Question 1

It is most important to remember that all assignment questions are SCALED INDIVIDUALLY. In other words, marks per assignment are aggregated on the SCALED MARKS PER QUESTION. A low mark on any particular question does not necessarily indicate sub-par performance relative to the expectations of this course, nor does it represent a poor performance as the cohort average may be low.

1.1

- Almost everyone attained full marks.

1.2

- Students generally did not perform well on this question.
- A surprising majority of students correctly reason that we must minimise the number of splits while also maximising the height of the left-substack post split, but give an almost arbitrary formula for computing the maximal height of the resulting left-most substack. Most incorrect formulas given are easily proven wrong by very simple examples.
- Following from this, it seems that many students simply did not consider the correctness of their algorithms. An attempt at a formal justification of correctness would've quickly shown that any of the aforementioned methods of computing the optimal height of the left-most sub-stack was not correct.
- Some students provided an entirely correct answer but no justification of their correctness at all. These responses were generally awarded less than 10 marks. To provide no working out at all, and omit the justification of correctness which constitutes almost half the marks, leads to any such submissions being worth very little merit.
- Some students provide reasoning to both how they attain their formula, and why their algorithm is correct, but do not attempt to prove this. We have been generous and considered such submissions to have made progress to a proof of correctness and as such, correct answers which did so were generally awarded above 10 marks but below 16.
- Many students prove some sections of their algorithm correctly, or prove many necessary lemmas but do not combine them to show the correctness of their algorithm. Examples of lemmas that people often prove are: "Given the minimum height of all blocks to the right of the current block is b while the height of the current block is h, we must split the current block into a minimum of ceil(b/h) blocks to maintain the non-decreasing property", or "The height of the left sub-stack post split is maximally floor(h/ceil(b/h))". Proofs of these 2 lemmas are necessary but independently, do not show the correctness of the algorithm. We considered the proof of these lemmas to have made progress to a complete proof and submissions which correctly do so were generally awarded above 13 marks but below 18.
- Some students almost proved their algorithm correctly, but omit certain necessary lemmas/parts of the proof. The marks given/penalised would be dependent on the degree to which a submission omitted non-trivial elements of the proof. Generally, even the top submissions did not prove the algorithm entirely correctly.

Question 2

2.1

- Most students did well in this part, properly proving the required results.

- The most common mistake was approaching the proof in the wrong direction, by assuming the statement and showing something true. As a simple example as to why this fails, if I start with $2 \leq 1$ and multiply both sides by 0, then $0 \leq 0$ which is true – you need to be aware of destructive steps, or simply approach the proof from true statements to begin with.
- Some students lost marks for general clarity of their solutions. Please take a look at some of the simple proofs in first year mathematics notes or textbooks – they aren't entirely symbols, English is used to add clarity to the proof and guide the reader.
- It is important to note that the condition $a \leq b \leq c \leq d$ does allow for $d - a$ and similar to be zero, so dividing through by this isn't guaranteed to be possible.

2.2

- The majority of students gave a correct algorithm, but failed to properly justify correctness of their algorithm.
- Some students quoted rearrangement inequality as justification without reproducing a proof in their own words and providing a reference. It is expected that students using results from outside the course cite the source and reproduce the proof to the best of their ability to demonstrate understanding of the result.
- Most students approaching the proof as an exchange argument could use the result from 2.1 to conclude that their solution was locally maximal (that is, one swap won't decrease the sum), but didn't include a complete exchange proof to prove that this optimality is global. An exchange argument requires two parts: showing that your solution is locally optimal, and showing that local optimality extends to global optimality, generally by showing that every optimal solution can be exchanged using local optimisations into your solution.
- Some exchange arguments attempted to justify global optimality, but failed to provide a clear explanation for why an optimal solution can be converted into their solution. It is expected that you provide a clear measure for how close your solution is to the optimal solution, and explain why this decreases after every exchange. For example, by taking an optimal solution and finding the first differing pair, then exchanging it to remove this difference – the number of matching pairs then increases after every exchange until the entire solution is the same.
- Inductive proofs of the generalisation of 2.1 to k pairs were rarer, but generally failed to actually prove an inductive step. Many students opted to state the base case from 2.1 and claim without proof that this could be generalised.
- Some attempted to prove the inductive generalisation by proving the result for 3 pairs, but not going any further. This is not induction.
- Steps like iterating an array to calculate the sum still form part of complexity analysis as well, you should mention the complexity of this as part of your overall analysis.
- Quicksort is not an $O(n \log n)$ algorithm.

2.3

- The comments for 2.2 apply to this subquestion as well.
- Some students attempted to break up the array into groups of four and use 2.1. This works for 2.2 coincidentally, but pairing up in groups of four does not work for this subquestion, since the optimal solution involves pairing up opposite elements in the array of marks. This is a nice example of where local optimisations don't extend to a global minimum.

Question 3

3.1 Most students received full marks on this question. There were no marks for proof of correctness, if a valid counter-example was given, all 5 marks were received. A few students provided trivial examples, i.e. examples using backward edges, and a few provided examples where there actually was a common topological sort, but most were correct. In a few rare cases, marks were deducted for extremely complex responses. Many students however had responses more complex than necessary. In general, best practice is to make counterexamples as clear and simple as possible. The best solutions here only had 3 nodes, 2 edges in one graph, and 1 edge in the other.

3.2 Students did okay on this question overall. Many struggled with rigorous proof of correctness, but a good proportion came up with a correct algorithm. As well as the merging of graphs solution presented in the solutions, many students submitted variations on an equivalent solution, which was to run Kahn's algorithm but at every step only take nodes that have no incoming edges in both X and Y. This is functionally identical to the merge solution, but was a bit more difficult to justify, both in terms of correctness and time complexity.
Most incorrect responses were one of the following

- Students attempting to enumerate all topological sorts of both graphs, and comparing the pair. Responses of this type received very low marks, between 0 and 5. The main issue is that there are $O(n!)$ topological sorts of a given graph in the worst case (i.e. a very sparse graph), and therefore comparing two lists of them cannot be done more efficiently than $O(n!)$ time. Several students quoted a geeksforgeeks webpage incorrectly claiming this could be done in $O(V * (V + E))$ time using 'backtracking'. This page was incorrect, and an instructor on Ed had already pointed it out as incorrect. furthermore, even if that complexity WAS correct, that would still give $O(n^3)$ time as there at at most $O(n^2)$ edges. Therefore it would not be sufficiently efficient anyway.
- Students creating a single topological sort for each graph, and then comparing those. This approach would have way too many false negatives, as two graphs that have a common topological sort may well have many non-common topological sorts. A variation on this was students creating a topological sort for X and then checking if it was compatible with Y. Students received quite low marks for these approaches, as they are incorrect and show a lack of understanding of topological ordering.
- students who constructed some notion of "layers" in a topological sort. While in a single sort it is true that there are sometimes "layers" of nodes that could be swapped in arbitrary order, this is not a fundamental structure of topological sorting, and most students who attempted to use this notion had a solution that was not generalisable to all cases.

Question 4

For this question, most students either got (near) full marks or around 6-8 marks. This part was marked quite leniently. Therefore, please be aware that a mark review request could adjust your mark downwards. However, if you strongly believe that your mark is incorrect, you are encouraged to submit a review request with a clear explanation of the reason for your review request.

4.1

- This part of the question was pretty straightforward. Most students managed to claim the allocated 4 marks.
- Partial credit was given for reasonable attempts.
- A handful of students misunderstood the question. The question states "Provide an expression for the probability that Alice **arrives k minutes before** Bob" indicating that the

arrival times should be **exactly k minutes** apart. We are not interested in finding the probability that Alice arrives **at least** k minutes before Bob. Please read the question carefully and if you have doubts, please approach the teaching staff via the channels available.

4.2

- Most students who attempted this part with the prefix sum approach managed to get a good mark for this part.
- For solutions involving convolution using FFT, we expected a level of detail similar to slide 29 of Lectures 5-6 (Fast Fourier Transform). Please also refer to the sample solution as an example.
- In addition to presenting the workings of the method, students are expected to justify the correctness of the proposed method. Up to 4 marks would have been taken away for unclear or insufficient correctness justifications.

4.3

- Almost all the students who attempted this part, computed the probability that the person arrives at an odd-numbered time instant and an even-numbered time instant for Alice and Bob and then used the probabilities to compute the expected answer.
- Note that the arrival times need **not** be even. It is sufficient that they are even-number of minutes **apart**. That is, Alice can arrive at $t = 1$ and Bob can arrive at $t = [1, 3, 5, 7, \ldots]$. We want to consider these cases even though none of the values are even.