QUASAR ACCRETION DISKS ARE STRONGLY INHOMOGENEOUS

JASON DEXTER and ERIC AGOL

1 Department of Physics, University of Washington, Seattle, WA 98195-1560, USA; jdxter@uwashington.edu
2 Department of Astronomy, University of Washington, P.O. Box 351580, Seattle, WA 98195, USA

Received 2010 October 23; accepted 2010 December 14; published 2010 December 30

ABSTRACT

Active galactic nuclei have been observed to vary stochastically with 10%–20% rms amplitudes over a range of optical wavelengths where the emission arises in an accretion disk. Since the accretion disk is unlikely to vary coherently, local fluctuations may be significantly larger than the global rms variability. We investigate toy models of quasar accretion disks consisting of a number of regions, \( n \), whose temperatures vary independently with an amplitude of \( \sigma_T \) in dex. Models with large fluctuations (\( \sigma_T = 0.35–0.50 \)) in \( 10^2–10^3 \) independently fluctuating zones for every factor of two in radius can explain the observed discrepancy between thin accretion disk sizes inferred from microlensing events and optical luminosity while matching the observed optical variability. For the same range of \( \sigma_T \), inhomogeneous disk spectra provide excellent fits to the Hubble Space Telescope quasar composite without invoking global Compton scattering atmospheres to explain the high levels of observed UV emission. Simulated microlensing light curves for the Einstein cross from our time-varying toy models are well fit using a time-steady power-law temperature disk and produce magnification light curves that are consistent with current microlensing observations. Deviations due to the inhomogeneous, time-dependent disk structure should occur above the 1% level in the light curves, detectable in future microlensing observations with millimagnitude sensitivity.

Key words: accretion, accretion disks – black hole physics – galaxies: active – gravitational lensing: micro

Online-only material: color figures

1. INTRODUCTION

Optical emission from active galactic nuclei (AGNs) is thought to be due to thermal emission from a standard thin accretion disk (Shakura & Sunyaev 1973; Novikov & Thorne 1973). This model is successful in explaining many observations of black hole accretion disks, such as the peak frequencies of the thermal spectra of X-ray binaries and AGNs. It is also used to fit continuum spectra in the X-ray to measure black hole spin (e.g., Shafee et al. 2006) and to infer accretion disk size from gravitational microlensing, both in absolute terms (e.g., Morgan et al. 2010) and as a function of wavelength (Anguita et al. 2008; Poindexter et al. 2008; Eigenbrod et al. 2008; Morgan et al. 2010). Attempts to fit continuum AGN spectra with thin disk models have met with mixed success (Blaes et al. 2001; Davis et al. 2007). While multi-temperature blackbody spectra generally fit well in the optical, at UV wavelengths quasar spectra are brighter than predicted (e.g., Blaes et al. 2001). The X-ray size has been shown by microlensing to be extremely compact (Chartas et al. 2009; Dai et al. 2010), disfavoring spectral models invoking global Comptonizing atmospheres to explain the UV emission.

Quasar optical variability is also difficult to explain in the context of the thin disk model. Long-term monitoring of AGN has found almost simultaneous variability across optical wavelengths, with lags of less than 1–2 days (Cutri et al. 1985; Clavel et al. 1991; Korista et al. 1995; Wanders et al. 1997; Collier et al. 1998). Comparing these lags to the radii dominating the thin disk emission at these wavelengths gives a traveling speed of \( 0.1c \) for the variability mechanism (Krolik et al. 1991; Courvoisier & Clavel 1991). This would force the variability in AGN to be communicated at the local sound or Alfvén speed rather than on the infall timescale associated with disk instabilities.

Wambsganss et al. (1990) and Rauch & Blandford (1991) first measured the size of an accretion disk from microlensing and found that accretion disks were smaller than expected from the thin disk model for the observed \( L_\nu \). In recent years, the opposite trend has emerged (Pooley et al. 2007; Dai et al. 2010; Morgan et al. 2010; Blackburne et al. 2010). Sizes are now robustly found to be a factor of \( \sim 4 \) larger on average than expected from the thin disk model at IR/optical/UV wavelengths. The size versus wavelength relation predicted by the thin disk, \( r \propto \lambda^{3/4} \), is within the large range allowed by microlensing observations (Eigenbrod et al. 2008). At the same time, Kelly et al. (2009), Kozłowski et al. (2010), and MacLeod et al. (2010) have studied large samples of quasar light curves. They find that optical quasar variability is well described by a damped random walk which returns to a mean value on a typical timescale of 200 days with variability amplitudes of \( \approx 10\%–20\% \).

It is unlikely that the observed quasar variability is caused by a coherently varying accretion disk,\(^3\) but rather is probably the added effect of many smaller, independently varying regions. Many such models have been proposed to explain quasar variability (e.g., Lyubarskii 1997). In this Letter, we demonstrate that for the observed variability characteristics, such an inhomogeneous disk can simultaneously explain multiple discrepancies between AGN observations and accretion disk theory. Inhomogeneous disks can be large enough to explain the microlensing observations, while their temperature fluctuations on small spatial scales naturally explain the observed simultaneous variability across optical wavelengths. Temperatures exceeding the local thin disk value lead to broader spectra extending into the UV, consistent with quasar spectra without invoking a Compton scattering medium.

\(^3\) The implications of coherent variations for microlensing measurements are discussed by Blackburne & Kochanek (2010).
onto a Cartesian grid for plotting, and the color scale increases logarithmically from blue to red to yellow to white with a dynamic range of 100.

Figure 1. False color temperature maps of damped random walk models with \( n = 2200 \) (left), \( n = 550 \) (middle), and \( n = 140 \) (right). The images are interpolated onto a Cartesian grid for plotting, and the color scale increases logarithmically from blue to red to yellow to white with a dynamic range of 100.

(A color version of this figure is available in the online journal.)

2. INHOMOGENEOUS ACCRETION DISKS

We assume that (1) the optical/UV emission observed from AGN originates in an optically thick accretion disk. (2) Variations in the disk occur locally and are uncorrelated on large spatial scales. (3) Fluid in the disk is on circular Keplerian orbits. Assumption (1) allows us to model the disk emission using temperature alone, while (2) requires that the disk be inhomogeneous. Assumption (3) is only used for computing microlensing light curves, where the disk surface brightness map enters into the magnification light curve produced. Kelly et al. (2009) found that quasar light curves are well described as a CAR(1) process, a random walk that tends to return to a mean value on a typical timescale. The timescale they found was roughly 200 days, consistent with the thermal timescale in the predicted optical emission region of many AGNs. This behavior was confirmed for SDSS Stripe 82 quasars by MacLeod et al. (2010). The amplitude of variations in typical sources is 10%–20%. To produce this variability amplitude with multiple, independent regions in the disk requires larger local variations (total variance \( \propto N^{-1} \) for \( N \) independent zones). The local accretion disk flux and effective temperature will then no longer be a monotonic function of radius \( r \), but will fluctuate with azimuth and time. This will cause the disk spectrum within an annulus to have stronger emission at shorter wavelengths than if the same flux were emitted at constant temperature. Consequently, the outer portions of the disk will contribute more flux than for a uniform disk, causing the disk to appear larger at a particular wavelength.

The thin disk model\(^4\) is broadly consistent with many observed properties of black hole accretion flows and follows from the conservation of angular momentum and energy in the gravitational potential of the black hole. Its global time-averaged properties are likely correct. Assuming a geometrically thin, optically thick disk the local flux can be written \( F(r, \phi, t) = \pi B_\nu(T) \), where \( F \) is the flux at the frequency \( \nu \), \( B_\nu \) is the Planck function and \( T = (F/\sigma_B)^{1/4} \), and \( \sigma_B \) is the Stefan–Boltzmann constant. In the thin disk case, \( T \propto r^{-3/4} \) well outside the inner disk edge and is independent of \( \phi \) and \( t \) everywhere.

3. DAMPED RANDOM WALK DISK MODEL

We allow \( T \) to be a function of \( \phi \) and \( t \) as well as \( r \) by dividing the disk into \( n \) evenly spaced zones in \( \log r \) and \( \phi \)

\(^4\) In this Letter, we use “thin disk model” to mean that of Shakura & Sunyaev (1973). Inhomogeneous disks may also be geometrically thin.

\[ \sigma_{T
u}^2 \simeq \frac{3}{8n} \left( e^{3\sigma_T^2} - 1 \right) \langle L_\nu \rangle^2. \]
For a large number of zones, the disk can be significantly larger than a thin disk while producing little variability. For few (large) zones, the overall variability becomes unrealistically large before the accretion disk becomes large enough to explain the microlensing measurements. By design, power spectra and structure functions from this damped random walk model are consistent with studies of quasar variability (Kelly et al. 2009; Kozłowski et al. 2010; MacLeod et al. 2010). This model is by no means unique. We have also tried other time-variable toy models for inhomogeneous disks, which can reproduce the observed variability characteristics. All models considered give similar size–variability amplitude relations.

In the limit $n \to \infty$, the damped random walk model becomes time independent with a log-normal distribution of disk temperatures in each annulus whose variance is $(\ln 10 \sigma_T^2)/2$. In this limit, the flux from an annulus can be written as

$$F_\nu(r; \sigma_T) = \frac{2\pi h^3}{c^2} \int_0^\infty \frac{dw}{\sqrt{\pi} \ln 10 \sigma_T w} e^{-\left(\frac{\ln(w/e^{-w})^2}{\ln 10 \sigma_T^2}\right)} e^{w/w - 1}, \quad (2)$$

where $z \equiv h/\nuT$, $T$ is the thin disk temperature, and the local temperature in an infinitesimal piece of the annulus is $wT$. The specific luminosity can be computed from

$$L_\nu(\sigma_T) = \int_0^\infty 2\pi r F_\nu(r; \sigma_T) dr. \quad (3)$$

The size increase as a function of $\sigma_T$ is the ratio of half-light radii from this model assuming $T \propto r^{-3/4}$ to those computed from $L_\nu(0)$. Sizes are compared at fixed $L_\nu$ since the half-light radius, $r_h$, is a function of $L_\nu$ ($r_h \propto \sqrt{L_\nu}$ ignoring the disk inner edge). The relation,

$$r_h(\sigma_T)/r_h(0) \equiv e^{0.85(\ln 10 \sigma_T)^2}, \quad (4)$$

gives an excellent fit to numerical calculations. In models with $\sigma_T = 0.35–0.5$, 5%–10% of the zones in a given annulus produce 50% of its flux in agreement with the analytical log-normal ($n \to \infty$) model. The fractional area contributing half the flux decreases with increasing $n$ and $\sigma_T$.

### 3.1. Microlensing Simulations

Thus far, we have compared ratios of inhomogeneous to smooth model thin disk half-light radii at the same value of $L_\nu$. However, microlensing observations do not measure half-light radii. Instead, they measure uncorrelated variability between different images of strongly lensed quasars, caused by the motion of stars in the lens galaxy. Sizes can be measured by assuming a power-law thin disk with $T(r) \propto r^{-3/4}$ and fitting magnification light curves with surface brightness maps parameterized by the radius, $r_s$, where $h/\nu = kT(r_s)$ (Kochanek 2004). For the power-law disk, $r_s$ can also be calculated from the observed $L_\nu$ in the absence of all lensing effects. The discrepancy between the values of $r_s$ inferred from microlensing versus $r_s$ inferred from $L_\nu$ is 0.6 ± 0.3 dex (Morgan et al. 2010). In the power-law disk, $r_{hS} = 2.44r_s$. We would like to determine whether microlensing light curves from our inhomogeneous toy model can be fit well by the power-law disk model, and whether the values of $r_s$ inferred from microlensing simulations of the inhomogeneous model are consistent with our previous results for the half-light radius.

Using the code described by Wambsganss (1999), we simulate magnification patterns for QSO 2237+0305, the Einstein Cross, using lens galaxy parameters found by Kochanek (2004). Choosing random orientations and starting positions, many magnification light curves are produced from the time-varying toy models at multiple observed frequencies, and for the power-law disk at many values of $r_s$. The microlensing light curves are then fit to determine best-fit values of $r_s$. Sample light curves are shown in the top panel of Figure 3 along with best-fit power-law disk light curves, while the bottom panel shows fractional deviations between the two.

The time-steady, power-law disk model produces excellent fits to the microlensing light curves. In the damped random walk model, $r_s$ from microlensing is consistent with the half-light radius of the disk at various frequencies by a fixed fraction that depends on $n$. For many (few) zones, $r_s$ underestimates (overestimates) the time-averaged disk half-light radius. The largest deviations are of order 30% and will not dominate the error in estimating quasar accretion disk sizes from microlensing.

### 3.2. Spectra and Light Curves

Damped random walk model spectra are compared to the composite quasar spectrum from Zheng et al. (1997) in the top panel of Figure 4. The temperature normalizations have been chosen to provide the best fit. In all cases, the toy model spectra have larger $F_\nu$ at smaller $\lambda$ than the thin accretion disk model. The bottom panel of Figure 4 shows local spectra at $r = 20r_{ms}$, where $r_{ms}$ is the inner disk edge. The local inhomogeneous spectra peak at a smaller value of $\lambda F_\nu$ than the single temperature blackbody. To produce the same total flux at that wavelength, the emission must arise from a larger area. Thus, the disk appears larger at that wavelength.
These spectral differences will cause parameter values inferred from optical spectra to differ between the two models. We find that inferred black hole masses (accretion rates) are smaller (larger) when fitting inhomogeneous spectra with thin disk models. The discrepancy is typically less than a factor of two and is unlikely to be a significant source of error in estimating accretion disk parameters from spectral fitting.

The damped random walk model produces simultaneous optical variability, in that the cross-correlation between all wave bands peaks at a lag of zero time steps. The lag is zero regardless of the chosen time step, because neighboring zones are completely independent. Variability is stronger at higher frequencies, in qualitative agreement with observations. Quantitative predictions for this relation, for cross-correlations between light curves at different observed wavelengths, or for trends with luminosity or black hole mass will require a physical model for the variability mechanism. These quantities are highly sensitive to the particular inhomogeneous disk prescription chosen, many of which satisfy the existing observational constraints.

3.3. Combined Observational Constraints

Figure 5 summarizes our main results. The 68% allowed regions of the $\sigma_T$ versus $n$ parameter space from microlensing, spectral, and variability measurements are shown for the damped random walk model, along with contours of rms deviations between model and best-fit power-law disk microlensing light curves at 2500 Å. The disk size and variability constraints are the same as in Figure 2. The spectral constraint is found from fitting model spectra to the composite spectrum from Zheng et al. (1997), varying the temperature normalization in each case to obtain the best fit.

The independent spectral and microlensing size constraints require similar amplitudes of temperature inhomogeneity. Producing the observed level of variability requires larger $\sigma_T$ at larger $n$. All constraints intersect in the green region, giving best-fit parameter values of $0.35 \lesssim \sigma_T \lesssim 0.5$ and $100 \lesssim n \lesssim 1500$. The models predict short timescale deviations from smooth power-law disk models in microlensing light curves at or above the 1% level on average, with maximum deviations of 10% (see Figure 3). These are likely to be lower limits. Deviations from other toy models tried can be a factor of a few larger. Short timescale variations in microlensing light curves have previously been predicted to arise from accretion disk inhomogeneities (Gould & Miralda-Escude 1997), but could also be caused by an inhomogeneous intervening medium (e.g., optically thick broad line region clouds; Wyithe & Loeb 2002).

4. PHYSICAL MECHANISMS

Quasar accretion disks are expected to be inhomogeneous to some degree. However, to produce the observed sizes and broad spectra observed in AGNs, the inhomogeneities must be large—factor of three to five variations in temperature at a given radius. Disk instabilities are the most promising mechanism for producing inhomogeneous disks. The magnetorotational instability (MRI; Balbus & Hawley 1991) is now believed to operate in a wide range of accretion flows, causing the disk to become fully turbulent and providing a means for outward angular momentum transport and accretion. Local (e.g., Miller & Stone 2000) and global simulations of thin MRI disks have been performed both in pseudo-Newtonian potentials (Armitage et al. 2001; Armitage & Reynolds 2003) and recently in full general relativity (Shafee et al. 2008; Noble et al. 2009, 2010; Penna et al. 2010). None of these simulations included radiation forces, which likely dominate the dynamics of AGN accretion disks over a large range in radius (Shakura & Sunyaev 1973). Hirose et al. (2009a) studied a set of local simulations in radiative MHD, many of which were radiation dominated. Their simulations show temperature variations with factors of ~2 in a
local patch of disk. Global general relativistic MHD simulations from Fragile et al. (2007) and McKinney & Blandford (2009) have gas temperatures that can vary by factors of two to three at a single radius in the midplane, but are considerably thicker than AGN disks and neglect radiation altogether. These distributions correspond to $\sigma_T = 0.1$–0.2. It is unclear whether the MRI-driven fluctuations in radiation-dominated global disks will be larger.

Thin, radiation-dominated disks have long been believed to be subject to a strong thermal instability, since the heating rate depends on a higher power of the temperature than the cooling rate. Hirose et al. (2009b) find that the dynamics of shearing box simulations of radiation-dominated disks are well described by the alpha model on long timescales. The flow is highly variable, but thermally stable. Radiation-dominated disks are also subject to an inflow instability (Lightman & Eardley 1974). Radially extended, radiation-dominated MHD simulations are needed to determine whether this instability operates in AGNs. If so, it would likely create large local temperature gradients in the accretion flow between hot, tenuous gas and cooler optically thick regions. Finally, it is possible that a photon bubble instability could be present in radiation-dominated accretion disks (Turner et al. 2005), which can cause the flux escaping the disk to rise by factors of several locally. This may prevent inhomogeneous disks from becoming dynamically unstable, despite exceeding the Eddington limit locally (Begelman 2006). The disk would also remain stable if magnetic pressure provides the vertical support in quasar accretion disks (e.g., Begelman & Pringle 2007). Alternatively, the disk may drive a wind when the Eddington limit is exceeded.

5. CONCLUSIONS

Standard thin disk accretion has formed the basis for understanding X-ray binaries and AGNs for nearly 40 years. However, it has always been difficult to extend the model to explain optical quasar variability. In recent years, microlensing observations of multiply imaged quasars have provided a probe of the disk structure, finding that quasar microlensing sizes are robustly larger than the flux sizes predicted from thin disk theory. If the average temperature structure remains identical to that in thin disk theory but is highly inhomogeneous, accretion disks can be large enough to explain the sizes found by microlensing while matching the observed level of optical variability. The level of inhomogeneity ($\sigma_T = 0.35$–0.50 with $n = 10^2$–$10^3$) required to explain the discrepancy in microlensing sizes is in excellent agreement with that necessary to fit observed quasar spectra. Such inhomogeneous structure produces short timescale variations in microlensing light curves that should be larger than $\pm 1\%$. The amplitude of the temperature fluctuations can be further constrained from measuring the size of these deviations. The range of temperatures in small regions of the inhomogeneous disk explains the simultaneous variability observed across optical wavelengths.

We have demonstrated this idea with an unphysical toy model. Proper modeling of an inhomogeneous disk will require global MHD simulations of radiation-dominated accretion disks. It is unclear whether the MRI alone is sufficient to produce the required temperature fluctuations or whether additional disk instabilities or other variability mechanisms are also important.

We thank Robert Antonucci, Brandon Kelly, Joachim Wambsganss, and the referee, Chris Kochanek, for useful comments.

This work was partially supported by NASA Earth & Space Science Fellowship NNX08AX59H. Support for program HST-GO-11225.02-A was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555.

REFERENCES

Anguita, T., Schmidt, R. W., Turner, E. L., Wambsganss, J., Webster, R. L., Loomis, K. A., Long, D., & McMillan, R. 2008, A&A, 480, 327

Armitage, P. J., & Reynolds, C. S. 2003, MNRAS, 341, 1041

Armitage, P. J., Reynolds, C. S., & Chiang, J. 2001, ApJ, 548, 868

Balbus, S. A., & Hawley, J. F. 1991, ApJ, 376, 214

Begelman, M. C. 2006, ApJ, 643, 1065

Begelman, M. C., & Pringle, J. E. 2007, MNRAS, 375, 1070

Blackburne, J. A., & Kochanek, C. S. 2010, ApJ, 718, 1079

Blackburne, J. A., Pooley, D., Rappaport, S., & Schechter, P. L. 2010, arXiv:1007.1665

Blaes, O., Hubeny, I., Agol, E., & Krolik, J. H. 2001, ApJ, 563, 560

Chartas, G., Kochanek, C. S., Dai, X., Poindexter, S., & Garnire, G. 2009, ApJ, 693, 174

Clavel, J., et al. 1991, ApJ, 366, 64

Collier, S. J., et al. 1998, ApJ, 500, 162

Courvoisier, T. J.-L., & Clavel, J. 1991, A&A, 248, 389

Cutri, R. M., Wisniewski, W. Z., Rieke, G. H., & Lebofsky, M. J. 1985, ApJ, 296, 423

Dai, X., Kochanek, C. S., Chartas, G., Kozowski, S., Morgan, C. W., Garnire, G., & Agol, E. 2010, ApJ, 709, 278

Davis, S. W., Woo, J., & Blaes, O. M. 2007, ApJ, 668, 682

Eigenbrod, A., Courbin, F., Meylan, G., Agol, E., Anguita, T., Schmidt, R. W., & Wambsganss, J. 2008, A&A, 490, 93

Fragile, P. C., Blues, O. M., Anninos, P., & Salmonson, J. D. 2007, ApJ, 668, 417

Gould, A., & Miralda-Escude, J. 1997, ApJ, 483, L13

Hirose, S., Blaes, O., & Krolik, J. H. 2009a, ApJ, 704, 781

Hirose, S., Krolik, J. H., & Blaes, O. 2009b, ApJ, 691, 16

Kelly, B. C., Bechtold, J., & Siemiginowska, A. 2009, ApJ, 698, 895

Kochanek, C. S. 2004, ApJ, 605, 58

Korista, K. T., et al. 1995, ApJS, 97, 285

Kozlowski, S., et al. 2010, ApJ, 708, 927

Krolik, J. H., Horne, K., Killman, T. R., Malkan, M. A., Edelson, R. A., & Kriss, G. A. 1991, ApJ, 371, 541

Lightman, A. P., & Eardley, D. M. 1974, ApJ, 187, L1

Lyskarski, Y. E. 1997, MNRAS, 292, 679

MacLeod, C. L., et al. 2010, ApJ, 721, 1014

McKinney, J. C., & Blandford, R. D. 2009, MNRAS, 394, L126

Miller, K. A., & Stone, J. M. 2009, ApJ, 543, 398

Morgan, C. W., Kochanek, C. S., Morgan, N. D., & Falco, E. E. 2010, ApJ, 712, 1129

Noble, S. C., Krolik, J. H., & Hawley, J. F. 2009, ApJ, 692, 411

Noble, S. C., Krolik, J. H., & Hawley, J. F. 2010, ApJ, 711, 959

Novikov, I. D., & Thorne, K. S. 1973, in Black holes (Les astres occuls), ed. C. DeWitt & B. DeWitt (New York: Gordon and Breach), 343

Penna, R. F., McKinney, J. C., Narayan, R., Tchekhovskoy, A., Shafee, R., & McClintock, J. E. 2010, MNRAS, 408, 752

Poindexter, S., Morgan, N., & Kochanek, C. S. 2008, ApJ, 673, 34

Pooley, D., Blackburne, J. A., Rappaport, S., & Schechter, P. L. 2007, ApJ, 661, 19

Rauch, K. P., & Blandford, R. D. 1991, ApJ, 381, L39

Shafee, R., McClintock, J. E., Narayan, R., Davis, S. W., Li, L., & Remillard, R. A. 2006, ApJ, 636, L113

Shafee, R., McKinney, J. C., Narayan, R., Tchekhovskoy, A., Gammie, C. F., & McClintock, J. E. 2008, ApJ, 687, L25

Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337

Turner, N. J., Blaes, O. M., Socrates, A., Begelman, M. C., & Davis, S. W. 2005, ApJ, 624, 267

Wambsganss, J. 1999, J. Comput. Appl. Math., 109, 353

Wambsganss, J., Paczyński, B., & Schneider, P. 1990, ApJ, 358, L33

Wanders, I., et al. 1997, ApJS, 113, 69

Wyithe, J. S. B., & Loeb, A. 2002, ApJ, 577, 615

Zheng, W., Kriss, G. A., Telfer, R. C., Grimes, J. P., & Davidsen, A. F. 1997, ApJ, 475, 469