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

This paper is at the intersection of operator algebras, group theory and quantum information theory. It revolves around the groundbreaking result "MIP*=RE" which solved Connes' embedding problem in von Neumann algebras using quantum information techniques.

More precisely, this paper deals with so-called Hilbert-Schmidt stability for groups, which roughly asks to what extend a map from a group to a unitary group that is almost multiplicative (in normalized Hilbert-Schmidt norm) is close to an actual homomorphism. This question is mostly studied for a fixed infinite group in several distinct variants, but in each case the way we measure the *almost* and *close* is not very relevant. A notable other example is the investigation by Gowers-Hatami of finite groups where the almost homomorphism is computed on average over the whole multiplication table.

The main point of the paper is to initiate a more systematic study of quantitative forms of stability, that is interesting and far from being understood for finite groups. The main definition is Definition 2.3, that defines the modulus of stability of an arbitrary group presentation together with a probability measure on its relations. This notion is a more general form of a form of quantative stability for groups that was also recently studied in [dlS22]. The whole point of the paper is investigate to what extend the almost homomorphism property can be checked only on a small (hopefully polylogarithmic in the size of the group) part of the multiplication table, while keeping a reasonable $\varepsilon - \delta$ dependance. This amounts to finding not-too-large presentations of the group and good measures on them. To oversimplify, the main example of study are finite groups of the form $(\mathbb{Z}/2\mathbb{Z})^k$; it is very easy to obtain dependance that is polynomial in k, and the whole point here is to obtain examples with sub-polynomial dependance. The first two Sections of the paper are devoted to setting the correct definitions and notion for such a study.

A similar question for non-local games has already been studied in [JNV+20b] and [JNV+22] by the authors of MIP*=RE who, in an impressive tour-de-force, managed to obtain a single quite satisfactory example, constructed from low-degree polynomials–Reed-Muller error correcting codes. This played an absolutely crucial rôle in the MIP*=RE paper, to fix the mistake that was discovered by the authors in the first version of the paper. Synchronous strategies to non-local games are objets that deal with PVM's, that is families of self-adjoint projection that sum to the identity. By Fourier transform, there is a one-to-one correspondance between PVMs indexed by a finite abelian group and unitary representations of its Pontryagin dual. This correspondance has already been exploited in this context, notably in an unpublished note by Vidick *Pauli brading* and the more recent [dlS22]. It is therefore not surprising that the result from [JNV+20b,JNV+22] can be translated to very strong results about stability for groups. This somewhat tedious but elementary task is performed in Section 3-4 (Theorem 4.1).

Then in Section 5, the authors manage to use Theorem 4.1 to obtain a variant of a recent construction in [dlS22]. In the vocabulary introduced in Definition 5.6, in [dlS22] a sequence of $(k, O(\varepsilon))$ qubit test was constructed, with exponentially many questions and polynomially many answers. Here, in Corollary 5.12, a $(k, \delta'_k(\varepsilon))$ qubit test is constructed, with much less (subexponentially) questions, still a polynomial number of answers and with a reasonable control on δ'_k. This is a variant of result that had been proved/claimed in [NV18].

I believe that the body of the paper is useful. It sets the correct framework for future work, and translates important results to this language and raises interesting questions. I believe that several basic results, that are certainly clear to the authors and somewhat implicit in the arguments, should be added explicitely. See my suggestion below. Appendix A contains an original result, which I find quite interesting: it gives precise estimates on the stability of the most obvious presentation of $(\mathbb{Z}/2\mathbb{Z})^k$: it satisfies

$$\frac{1}{C}\min(k\varepsilon, 1) \leq \delta_k(\varepsilon) \leq C\min(k\varepsilon, 1). \tag{1}$$

The upper bound was already known, the content of the appendix is the lower bound. Using the large group of symmetries of the obvious presentation of $(\mathbb{Z}/2\mathbb{Z})^k$, the authors also obtain in § A.1 the same estimate for the L_∞-modulus.

The language of the paper is very good. The first two sections of the paper as well as the Appendix are nicely and very clearly written, but then the quality of writing deteriorates. There are quite a few typos, the conventions start to vary... I list below some specific comments, but the authors should have a careful reading of what they have written. In particular, following the guidelines that you sent me, *I recommend that the paper should not be accepted until the authors have made a significant effort to improve the writing.*

1. page 2, footnote 3: "incolution", "satisfied by out choice".

2. page 4, footnote 5: it should be Definition 2.3.

3. page 6: I am sorry to raise this annoying point... When you mention [NV18], are you refering to something that is actually correct? Because later (page 38) you mention that some of the argument in [NV18] is flawed. A comment or a pointer would be welcome here to clarify what is correct and what isn't in [NV18].

4. page 6, footnote 7: remove "(", or add ")".

5. page 8, line 1: perhaps you could say square-convergent? (defining ℓ_2 as the space of convergence sequences in $\mathbb{C}^{\mathbb{Z}}$ is strange).

6. page 8, line 3: I do not know what "operator topology" is. I am happy if you put any of the standard topologies except the norm, for example the weak operator topology.

7. page 9, Definition 2.3. It would be useful if you commented how your definition of almost homomorphims (and therefore of stability) compares with the point of view that was studied in [dlS22]. Your notion is more general, and essentially coincides with that of [dlS22] for triangular presentations, that is presentations where all relators have length 3. Indeed, in that case, $\mathbb{E}_{r\sim\mu}\|\phi(r)-1\|_\tau^2 = \mathbb{E}_{(g,h)\sim\nu}\|\phi(gh)-\phi(g)\phi(h)\|_\tau^2$, where ν is the image of μ by the map $r = abc \mapsto (a,b)$.

8. page 8, 3 lines before Definition 2.2: there is no μ_S and μ_R so far.

9. page 9, I am not sure I am fully convinced by Remark 2.4, but I will be happy if you can prove me wrong. What you write is fine if you only consider group presentations where all relations have a uniformly bounded length. But for general relations, I would say that the natural distribution induced by μ_R is the measure

$$\mu_{S,1} = \sum_{r\in R} |r|\mu_R(r) \sum_{i=1}^{|r|} \delta_{r_i},$$

where $|r|$ is the length of the relation, and $r = r_1 \dots r_{|r|}$. Or the measure $\frac{1}{\sum_r \mu_R(r)|r|^2}\mu_{S,1}$ if you insist to work with probability measures. Indeed, it is for this choice or $\mu_{S,1}$ that the reverse direction of stability (the one that should be always obvious) holds:

Lemma 1. *Let $G = \langle S : R\rangle$ be a group presentation and μ_R a probability on R. If a homomorphism $\phi : F_S \to \mathcal{U}(\mathcal{M})$ is $(\delta, \mu_{S,1})$-close to a unitary representation of G, then it is a $(O(\delta), \mu_R)$-almost homomorphism.*

10. page 9, after Definition 2.3. You should have a lemma stating that the stability is a group property, and quantifying how the modulus of stability changes when the presentation and the measure μ changes (making precise your discussion after Lemma 2.11).

11. The statement that appears after Lemma 2.12, saying that formula (1) above holds, is stronger than Lemma 2.12, and therefore more interesting. It deserves to be highlighted.

12. page 15, line -2: replace Lemma 2.11 by [CRSV17]. Indeed, Lemma 2.11 does not state how $O_k(\varepsilon)$ depends on k.

13. Page 16 and after: the writing is much less carefull starting there. The notation change from line to line between $\mathbb{Z}/2\mathbb{Z}, \mathbb{F}$ and $\mathbb{F}_2$. I could not understand whether there was a logic. I was often confused when q can be arbitrary, and when $q = 2$. You should really avoid q when you mean $q = 2$ only.

14. page 16, Lemma 3.1: I am certainly too picky, but you could write that these two groups are isomorphic, not equal. More importantly, where did you dedine $R''_{jj'}$?

15. Definition 3.3. Since you talk about parity check, I guess that $q = 2$ and $\mathbb{F} = \mathbb{F}_2$ here. Right?

16. Subsection 3.3 is almost impossibe to read. Line 5, you did not recall self-dual basis. In the next paragraph, a code (which was defined as a linear subspace of $\mathbb{F}_q^n$) suddently appears to become a linear map $\mathbb{F}_q^k \to \mathbb{F}_q^n$. After struggling a lot, I believe I understood that $\mathcal{C}'$ is simply the set of all functions $\widehat{\mathbb{F}_q} \to (\mathbb{Z}/2Z)^n$ (where $\widehat{\mathbb{F}_q}$ is the group of all group homomorphism from the additive group of $\mathbb{F}_q$ to $\mathbb{F}_2$, that, by choosing a self-dual basis, you identify with $\mathbb{F}_q$) of the form $\chi \mapsto (\chi(x_1), \dots, \chi(x_n))$ for $(x_1, \dots, x_n) \in \mathcal{C}$. Is that correct?

17. page 19, line -1: what is $\mathbb{F}$? page 20, line 2: what is ℓ?

18. page 21, line 1:n=N?

19. page 21, line 8: I did not manage to understand where $t_1, \dots, t_d$ are defined.

20. page 21, 6 lines before Theorem 4.1: I am not sure why you make this definition so complicated. If you took $v, v' \in (\mathbb{F}_2^t)^m$ uniformly at random, then you would get a distribution that is simpler but comparable with the one you define up a factor m. And since in Theorem 4.1 you allow any factor $poly(m, d, t)$, this additional factor does not matter.

21. page 21, theorem 4.1: you should clarify the notion $poly(\cdot, \cdot)$. Clearly you have something else in mind that an arbitrary polynomial function (otherwise the theorem would be empty with $\delta(\varepsilon) = 100$). Does it mean a polynomial without constant factor, a polynomial that is divisible by ab, or something else? Is $\varepsilon^{\frac{1}{8}} + q^{-\frac{1}{100}}$ a polynomial in ε and q^{-1}? Also, are you sure that the modulus δ that you obtain in Theorem 4.1 satisfies the assumption $\lim_{t \to 0} \delta(t) = 0$ that you have prescribed in Definition 2.3?

22. page 21, line -4: scakes.

23. page 21, line -3: maximum length of a relation is $d + 2$ (not d), I think.

24. page 22, proof of Claim 4.2: this proof is ok, but it is much faster than the rest of the article, where you usually put all the details. In particular, you seem to be implicitely using something like Lemma 1 above. (another reason to state it).

25. page 23: the proof is way too long, I believe. You should use Corollary 2.10 instead of Lemma 3.5, no?

26. page 23, Claim 4.4: I am not convinced by your choice to leave the dependance on u and j implicit.

27. page 25: same comment as 25., with Theorem 2.6 instead of Lemma 3.5.

28. page 28, footnote 14: the symmetrization is useless, because $D(x, y, a, b)\tau(P_a^x P_b^y) = D(y, x, b, a)\tau(P_b^y P_a^x)$ by the assumption on D and the trace property. Strictly speaking, your formula is even incorrect, because $D(x, y, a, b)$ might be undefined while $\mu(y, x) + \mu(x, y) > 0$, for example if $\mu(y, x) > 0 = \mu(x, y)$.

29. page 29: It is fine to include the proof, but you could say that Lemma 5.5 is analogous (even a particular case?) of Lemm 5.28 in [JNV+21].

30. page 32, in (Soundness) of Definition 5.6. are μ' and $\tilde{\mu}$ the same things?

31. page 33, Definition 5.8. Read what you have written, and correct all the mistakes: $q = 2$? Is "=" missing in $M = (h, \nu)$? Are you assuming that no line in h is zero, otherwise what do you mean by $\frac{1}{|h_{j.}|}$? The measure μ that you define has total mass 2 (so divide by 2), and does not depend on ν! Also, the symbol μ is used again 2 lines after the definition for another meaning.

32. page 36, footnote 16: I did not understand this footnote. And if this is a hypothesis to the theorem, it should appear in the statement and not in a footnote.

33. page 36, line -9: a bit more details could be added about the way (U, V) can be completed: this is where you need to tensor with $1_{\mathcal{H}}$ in Definition 5.6.

34. page 38: what is "the work ??"? What is this important application? Is it Proposition 5.13? What are the motivations for Corollary 5.12? Does it allow to recover all the results claimed in [NV18]? How does Corollary 5.12 compares with Theorem 7.14 in [JNV+20a].

35. Appendix A: I suggest to spell out explicitely the following general phenomenon, that is behind the example. If a presentation $G = \langle S : R \rangle$ if (δ, μ_S, μ_R)-stable, then it implies that any (ε, μ_R)-almost representation of G is also a $(O(\delta(\varepsilon)), \nu)$-representation of G, for any other probability measure ν on R such that $\nu_{R,1} \lesssim \mu_S$ (this is immediate from Lemma 1, and essentially what you are proving in Claim A.3).

36. page 43: mathcing

37. ref [JNV+22] is a reference to a short paper behind a paywall (and that I could not access). Given the importance that this paper plays in the present work, I believe you should cite the longer version published in Discrete Mathematics.