* **Dear Colleagues, Here are some notes on these remarks. Boyan, perhaps you can use these when you do the first re-write. Warmly, Dennis**
* **Limited contribution**
	+ Generally, there’s not a lot we can do, but maybe stress more on the contribution, e.g. in the introduction of Section 4
	It’s usually good to mention the contribution in the beginning and then maybe repeat in section 4. Somehow, I thought we had done that.
	+ Link motivation to problem statement, e.g.
		- We need a highly scalable solution, because… There are increasing numbers of data streams as it becomes easier and easier to build sensors, create new financial instruments, introduce new social media sources.
		- We need online response times per window, because… reacting quickly to new trends can contribute to better services, improved economic efficiency etc.
		- We can sacrifice 5% of the accuracy in favor of the above, because… many times we don’t need to find all connections, but the connections we do find should be right. In trading for example, there is no loss in missing an opportunity cost, but the opportunities we see should be real.
		- Linear distance (Pearson, Euclid, …) is fine, because… In the short term even complex functions are piecewise linear.
* **Poor presentation**
	+ Improve problem statement
		- Formalize, e.g. given X, find Y, optimizing Z Given a set of time series, find the most highly positively or negatively correlated pairs as quickly as possible.
		- Define parametrization
		- Define the connection with parallelization and similarity measures Parallelization is just for speed.
	+ Define terms, e.g. cooperative time series Don’t we do that? Good encoding with fourier transforms (fourier transforms capture most of the energy).
	+ Implementation details
	+ Formalize algorithms / pseudo-code?
* **Choice of parameter values**
	+ Justify the choice of threshold 0.7, fraction 2/3, etc. Don’t we try several already?
	+ Recall bounds, recall to performance tradeoff
	+ Discussion in the experimental section