit a $\frac{1}{400}$-dense subset of I_1 on which $L(x) = (1-x^2)G''(x) \geqslant 0.04 \geqslant \frac{15.5}{400}$. In Table 1 below we compute the values of L on each multiple of $\frac{1}{200}$ from 0.6 to 0.77 inclusive. We find that $L(x) \geqslant 0.09$ holds at all of these points, completing the numerical verification on I_1. □

x	$L(x)$	x	$L(x)$	x	$L(x)$	x	$L(x)$	x	$L(x)$	x	$L(x)$
0.600	0.1020	0.630	0.1117	0.660	0.1173	0.690	0.1182	0.720	0.1137	0.750	0.1032
0.605	0.1039	0.635	0.1130	0.665	0.1178	0.695	0.1178	0.725	0.1124	0.755	0.1009
0.610	0.1057	0.640	0.1141	0.670	0.1182	0.700	0.1173	0.730	0.1109	0.760	0.0983
0.615	0.1074	0.645	0.1151	0.675	0.1184	0.705	0.1167	0.735	0.1093	0.765	0.0955
0.620	0.1089	0.650	0.1159	0.680	0.1185	0.710	0.1159	0.740	0.1075	0.770	0.0925
0.625	0.1104	0.655	0.1167	0.685	0.1184	0.715	0.1149	0.745	0.1054		

Table 1: Evaluations of L to precision 10^{-4}. All values appear to be at least 0.09, and it suffices for all values to be at least 0.04.

A.2 Verification on I_2

Our verification for $x \in I_2$ is based on evaluating G. We write $G(x) = g_1(x) - g_2(x)$ for

$$\begin{aligned}g_1(x) &= \varphi H(x^2), \\ g_2(x) &= xH(x).\end{aligned}$$

Note that g_1 is clearly decreasing on I_2. The next lemma shows the same for g_2.

Lemma 12. *g_2 is decreasing on $[5/7,1] \supseteq I_2$.*

Proof. First we claim that it suffices to show $g_2'(5/7) \leqslant 0$. This is because

$$g_2'(x) = H(x) + x\log\frac{1-x}{x} = 2x\log\frac{1}{x} - (2x-1)\log\frac{1}{1-x}$$

so $g_2'(x) \leqslant 0$ if and only if

$$\left(1 - \frac{1}{2x}\right)\log\frac{1}{1-x} \geqslant \log\frac{1}{x}. \tag{6}$$

Indeed both terms on the left-hand side are increasing while the right-hand side is decreasing.

It remains to show that $g_2'(5/7) \leqslant 0$ which in light of (6) is equivalent to showing

$$\frac{3}{10}\log(7/2) \geqslant \log(7/5),$$

i.e. $(7/5)^{10/3} \leqslant 7/2$. This holds since $(7/5)^3 \leqslant 2(7/5) = 14/5$ and $7/5 \leqslant \left(\frac{5}{4}\right)^3 = \left(\frac{7/2}{14/5}\right)^3$. ☐

Claim 13. *Claim 3 holds for $x \in I_2$.*

Proof. We computer-evaluate g_1, g_2 at a finite set of values $x_1 < x_2 < \cdots < x_{97}$ with $5/7 < x_1 < 0.76$ and $x_{97} = 0.98$ and verify that $g_1(x_{i+1}) \geqslant g_2(x_i)$ for each i. The values are shown in Table 2; note that in all cases $g_1(x_{i+1}) - g_2(x_i) \geqslant \frac{2}{1000}$ holds, modulo rounding to four decimal places. The intervals $[x_i, x_{i+1}]$ cover I_2, and for all $x \in [x_i, x_{i+1}]$ we have

$$g_2(x) \leqslant g_2(x_i) \leqslant g_1(x_{i+1}) \leqslant g_1(x)\,.$$ ☐

A.3 Verification on I_3

Proposition 14. *Claim 3 holds for $x \in I_3$.*

Proof. Taylor expansion of $\log(1-\varepsilon)$ gives that for all $\varepsilon \in (0,1)$,

$$\varepsilon\left(\log\frac{1}{\varepsilon} + 1 - \varepsilon\right) \leqslant H(\varepsilon) \leqslant \varepsilon\left(\log\frac{1}{\varepsilon} + 1\right).$$

Let $x = 1-\varepsilon$ for $\varepsilon \in [0, 0.02]$. Then

$$g_1(x) = \varphi H(2\varepsilon - \varepsilon^2) \geqslant \varphi\varepsilon(2-\varepsilon)\big(\log\frac{1}{\varepsilon} - \log(2-\varepsilon) + (1-\varepsilon)^2\big)\,,$$
$$g_2(x) = (1-\varepsilon)H(\varepsilon) \leqslant \varepsilon(1-\varepsilon)\big(\log\frac{1}{\varepsilon} + 1\big)\,.$$

x	$g_1(x)$	$g_2(x)$	x	$g_1(x)$	$g_2(x)$	x	$g_1(x)$	$g_2(x)$	x	$g_1(x)$	$g_2(x)$
0.7598	0.4210	0.4189	0.7797	0.4139	0.4111	0.8472	0.3678	0.3622	0.9350	0.2338	0.2249
0.7600	0.4209	0.4188	0.7814	0.4131	0.4103	0.8507	0.3643	0.3586	0.9380	0.2270	0.2180
0.7603	0.4208	0.4187	0.7832	0.4124	0.4095	0.8543	0.3606	0.3547	0.9409	0.2202	0.2112
0.7606	0.4207	0.4186	0.7851	0.4115	0.4085	0.8579	0.3567	0.3507	0.9437	0.2134	0.2045
0.7609	0.4206	0.4185	0.7871	0.4106	0.4075	0.8615	0.3528	0.3465	0.9465	0.2065	0.1975
0.7613	0.4205	0.4184	0.7892	0.4095	0.4064	0.8651	0.3486	0.3422	0.9492	0.1996	0.1907
0.7617	0.4204	0.4183	0.7913	0.4085	0.4053	0.8688	0.3442	0.3377	0.9518	0.1927	0.1839
0.7621	0.4203	0.4181	0.7935	0.4074	0.4041	0.8725	0.3397	0.3330	0.9543	0.1860	0.1772
0.7626	0.4201	0.4180	0.7958	0.4062	0.4028	0.8762	0.3350	0.3281	0.9567	0.1793	0.1706
0.7631	0.4200	0.4178	0.7982	0.4048	0.4014	0.8799	0.3301	0.3230	0.9590	0.1728	0.1641
0.7637	0.4198	0.4176	0.8007	0.4034	0.3999	0.8836	0.3251	0.3178	0.9612	0.1663	0.1577
0.7643	0.4196	0.4174	0.8033	0.4019	0.3983	0.8873	0.3198	0.3124	0.9633	0.1600	0.1515
0.7650	0.4194	0.4171	0.8060	0.4003	0.3965	0.8909	0.3146	0.3070	0.9654	0.1535	0.1452
0.7657	0.4191	0.4169	0.8088	0.3985	0.3947	0.8945	0.3092	0.3014	0.9674	0.1472	0.1390
0.7665	0.4189	0.4166	0.8116	0.3967	0.3927	0.8981	0.3035	0.2957	0.9693	0.1411	0.1330
0.7673	0.4186	0.4163	0.8145	0.3948	0.3907	0.9017	0.2977	0.2897	0.9711	0.1351	0.1271
0.7682	0.4183	0.4159	0.8175	0.3927	0.3884	0.9052	0.2919	0.2838	0.9728	0.1293	0.1215
0.7692	0.4179	0.4156	0.8206	0.3904	0.3861	0.9087	0.2859	0.2776	0.9744	0.1237	0.1160
0.7702	0.4176	0.4152	0.8237	0.3881	0.3836	0.9122	0.2797	0.2713	0.9759	0.1183	0.1109
0.7713	0.4172	0.4147	0.8269	0.3857	0.3810	0.9156	0.2734	0.2650	0.9773	0.1132	0.1059
0.7725	0.4167	0.4142	0.8301	0.3831	0.3783	0.9190	0.2670	0.2584	0.9787	0.1080	0.1009
0.7738	0.4163	0.4137	0.8334	0.3803	0.3754	0.9223	0.2606	0.2519	0.9800	0.1030	0.0961
0.7752	0.4157	0.4131	0.8368	0.3774	0.3723	0.9256	0.2539	0.2452			
0.7766	0.4152	0.4125	0.8402	0.3744	0.3691	0.9288	0.2473	0.2385			
0.7781	0.4145	0.4119	0.8437	0.3711	0.3657	0.9319	0.2406	0.2318			

Table 2: Evaluations of g_1 and g_2 to precision 10^{-4}. We require that for consecutive inputs $x_i < x_{i+1}$ in the table, $g_1(x_{i+1}) - g_2(x_i) \geqslant 0$. The values shown in fact satisfy $g_1(x_{i+1}) - g_2(x_i) \geqslant \frac{2}{1000}$ modulo rounding.

Dividing by ε, it suffices to prove

$$\big((2\varphi-1)+(1-\varphi)\varepsilon\big)\log\frac{1}{\varepsilon} \geqslant (1-\varepsilon)\left(1-\varphi(1-\varepsilon)(2-\varepsilon)\right)+\varphi(2-\varepsilon)\log(2-\varepsilon).$$

Noting $\varphi(1-\varepsilon)(2-\varepsilon) \geqslant 1$ in the first line below, we next find

$$\begin{aligned}(1-\varepsilon)\left(1-\varphi(1-\varepsilon)(2-\varepsilon)\right)+\varphi(2-\varepsilon)\log(2-\varepsilon) &\leqslant 2\varphi\log 2 = (\sqrt{5}-1)\log 2,\\ ((2\varphi-1)+(1-\varphi)\varepsilon)\log\frac{1}{\varepsilon} \geqslant (2\varphi-1)\log\frac{1}{\varepsilon} &\geqslant (\sqrt{5}-2)\log 50.\end{aligned}$$

Finally $(\sqrt{5}-2)\log 50 \geqslant (\sqrt{5}-1)\log 2$ because

$$\begin{aligned}\log_2(50) &\geqslant \log_2(2^5\cdot 1.5) \geqslant 5.5\\ &\geqslant 3+\sqrt{5} = (\sqrt{5}-1)/(\sqrt{5}-2).\end{aligned}$$

Hence the proof is complete. Equality holds if and only if $\varepsilon = 0$, i.e. $x = 1$. □

Proof of Claim 3. Follows by combining Claims 11, 13 and Proposition 14. □