I like how the sandwiching worked out. Here are some thoughts:

1. Section 5 looks cramped. I think it should be cut out and then later become a paper on its own. As it stands it might get overlooked. Or somebody might want to implement it but then the paper won't have enough support. Also, a longer paper with enough supporting detail won't be novel enough for a good journal.

   (a) The followup paper should have a title to convey the "fixed relative error" idea so it won't get lost.

   (b) There is a lot of careful detail involved in setting up the hybrid of absolute and relative error. I think that is too much to squeeze into a page or so. Most people will skim it. It should be a section in a larger paper.

   (c) Then there should be extensive numerical evaluation and some thought on how best to pick the infinite sequence of error tolerances.

   (d) The infinite sequence calls to mind sequential analysis and processes that exit either upper or lower boundaries. Also, it starts to look reminiscent of Chow and Robins (infinite number of 'looks').

   (e) The algorithm should be moved out of the theorem statement.

2. The paper is getting long. So my earlier remark on Chow and Robbins and related papers can be abbreviated. How about

   > Chow and Robbins (1965) develop a sequential sampling fixed width confidence interval procedure for the mean. Their procedure attains the desired coverage level in the limit as $\varepsilon \to 0$. But it does not provide coverage gaurantees for fixed $\varepsilon > 0$.

   We don't necessarily have to mention all the later papers.

3. p2 CLT provides confidence interval
   CLT provides a confidence interval

4. p2 inequalities Hoeffding's
   inequalities such as Hoeffding's

5. at (4) we should say something for the -3 crowd

6. p2 below (4) has: the the

7. p2 "Firstly" and "Second of all,"
   Firstly and Second,
   (or) First and Second,

8. p3 (maybe)
   some d-variate function f
   some function f
   [the domain of f is already given as $\mathbb{R}^d$] (I'm less than sure which way this should go; sometimes redundancy helps.)

9. p3 practical error estimation for these methods remains a challenge
these methods do not provide fixed-width confidence intervals.
(Perhaps separate QMC from RQMC. QMC has no practical error estimates. RQMC does, but they're not fixed-width.)

10. p4 I've not previously seen a convention for taking $\gamma = 0$ and $\kappa = -2$ when $\sigma = 0$. Do we need $\sigma = 0$?

11. p4 after first display: *sigma*
σ

12. p4 Slutsky's theorem is in Lehmann and Romano (2005) page 433.

```
@book{lehm:roma:2005,
 author = {E. L. Lehmann and J. P. Romano},
 year   = 2005,
 title  = {Testing Statistical Hypotheses},
 publisher = {Springer},
 edition = {3rd},
 address = {New York}
}
```

13. p5 Good catch. I had the wrong Hall paper. The right one is

```
@article{hall:1988,
author  = {P. Hall},
year    = 1988,
title   = {Theoretical comparisons of bootstrap confidence intervals},
journal = {The Annals of Statistics},
volume  = 16,
number  = 3,
pages   = {927--953}
}
```

The coverage errors are on page 948. The usual intervals correspond to π_{STUD}.

14. Depending on your appetite for more notation fiddling, it occurs to me that we might want to write n_σ and n_μ instead of n_σ and n.

15. (17) needs a period (not comma).

16. (19) has a new wrinkle: making $n \geq n_\sigma$. Was that condition enforced in the simulations? I cannot tell but suspect it is not used in Section 3.4 on the cost of the algorithm. I think it is reasonable advice but we should be sure to spell out when it is used and when it is not.

17. Theorem 5 statement:
fixed width confidence interval
fixed width confidence interval condition

18. Remark 1:
    should that be $\sigma \le \varepsilon\sqrt{\alpha n_\sigma}$? (instead of σ^2)

19. p9 top:
    one popular case is occurs when Y is a d-variate function
    one frequently encountered case has Y a d-variate function
    In this case μ
    Then μ

20. Equation (22): as mentioned above it is not clear that this enforces $n \ge n_\sigma$. The n_σ clause is in the definition of N_μ but I don't see how that clause gets honored in the displayed math at the bottom of page 9. That might be the same derivation as before the clause was added. Then it was ok to just use N_{BE} because it was inside a min(). Now it is also inside a max(). Please double check.

21. Figure 1's caption does not describe the many curves that appear. It should therefore say something about how those curves are defined in the surrounding text for the benefit of those who read figures first and the article second. Their full description is too bulky to go into the caption.

22. p11 end of Section 3. Add some examples. A sentence like "For example with $\mathfrak{C} = 1.5$ and $\tilde{\kappa} = 10, 100, 1000$ we get $n_\sigma =$ xxx, xxx, and xxx respectively."

23. Somewhere we should have a comment like the following:

    > In certain cases our procedure multiplies the computational cost by a large factor such as 4 or 10 or even 100 compared to what one might spend with a known value of σ. While this seems inefficient, it is well to remember that the total elapsed time may still be well below one second.

    I leave it to you to put in more appropriate numbers perhaps after seeing how the finance example comes out. The comment might belong in the introduction or the conclusions. It came to mind when I was reading Section 3 though that is probably not where it belongs.

24. Section 4 title contains "Algorighm"

25. p11 has: the the absolute error

26. The second full sentence on p12 has a grammatical problem.

27. 131 072
    131,072

28. Figure 3: the vertical dashed line and the horizontal Failure line are prominent and important. But it is not clear from the caption how they are defined. Something should be said.

29. As mentioned above, please double check whether $n \geq n_\sigma$ was enforced for the numerical examples.

30. As described above I think this section should expand into its own paper. I think there are a few things to tweak

    (a) p14 form of this criterion would be
    form of this criterion is

    (b) p15 one must have $(1-\theta)+\theta|\mu| \neq 0$. Sometimes it is ok to write ratio criteria $A/B \leq \epsilon$ as $A \leq B\epsilon$. But then you have to watch like a hawk/lawyer to see whether $\leq$ or $<$ is correct, both inside and outside the probability function.

    (c) p15 The sentence containing "happy" is not grammatical.

    (d) p16 It then follows then by
    It then follows by

31. In the discussion where we point out that non-convexity absolves us from those theorems on adaptive methods we might also mention that it also lets us escape a Bahadur and Savage condition.