I think I can correct #578 and put it together with the Fourier-based calculations. The following alternate argument for Sperner still feels “wrong” to me (I will say why at the end), but I can’t see the “right” one yet so I’ll write this one up for now.

Let $(x_1, y_1)$ be chosen jointly from $\{0, 1\} \times \{0, 1\}$ as follows (I use boldface to denote random variables): The bit $x_1$ is 0 or 1 with probability 1/2 each. If $x_1 = 1$ then $y_1 = 1$. If $x_1 = 0$ then $y_1 = 0$ with probability $1 - \epsilon$ and $y_1 = 1$ with probability $\epsilon$. Finally, let $(x, y)$ be chosen jointly from $\{0, 1\}^n \times \{0, 1\}^n$ by using the single-bit distribution $n$ times independently. Note that $x \leq y$ always. This distribution is precisely the one Tim uses, with $p = 1/2$, $q = 1/2 + \epsilon/2$.

Let $f : \{0, 1\}^n \rightarrow [-1, 1]$ and let’s consider $E[f(x)f(y)]$. (We are eventually interested in $f$’s with range $\{0, 1\}$ and mean $\delta$, but let’s leave it slightly more general for now.) As per Tim’s calculations,

$$E[f(x)f(y)] = \sum_{S \subseteq [n]} \hat{f}(S)\hat{f}_\epsilon(S)(1 - \epsilon')^{|S|},$$

where $\hat{f}_\epsilon$ denotes Fourier coefficients with respect to the $(1/2 + \epsilon/2)$-biased measure and $\epsilon'$ is defined by $1 - \epsilon' = \sqrt{(1 - \epsilon)/(1 + \epsilon)}$ (AKA $\lambda_{1/2, 1/2 + \epsilon/2}$). Separate out the $S = \emptyset$ term here and use Cauchy-Schwarz on the rest to conclude

$$\left| E[f(x)f(y)] - \hat{f}(\emptyset)\hat{f}_\epsilon(\emptyset) \right| \leq \sqrt{\sum_{|S| \geq 1} \hat{f}(S)^2(1 - \epsilon')^{|S|}} \sqrt{\sum_{|S| \geq 1} \hat{f}_\epsilon(S)^2(1 - \epsilon')^{|S|}}. \quad (1)$$

Let’s compare $\hat{f}_\epsilon(\emptyset)$ to $\hat{f}(\emptyset)$. Write $\pi$ and $\pi_\epsilon$ for the density functions of $x$, $y$ respectively. Then

$$\hat{f}_\epsilon(\emptyset) - \hat{f}(\emptyset) = E[f(y)] - E[f(x)] = E\left[ \frac{\pi_\epsilon(x)}{\pi(x)} f(x) \right] - E[f(x)] = E\left[ \left( \frac{\pi_\epsilon(x)}{\pi(x)} - 1 \right) f(x) \right].$$

By Cauchy-Schwarz, the absolute value of this is upper-bounded by

$$\sqrt{E\left[ \left( \frac{\pi_\epsilon(x)}{\pi(x)} - 1 \right)^2 \right]} \cdot \|f\|_2,$$

where $\|f\|_2$ denotes $\sqrt{E[f(x)^2]}$. One easily checks that

$$E\left[ \left( \frac{\pi_\epsilon(x)}{\pi(x)} - 1 \right)^2 \right] = E\left[ \frac{\pi_\epsilon(x)^2}{\pi(x)^2} \right] - 1,$$  \quad (2)

and since $\pi_\epsilon$ and $\pi$ are product distributions the RHS of (2) is easy to compute. One can check explicitly that

$$E[\pi_\epsilon(x_1)^2/\pi(x_1)^2] = 1 + \epsilon^2$$
and therefore (2) is $(1 + \epsilon^2)^n - 1$. Naturally we will be considering $\epsilon \ll 1/\sqrt{n}$, and in this regime the quantity is bounded by, say, $4\epsilon^2 n$. Hence we have shown

$$|\hat{f}_\epsilon(\emptyset) - \hat{f}(\emptyset)| \leq 2\epsilon \sqrt{n} \cdot \|f\|_2$$

and in particular if $f$ has range $\{0, 1\}$ and mean $\mu$ (AKA $\delta$) then

$$|\hat{f}_\epsilon(\emptyset) - \mu| \leq 2\epsilon \sqrt{n} \cdot \sqrt{\mu}.$$  (3)

This is not very interesting unless $2\epsilon \sqrt{n} \cdot \sqrt{\mu} \leq \mu$, so let’s indeed assume $\epsilon \leq \sqrt{\mu}/\sqrt{n}$ and then we can also use $\hat{f}_\epsilon(\emptyset) \leq 2\mu$.

We now these deductions in (1). Note that the second factor on the RHS in (1) is at most the square-root of

$$\sum_S \hat{f}_\epsilon(S)^2 = \mathbb{E}[f(y)^2] = \mathbb{E}[f(y)] = \hat{f}_\epsilon(\emptyset) \leq 2\mu \leq 4\mu.$$

Also, using (3) for the LHS in (1) we conclude

$$|\mathbb{E}[f(x)f(y)] - \mu^2| \leq 2\mu^{3/2} \cdot \epsilon \sqrt{n} + 2\sqrt{\mu} \sqrt{S_{1-\epsilon'}(f) - \mu^2},$$  (4)

where

$$S_{1-\epsilon'}(f) = \sum_S \hat{f}(S)^2 (1 - \epsilon')^{|S|}.$$  (5)

Let’s simply fix $\epsilon = (1/8) \sqrt{\mu}/\sqrt{n}$ at this point. Doing some arithmetic, it follows that if we can bound

$$S_{1-\epsilon'}(f) - \mu^2 \leq \mu^3/64 (?)$$

(AKA $f$ is “uniform at scale $\epsilon'n$” as Terry might say) then (4) implies

$$\mathbb{E}[f(x)f(y)] \geq \mu^2/2.$$

So long as $\mathbb{P}[x = y] < \mu^2/2$ we’ve established existence of a Sperner pair (AKA non-degenerate combinatorial line). Since this probability is $(1 - \epsilon/2)^n \leq \exp(-\epsilon n/2) = \exp(-\Omega(\sqrt{\mu}/\sqrt{n}))$, we’re done assuming

$$n \geq O(\log^2(1/\mu)/\mu).$$  (6)

Thus things come down to showing (5). Now in general, there is absolutely no reason why this should be true. The idea, though, is that if it’s not true then we can do a density increment. More precisely, it is very easy to show (one might credit this to an old result of Linial-Mansour-Nisan) that $S_{1-\epsilon'}(f)$ is precisely $\mathbb{E}_V[\mathbb{E}[f|V]^2]$, where $V$ is a “random restriction with wildcard probability $\epsilon''$ (and the inner $\mathbb{E}[\cdot]$ is with respect to the uniform distribution). In other words, $V$ is a combinatorial subspace formed...
by fixing each coordinate randomly with probability $1 - \epsilon'$ and leaving it “free” with probability $\epsilon'$. Hence if (5) fails then we have

$$E_V[E[f|V]^2] \geq \mu^2 + \mu^3/64.$$ 

In particular, since $f$ is bounded it follows that $E[f|V]^2 \geq \mu^2 + \mu^3/128$ with probability at least $\mu^3/128$ over the choice of $V$. It’s also very unlikely that $V$ will have fewer than, say, $(\epsilon'/2)n$ wildcards; a large-deviation bound shows this probability is at most $\exp(-\Omega(\epsilon'n))$. Since $\epsilon' \approx \epsilon = (1/8)\sqrt{\mu}/\sqrt{n}$, by choosing the constant in (6) suitably large we can make this large-deviation bound strictly less than $\mu^3/128$. Thus we conclude that there is a positive probability of choosing some $V = V$ which both has at least $(\epsilon'/2)n = \Omega(\sqrt{\mu}\sqrt{n})$ free coordinates and also has

$$E[f|V]^2 \geq \mu^2 + \mu^3/128 \Rightarrow E[f|V] \geq \mu + \mu^2/500.$$ 

I.e., we can achieve a density increment.

If I’m not mistaken, this kind of density increment (gaining $\mu^2/C$ at the expense of going down to $c\sqrt{\mu}\sqrt{n}$ coordinates, with (6) as the base case) will ultimately show that we need the initial density to be at least $1/\log \log n$ (up to log log log $n$ factors?) in order to win. Only a couple of exponentials off the truth :)

The incorrect quantitative aspect here isn’t quite the reason I feel this argument is “wrong”. Rather, I believe that no density increment should be necessary. (Actually, we probably know this is the case, by Sperner’s proof of Sperner.) In other words, I believe that $E[f(x)f(y)] \geq \Omega(\mu^2)$ for any $f$, assuming $\epsilon \ll \sqrt{\mu}/\sqrt{n}$. 

I.e., we can achieve a density increment.