In this submission, the author gives necessary and sufficient criteria when a quantum channel can be approximately simulated by another quantum channel. More precisely, first, statistical deficiency is defined as the minimal simulation distance under different classes of allowed simulation schemes. Second, this is operationally characterised in terms of equations featuring operational guessing probabilities, as well as generalised notions of min-entropies for quantum channels (that were studied previously by various authors). These findings can be seen as the full extension of Le Cam's classical randomisation criteria for statistical experiments to the non-commutative regime.

The results are first derived in the setting of general probabilistic theories (GPTs) by employing the mathematical framework of base section norms (Section 2). This is then specialised to quantum channels and corresponding simulation schemes in terms of post-processing, pre-processing, and general super channels (Section 3). Norm duality and minimax theorems are tools used prominently.

Comments:
-The manuscript is well-written.
-The results are worked out with great care for mathematical detail and look correct. Proofs are often assisted using a helpful graphical representation of quantum states and quantum channels.
-The author represents the literature well and gives appropriate references throughout the manuscript.
-The results are timely for the study of resource theories of channels in the quantum information theory community. In particular, the author uncovers some of the underlying mathematics about the statistics of quantum channels and as such the results have great promise for future applications (this concerns in particular the study of the generalised notions of conditional min-entropy).
-I only have one major point of criticism, which is that one has to go through some heavy mathematical formalism in order to understand the results about quantum channels (first in Section 2 and then in Section 3). I understand that coming from GPTs makes the story mathematically more general and the proofs more elegant. However, for an information theory audience there might be more straightforward ways to present material (this is in contrast to a mathematical physics journal). As an example, the quantum content of the manuscript was presented at the Algebraic and Statistical ways into Quantum Resource Theories workshop at BIRS (as also mentioned in the acknowledgements). I found the presentation there very accessible, basically giving full proofs without too heavy notation. I suggest to maybe shift some of the abstract and the introduction in this direction? Personally, I would also find it helpful to see a version of the quantum results (Theorems 4-6) sketched in the introduction.

Minor comments:
-The main result is named a "randomisation theorem". Just curious, where does this name come from (as one might argue the result is about channel simulations)?
-Before Equation (4) there seems to be some inconsistent notation with Lambda_1 versus Lambda_pre? Please check throughout manuscript.
-In Lemma 3 there seems to be some typos concerning the name of the channels?
-When referring to Examples 1 and 2 it might be helpful to put a reminder that those

are concerning the GPT setting (as I was reading the quantum section on its own)?
-As noted in the manuscript, Section 3.3 was already derived by the same author in the IEEE ISIT proceedings [10]. I assume that full proofs can be found in these proceedings? If not it might be worthwhile to include these proofs here anyway.
-In Section 3.4 it is noted that Theorem 5 was already partially proven in the arXiv preprint [11] by the same author. Please clarify if [11] is or will be published anywhere else.
-As mentioned in the conclusions, it would be interesting from an information-theoretic viewpoint to explore if certain cases of restricted super channels (such as shared randomness based, non-signalling, positive partial transpose, entanglement-breaking, LOCC, etc.) lead to neat criteria for channel simulation. Maybe one could state under what exact mathematical conditions the current results could be applied as a starting point?
-As mentioned in the conclusions, the results discussed here cover the one-shot setting. It happens that channel simulation is also prominently studied in classical and quantum Shannon theory in the asymptotic regime. From and information-theoretic viewpoint it would then be interesting to explore if the characterisations derived here could be directly useful for finding single-letter asymptotic expansions for certain settings?

Overall, I recommend to accept the manuscript under review for publication in IEEE Transactions on Information Theory. Nevertheless, I do encourage the author to take my comments into account.