Reviewer's report: MN-18-1820-MJ, Ross et al

This is an interesting paper. I apologize for the tardy report. I have been traveling for the past 10 days by unconventional means that have spared me limited energy or internet access.

The phenomenology reported here is enough to justify publication. The theoretical interpretation is plausible in the large but could be tightened.

1. Too much credit is given to Sirko & Goodman 2003. That paper was mainly concerned with the implications of self-gravity for the spectrum emitted by the *outer* parts of the disc $r \gg r_g$. Little if any consideration was given to the boundary conditions at the inner edge of the disc (zero vs. nonzero torque).

2. Because the present authors are concerned with inner parts of the disc, they should be more careful in considering relativistic effects. In particular, the ISCO probably lies at $r < 6GM/c^2 = 3r_{\rm S}$ because AGN black holes have significant rotation (e.g., Reynolds 214, SSRv 183, 277; Capellupo et al. 2016, MN 460, 212). The specific binding energy at the ISCO is therefore probably larger than $1 - \sqrt{8/9} \approx 0.057$, so that even if there were no torque at the ISCO, the ratio of the bolometric luminosity to the monochromatic luminosity at any point in their spectra (e.g. $\lambda = 5100\mathring{A}$) would be larger than they suppose even in a steadily accreting thin disc. Naturally, however, the ISCO would not be expected to change on human timescales.

3. Although I agree that a thermal or viscous instability is a more plausible interpretation of their data, the arguments given here against transient absorption/obscuration are not wholly convincing and could be improved. For the authors' values of the black-hole mass ($7 \times 10^8\, M_\odot$) and Eddington ratio (0.07), the rest-frame emission at $\lambda \leq 5100\mathring{A}$ comes from $r \leq 0.01$ pc ($\approx 225\, r_g$, as they estimate). An obscuring cloud just large enough to block our view of this region, but in circular orbit at the sublimation radius (0.4 pc) would cross its own diameter in approximately six years, which is comparable to the time interval spanned by the spectroscopic observations (≈ 7.7 yr). However, were the "UV collapse" caused by absorption, there would be other implications. The authors should determine what (non-grey) reddening law would be needed to explain the differences among the spectra, and whether, for any reasonable gas-to-dust ratio and turbulent broadening, the absorber would produce transient optical absorption lines (since the optical emission is obscured by only a factor ~ 2); and finally, whether such a small absorber could affect the infrared flux (probably not).

4. The bald statement (p. 8, col. 2) "The 2010 spectrum can not be fit with…a simple MTB model with an alternate temperature profile" requires justification. Really? The continua cannot be fit even if $T_{\rm eff}(r)$ is a completely free function? Why not? Of course no superposition of local blackbodies will explain the emission lines, but that is already a known limitation of thin-disc models for "ordinary" QSOs.

5. The authors are too quick to dismiss thermal instabilty along the lines of the works by Hameury, Lasota, and colleagues for dwarf novae and X-ray binaries. It is known that radiation-pressure-dominated discs are thermally unstable (Lightman & Eardley 1974; Jiang, Stone, & Davis 2013). Whereas the authors consider (and then dismiss in favour of cold absorbers) simple thermal instability in the "hot" direction—leading to an RIAF—it could equally go in the opposite direction, leading to a cold, low-$\dot{M}$, gas-pressure-dominated state, since the growth rates of linear instabilities are independent of the sign of their initial perturbations. The thermal timescale $(\alpha\Omega)^{-1}$ at $225r_g$ is ~ 4 yr, in the right range to explain these observations. Admittedly, it is worth asking why *all* QSOs do not undergo propagating thermal instabilities of this sort. But perhaps they do: it would be interesting to compare the numbers of changing-look QSOs with what might be expected on such a model. To do this, however, would be beyond the scope of the present paper.