# Review for Margus Niitsoo's seminar paper *"Some applications of pairwise independence"*

Liina Kamm

November 22, 2007

## 1 General Assessment

| Evaluation criterion | No | Rather no than yes | More or less | Rather yes than no | Yes |
|---|---|---|---|---|---|
| The paper is well readable | | | | | √ |
| Language used in the paper is correct | | | | √ | |
| The paper is logical and well structured | | | | √ | |
| The general typeset of the paper is correct | | | | √ | |
| The paper was interesting to read | | | | | √ |
| The paper gives a good overview of the topic | | | | | √ |
| The material in the paper is mathematically correct | | | | | √ |
| References to the external sources are presented correctly | | | | √ | |
| All the relevant references are present | | | | √ | |
| The formulae are typed correctly | | | | √ | |

## 2 Comments

1. *How to evaluate the selection of the topic?* The topic is suitable for a graduate seminar in cryptography.
2. *How to evaluate general presentation style?* The general presentation style of the paper is good. The author explains the notions very loquaciously and makes it seem as though the topic is fairly easy (which of course it is not).
3. *How to evaluate the selection of the information given in the paper?* It seems that, what was intended to be said, has been said.
4. *How to evaluate typesetting of the paper?* The general picture is good, but there are a few things that catch the eye.
5. *What was new and interesting to me in the paper?* The topic of derandomisation is quite unfamiliar to the reviewer. It was actually more interesting than expected.
6. *What else would I have liked to read?* It is quite difficult to point anything out.

# 3 Specific Shortcomings

1. The paper has the following misprints:
   - On page 3, the probability $\Pr[Z_u = t]$, should probably be $\Pr[Z_u = \alpha]$.
   - On page 4, $E[c(\chi)]$ should be equal to $\frac{|E|}{2}$ not $\frac{E}{2}$.
   - On page 5, "wether" should be "whether".
   - On page 6, "happends" should be "happens" or possibly "happened".
   - On page 8, $E[C]$ should be $E[c]$.
   - On page 10, $Z_1$ should probably be $Z_i$.
   - On page 12, "$i$-h" should be "$i$-th".
   - In reference 1, Avi Wigderson's name is misspelled "Widgerson".
2. The paper has the following mistakes in wording and style:
   - The title of Table 1 on page 2 is unclear. Is this perhaps the title of the table on page 9? The table on page 2 is not referenced anywhere in the text.
   - Wording
     - On page 1, the phrase "then exhibit different ways of recycling random bits" is a bit strange. Maybe another word would be better instead of "exhibit".
     - On page 1, it would be better to substitute the word "group" with "set".
     - On page 2, it is written "We call the set $\mathbf{Z}$ *pairwise independent*". Should it perhaps be the distribution that is pairwise independent?
     - On page 2, the sentence "we have a joint distribution for $Z_a$, $Z_b$, $Z_c$ that is given by where every triplet has a probability of $\frac{1}{4}$." has some words missing.
     - On page 2, there is an expression "follow that distribution". It might be better to say "have that distribution".
     - On page 3, the phrase "construct pairwise independent distributions that are far smaller than their fully independent counterparts" indicates that one distribution is smaller than another.
     - On page 3, the phrase "Then use a uniform distribution on $S$ to choose $s \in S$", could be reworded as "Then uniformly choose $s \in S$"
     - On page 4, the phrase "at least as good or better than" should be "at least as good as or better than".
     - On page 4, in the phrase "the same argument goes through if" and on page 10 in the phrase "Analogous analysis goes through for", "goes through" could be substituted with "applies".
     - On page 5, "array" could be substituted with "range".
     - On page 6, "$r \to \infty$" could be written out in words.
     - On page 6, the phrase "the probability $c_{yes}$ can be driven up" is colloquial.
     - On page 6, a word could be added to the phrase "gives us a new $c'_{yes}$" to indicate what $c'_{yes}$ is. Otherwise, the phrase seems a bit incomplete.
     - The Chebychev inequality should have the definite article.
     - On page 7, the word "then" could be removed in the phrase "it is rather easy to prove that then".

- On page 9, the phrase "precision needed of the algorithm" could be "precision needed from the algorithm".
- On page 10, "We look each of" should be "We look at each of".
- On page 12, it is unclear what is meant by the phrase "taking $y$ to be the second endpoint of one of the edges leaving it in the graph"
- The KPS algorithm should also have the definite article (page 13).
- On page 13, the phrase "It follows quite straightforwardly from the EML" could be reworded as "It is quite straightforward to see that the EML implies".
- On page 13, the sentence "The AKS generator also uses the expander graph, but in a more subtle way - namely, instead of taking all the neighbours of a single node, it instead chooses nodes traversed along a random walk in the Expander graph." has too many "instead"-s in it. In addition the "-" sign should be substituted with "–" and should not be the first thing on a new line of text.

- Style
  - On page 3, there are spaces between the parenthesis and the text within.
  - In Chapter 3, it would be better to make sentences that introduce what the letters $P$, $NP$, etc. stand for. At the moment it is a bit too colloquial.
  - The seminar paper does not have a limit for the number of pages so some of the used abbreviations (like TM) could be written out. On the other hand, if the author insists on keeping the acronyms, they should first be given next to the word, so the reader does not have to guess what the letter combinations mean. TM is admittedly not a good example for this, because most of the readers of this paper know what is meant, but on the whole, it is very difficult to read papers where a lot of words have been abbreviated.
  - The title of Chapter 4 could have more words.
  - The reasoning in Chapter 4 could be given in the form of a theorem and its proof.
  - In Chapter 5, in the comparison of the methods, there are abbreviations KPS and AKS. At this point in the text these methods have not been mentioned (admittedly the names of the authors are given in the Abstract, but it has been 9 pages since then). As mentioned before, the abbreviations should first be given next to the notions they stand for, so the reader knows how they will be referenced in the rest of the paper.
  - After a formula or a table that is in the middle of text, it is not necessary to have indentation (page 9, after the table).
  - On page 9, the sentence "We now go on to the algorithms themselves." is too colloquial.
  - It might have been a better idea to have the descriptions of the generators as subsections of one chapter. It seems the author is trying to use metadiscourse, but there seems to be no place for it between the chapters, thus, leaving the introduction to the next chapter at the end of the previous one.
  - On page 11, it is unclear in the phrase "They consider the complexity class $RP$" who "they" are.
  - The sentence on page 11-12 is too long. ("This can be used to show that increasing $l$ by one roughly squares the error probability (plus-minus a few $\epsilon$) as increasing $l$ by one doubles the number of potential witnesses and the lemma allows us to show that the new witnesses created with $h_{l+1}$ behave in roughly the same way as the ones we had before but are independent of them.").

– In a formal paper, it is better to refrain from using references to Wikipedia or any other non-published source. A good alternative would be a more widely known handbook.

3. The paper has the following mathematical mistakes:
   - On pages 1 and 10, the variable $i$ is not defined.
4. The paper has the following mistakes in typesetting:
   - It is nice and more readable if longer equations are put on separate lines. There is, of course, no measurement to distinguish between longer and shorter equations. On page 3 the definition of $h_S(x)$ could be on a separate line, for example.
   - The inequality on page 4 "$c(\chi) \leq \frac{|E|}{2}$" definitely belongs to the longer equation category and should, therefore, be given on a separate line.
   - The table on page 9 might perhaps look better if it were in the center.
   - On page 10, "1/4" should be $\frac{1}{4}$.
   - The probabilities on page 10 should have larger parenthesis that cover all of the "layers" of the probability. Instead of
   $$\Pr[\sum_{i=1}^{k} Z_i < \frac{k}{2}]$$
   it is better to have
   $$\Pr\left[\sum_{i=1}^{k} Z_i < \frac{k}{2}\right] \quad .$$
   - On page 11, the asterisk indicating the footnote should be before the full stop.
5. The paper will be more readable if the author makes the following changes:
   - On page 4, it might be good to indicate that one is dealing with vertex colouring in the phrase "find a two-valued colouring". It is fairly obvious from the rest of the sentence, but if the reader is not familiar with graphs, it will be helpful for finding further information.
   - On page 4, it would be good to define what $E[c(\chi)]$ is.
6. The paper misses the following elements (topics, references, figures, proof steps, etc.):
   - Even though it is clearly stated, that the paper is based on a series of lectures by Luby and Wigderson, it is customary to have a reference to the source as well. All of the references given in the bibliography section have to be cited somewhere in the text.
   - The author mentions that the Impagliazzo-Zuckerman generator will be discussed in the paper and also in Chapter 5, five different methods are mentioned. the Impagliazzo-Zuckerman generator is either very well hidden or simply not there.
   - Maybe it would have been better to include references to the generators as well, so the reader could find further information about them more easily.
7. The following elements could be removed from the paper: The reviewer did not find anything that should be removed.
8. Other comments: