Report on the revision of "Efficiently stable presentations from error correcting codes" by Michael Chapman, Thomas Vidick, and Henry Yuen submitted to Discrete Mathematics

I am not fully happy with the revision. Most of my comments and suggestions have been taken into account, but not all. Several have been ignored for a good reason (14 15 20 were completely wrong and 7 was perhaps not relevant), but it seems that the authors have not got what I meant in my comments 34 and 35.

In 34, I was just asking that you explain in the text of your paper how Corollary 5.12 improves Theorem 7.14 in [JNV+20a]. What you wrote *The statement is stronger than Theorem 7.14 in [JNV+20a], which involves a game that has a question with long answers* is exactly the kind of thing I was asking. Would you agree to add it to the text?

In 35, I was just suggesting a different way of phrasing your proof. I guess that I was implicitely complaining that you were reproducing again the same argument that you had written many times earlier (and that I called "Lemma 1" in my first report). And that it would be kind to the reader if you encapsulated the idea as a lemma so that you can only refer to it each time you use it. Now that you did it in you new Lemma 2.7, you'd better use it rather than reproving it. See 21 below.

More importantly, the comments 9 10 13 14 19 21 have not been fully addressed, or resulted in adding new inaccuracies (see points 2–5, 8, 7, 10, 18). In particular I fully disagree with your reply to 21, see point 18 below.

Here are some comments. Some are new and some are about previous comments.

1. page 9, Definition 2.4. You made the choice to only discuss what is usually called flexible stability and not the more restrictive notion where you do not change de dimension/algebra of the representation. This choice is understandable because flexible stability is much more common that non-flexible stability, but maybe it's worth pointing out the stability results that are true for the stronger form of non-flexible stability. In particular, this seems to be the case for Theorem 4.1. This deserves to be written.

2. page 9, Claim 2.3. There are squares missing. (3) is certainly correct (perhaps with a different constant than 3), but I did not follow, in the proof, how you take into account the renormalization of the trace on $\mathcal{N}$. (4) is obvious with 5 replaced by 1, because by the trace property, if $X = U_i - w^* V_i w$,
$$\|X\|_\tau = \|X^*\|_\tau = \|U_i^{-1} - w^* V_i^{-1} w\|.$$

3. page 10, Lemma 2.5: δ sould be $\sqrt{\delta}$. The fact that you do not put parentheses is misleading, it could be understood as $(\sum_{i=1}^r \|...\|_\tau) + 2\delta$. Put parentheses in the sum, or write $(r+1)\sqrt{\delta} + \sum_{i=1}^{|r|} \|...\|_\tau$.

4. page 12, (7): μ_S is not a probability measure.
5. page 12, proof of Lemma 2.7. You are not proving the lemma: your definition of almost-homomorphism was in terms of $\mathbb{E}\|\phi(r)-1\|_\tau^2$, whereas you prove something about $\mathbb{E}\|\phi(r)-1\|_\tau$.
6. page 20, line 6: the phrasing "to guarantee that R_i is independent of the order..." is misleading. Probably what you mean it that the presentation (20) defines the same group, whatever the order in which the x_i are multiplied. But as a relation, the order always matters, even after adding the relations $R'_{ijj'}$. For example, the notion of (ε,μ)-almost homomorphism changes is the order changes.
7. This is really unimportant, but I meant that the two groups in the statement of the lemma are only isomorphic, and I kind of agreed that the two groups in the proofs were equal.
8. page 20, Definition 3.3 and above. You write "we state the definition for the case of a general prime power q" but then there is no q, just 2. So both occurancies of 2 should be replaced by q.
9. page 22, line 1-2: you should write down the assumption you have on d (for $d \geq q$, the dimension you write is clearly incorrect).
10. page 24, line 8. You could say more clearly that you fix t_i one and for all, and that the way they are chosen has no impact on the statements (it was not clear to me whether they are allowed to change or not, and how they are chosen).
11. page 24, line 3 after Figure 3: is the formula for M correct? I count $q^m \times m \times q \times q + q^m \times 2^t \times 2^t = (m+1)q^{m+2}$, so either I did not understand what *the tester implied by the figure* means, or there is a typo.
12. page 24, line -7: $k = t(d+1)^m$?
13. page 25, Theorem 4.1. There is a typo in the formula for $\delta(\varepsilon)$ (your formula gives $\delta(\varepsilon) \geq C$; this form of the theorem is of course true, but not very impressive).
14. page 25, line 1 in the proof of Theorem 4.1: $\mu_S = \mu_R$.
15. page 25, Claim 4.2: do you need Claim 4.2? Corollary 2.15 that you use later does not require that ϕ takes its values in the hermitian unitaries.
16. page 26, Claim 4.4: why the $\sqrt{\varepsilon}$? Claim 4.3 exactly says that ϕ is (δ,μ_S)-close to U with $\delta = O(\varepsilon)$, and Lemma 2.7 (or its proof) says that $\mathbb{E}\|[U_{u,\alpha},U_{v,\alpha'}] - [\phi(s_{u,\alpha}),\phi(s_{v,\alpha'})]\|^2 \leq 16\delta$ for any probability measure on pairs whose marginals are μ_S.
17. page 28, line 3: what you write is not true for a fixed u, but it is true on average over u.

18. page 29, line -5– -1: the conventions for the notation poly here (given in footnote 7) do not match those of [JNV+22]. For example, on the last line of page 23 of [JNV+22], you have an expression of the form $\mathrm{poly}(\varepsilon, n^{-1}) + e^{-\Omega(r/m^2)}$ that becomes, in the conclusion of [JNV+22, Theorem 4.1] $\mathrm{poly}(\varepsilon, n^{-1}, e^{-\Omega(r/m^2)})$. And similarly all over [JNV+22]. So, the way you prove Theorem 4.1 has two problems: (a) you do not prove that the presentation given in (22) is stable in the sense of Definition 1.2 (because you do not prove $\lim_\varepsilon \delta(\varepsilon) = 0$) and (2) your convention of poly contradicts your use of it.

    When you write *Yes, we are sure that the modulus goes to 0 with epsilon*, I agree that this is true, but where is it proven? I did not find the proof written, neither in [JNV+22] nor in your submission. I thought that what is proven in [JNV+22] is with $\delta(\varepsilon) \leq C(m^C + d^C)(\varepsilon^{1/C} + q^{-1/C})$ for some universal C, which is $C(m^C + d^C)q^{-1/C} > 0$ for $\varepsilon = 0$.

    If you want to stick to the convention that $\lim_\varepsilon \delta(\varepsilon) = 0$, one cheap way is to exploit Lemma 2.6 and [GH17] to explain that the modulus of stability in Theorem 4.1 is something like $\min(C_1\varepsilon, C_2\mathrm{poly}(\varepsilon, 1/q))$ for a good (=polynomial in the parameters) constant C_2 and a bad (=exponential in the parameters) constant C_1.

19. page 29, line -2: in [JNV+22], there is an assumption $r \geq 12mt$ (which in the notation here becomes $r \geq 12m(d+1)$). Do you have assumptions that ensure that $\Omega(tm^2)$ satisfies this inequality?

20. page 29, line -1: you could explain why the existence of G_c implies the conclusion of the theorem: write the homomorphism that is $\delta(\varepsilon)$-close to ϕ.

21. Here is what I had in mind in my comment 35.

    Lemma 1. *Assume that $G = \langle S, R\rangle$ is (δ, μ_R, μ_S)-stable, and let ν_R be another probability measure on R such that $\nu_S = \mu_S$.*[1] *Then every (ε, μ_R)-almost representation is a $(C\mathbb{E}_{r\sim\nu}|r|^2\delta(\varepsilon), \nu_R)$-almost representation.*[2]

    Proof. By the definition of stability, an (ε, μ_R)-representation is $(\delta(\varepsilon), \mu_S)$-close to a representation, that is $(\delta(\varepsilon), \nu_S)$-close to a representation ($\nu_S = \nu_R$). Lemma 2.7 implies that it is a $(C\delta(\varepsilon), \nu_R)$-almost representation. $\square$

    Let us deduce the new inequality in Corollary 2.8:

    $$\forall \varepsilon \leq \frac{1}{k-1}, \delta_k(\varepsilon) \geq \frac{k\varepsilon}{C'}.$$

    Let $G = ([k], E)$ be a perfect matching on k vertices, consider ν_R the uniform probability measure on $\{[x_i, x_i] \mid ij \in E\}$. Then $\nu_S = \mu_S$ is the uniform probability measure $\{x_1, \dots, x_k\}$.

[1] We use the notation from your remark 2.8.

[2] C is the universal constant in your Lemma 2.7.

Consider A the map from Claim 4.2. It is a $(\frac{1}{k-1}, \mu_R)$-almost representation. So $\Phi : x_i \mapsto (A_i, 1)$ is a (ε, μ_R)-almost representation if $\mathcal{M} \oplus \mathbb{C}$ is equipped with the trace $\tau^\varepsilon(x, \lambda) = (k-1)\varepsilon\tau^{\mathcal{M}}(x) + (1-(k-1)\varepsilon)\lambda$. Moreover, for every edge ij, $\|[\Phi(x_i), \Phi(x_j)] - 1\|_{\tau^\varepsilon}^2 = 2(k-1)\varepsilon$, and the lemma implies

$$2(k-1)\varepsilon = \frac{1}{|E|}\sum_{ij\in E} \|[A_i, A_j] - 1\|_{\tau^\varepsilon}^2 \leq C16\delta_k(\varepsilon).$$