Draft of 2018.08.08–00:15 -0400. **CONFIDENTIAL**

MNRAS , 1–8 (2018) Preprint 8 August 2018 Compiled using MNRAS LATEX style file v3.0

The major contributor to systematic differences between Gaia DR2 and VLBI positions

L. Petrov[1,2⋆], Y. Y. Kovalev[2,3,4] and A. V. Plavin[2,3]
[1]*Astrogeo Center, 7312 Sportsman Dr., Falls Church, VA 22043, USA*
[2]*Astro Space Center of Lebedev Physical Institute, Profsoyuznaya 84/32, 117997 Moscow, Russia*
[3]*Moscow Institute of Physics and Technology, Dolgoprudny, Institutsky per., 9, Moscow, Russia*
[4]*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany*

Accepted xxx. Received yyy; in original form zzz

ABSTRACT

We have analyzed the differences in positions of 9030 matching sources between the Gaia DR2 and VLBI catalogues. We found that position errors should be re-scaled by a factor of 1.3 for VLBI and 1.06 for Gaia, and in addition, Gaia errors should be multiplied by the square root of chi square per degree of freedom. Position uncertainties of matching sources in the VLBI catalogue is a factor of two worse than position accuracy in the Gaia DR2. There are 9% matching sources with statistically significant offsets. We have established that the major contributor to statistically significant position offsets is the presence of optical jet. Among the sources for which the jet direction was determined, the position offsets are along the jet directions for 62% outliers. We found evidence of position jitter along the jet direction, probably associated with optical variability. The jitter results in statistically significant proper motions of quasars that we have predicted earlier. Therefore, an assumption that quasars are unmovable objects and therefore, differential proper motions determined with respect to quasars can be regarded as absolute proper motions, should be treated with a great caution.

Key words: galaxies: active – galaxies: jets – quasars: general – radio continuum: galaxies – astrometry: reference systems

1 INTRODUCTION

Since 80s very long baseline interferometry (VLBI) was the most accurate absolute astrometry technique. Accuracy of VLBI absolute positions can reach 0.1 mas level. With few exceptions, VLBI is able to provide absolute positions only for active galactic nuclea (AGNs). In 2016, the Gaia Data Release 1 (DR1) (Lindegren et al. 2016) ushered an appearance of the technique that rivals VLBI in accuracy. Quick analysis by Mignard et al. (2016b) found that in general, the differences between common AGNs in VLBI and Gaia DR1 catalogues are close to their uncertainties, except 6% common objects. Mignard et al. (2016b) claims that "close examination a number of these cases shows that a likely explanation for the offset can often be found, for example in the form of a bright host galaxy or nearby star". They conclude (page 13) that "the overall agreement between the optical and radio positions is excellent". We do not think that if two independent observing campaigns produced small (negligible) differences, such an outcome should be called excellent, because it implies that the contribution of a new campaign is also small (negligible) with respect to what has been known before. Science does not emerge from agreements. It emerges from disagreements. Therefore, we focused our analysis on differences between VLBI and Gaia AGN positions.

Our analysis of Gaia DR1 confirmed the existence of a population of sources with a statistically significant VLBI/Gaia offsets (Petrov & Kovalev 2017a). We found that such factors as failures in quality control in both VLBI and Gaia, blended nearby stars, or bright host galaxies can account at maximum 1/3 of that population. This analysis, as well as work of others (Mignard et al. 2016a; Makarov et al. 2017; Frouard et al. 2018; Liu et al. 2018a,b,c), used arc lengths of VLBI/Gaia differences. Including the second dimension, position angle of VLBI/Gaia offsets, resulted in a breakthrough. Though the distribution of the position angle counted from the declination axis turned out to be close to uniform, the distribution of the position angles of respect to the jet direction determined from analysis of VLBI images of matching sources revealed a strong anisotropy (Kovalev et al. 2017): they have have a preferable direction along the jet, and at a smaller extent along the opposite direction. We interpret it as a manifestation of a presence of optical jets at scales finer than the Gaia point spread function (PSF), i.e. 100–300 mas. It should be noted that even if radio and optical jets are perfectly co-spatial, as ground observations of some AGNs with very large jets resolved by the HST suggest (Gabuzda et al. 2006; Perlman et al. 2010; Meyer et al. 2018), still there will be position differences. Since a

⋆ E-mail: Leonid.Petrov@lpetrov.net

response to an extended structure of a power detector used by Gaia and an interferometer that records voltage is fundamentally different, as it was shown in (Petrov & Kovalev 2017b). Gaia positions correspond to location of the optical centroid, while VLBI position are associated with the most compact feature of the jet base. Therefore, the physical meaning of the VLBI/Gaia offset is a displacement of the optical centroid with respect to the jet base.

In April 2018, the Gaia DR2 was published (Lindegren et al. 2018). It has 48% more sources than Gaia DR1 and significantly higher accuracy. Mignard et al. (2018) found that in general, the differences VLBI/Gaia DR2 are small with some exceptions. They set out five reasons for discrepancies (page 10): 1) real offsets between the centres of emission at optical and radio wavelengths; 2) error in matching of VLBI and Gaia objects; 3) an extended galaxy around the quasar; 4) double or lensed quasars; or 5) simply statistical outliers. Though the authors were aware of results in Kovalev et al. (2017), they did not mention the presence of optical jet as a likely explanation, tacitly assuming this factor so insignificant that it is not worth mentioning.

In Petrov & Kovalev (2017b) we examined consequences of our interpretation of VLBI/Gaia offsets due to presence of optical jets. Among others, we made two predictions: 1) "further improvement in the position accuracy of VLBI and Gaia will not result in a reconciliation of radio and optical positions, but will result in improvement of the accuracy of determination of these position differences", 2) "we predict a jitter in Gaia centroid position estimates for radio- loud AGNs" (pages 3785–3786). Since accuracy of Gaia DR2 is noticeably better than accuracy of Gaia DR1, this motivated us to extend our previous analysis to Gaia DR2 and check whether these predictions came true. We predicted the impact of optical structure in VLBI/Gaia DR2 differences will be more significant than in VLBI/Gaia DR1, while Mignard et al. (2018) tacitly assume this factor is insignificant at all. Answering the question which interpretation is correct and what is the most significant contributor to systematic position differences is the goal of this article.

2 COMPARISON OF VLBI/GAIA POSITIONS

We matched Gaia DR2 catalogue of 1,692,919,135 objects against the Radio Fundamental Catalogue rfc_2018b (Petrov and Kovalev in preparation, 2018)[1] (RFC) of 15,155 sources. The RFC catalogue is derived using all VLBI observations under astrometric programs publicly available by August 01 2018. We used the same procedure of matching of Gaia objects against the VLBI catalogue describe in more details in Petrov & Kovalev (2017a) and got 9081 matches with the probability of false association below $2 \cdot 10^{-4}$ level. Immediate comparison of formal uncertainties *among matches* showed that Gaia uncertainties are smaller (see Figure 1). The median semi-major error ellipse of the VLBI sample is 0.74 mas against the 0.34 mas of the Gaia sample. Although VLBI can reach accuracy 0.1 mas in absolute positions of strong sources, the majority of the sources were observed only once for 60 seconds, which is insufficient to derive their position with that level of accuracy. It is fair to say Gaia uncertainties of matches are roughly 2 smaller than errors of VLBI matches, though there is no grounds for generalization of this statement to the entire Gaia or VLBI catalogues.

[1] Available online at http://astrogeo.org/rfc

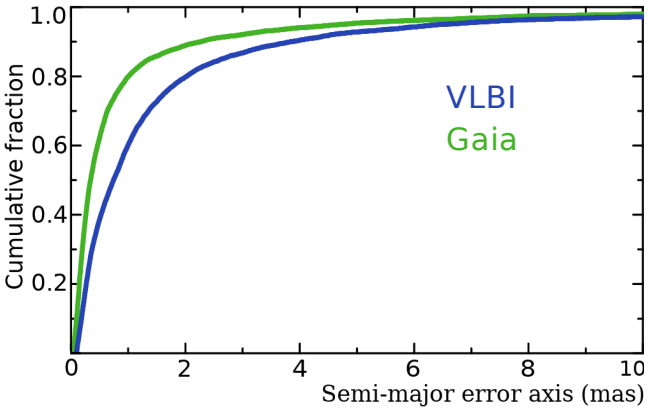

Figure 1. Cumulative distribution function of semi-major error axes $P(\sigma_{\rm maj} < a)$: green (upper) curve for Gaia and blue (low) curve for VLBI.

Among 9081 matches, 8080 have radio images at milliarcsecond resolution. Using these images, we have evaluated jet directions for 4030 sources, i.e. for one half. We removed 39 radio stars, supernova remnants in nearby star-forming galaxies, and known double objects, such as gravitational lenses, from further analysis

2.1 Analysis of VLBI/Gaia position angles with respect to the jet direction

We examined arc length a between VLBI and Gaia source position estimates and the position angle of VLBI offset with respect to Gaia offset ϕ counted counter-clockwise with respect to the declination axis. Using reported position uncertainties and correlations between right ascensions and declinations, we computed semi-major, semi-minor axes error ellipse, and position angles θ for both VLBI and Gaia position estimates. Then, assuming VLBI and Gaia errors are independent, we computed uncertainties of arc lengths σ_a and position offsets σ_ϕ in the linear approximation:

$$\begin{aligned}
\sigma_a^2 &= \frac{1+\tan^2(\theta_v-\phi)}{1+\dfrac{\sigma^2_{\rm v,maj}}{\sigma^2_{\rm v,min}}}\,\sigma^2_{\rm v,maj} + \frac{1+\tan^2(\theta_g-\phi)}{1+\dfrac{\sigma^2_{\rm g,maj}}{\sigma^2_{\rm g,min}}}\,\sigma^2_{\rm g,maj} \\
\sigma_\phi^2 &= \Delta(\alpha_g-\alpha_v)^2(\sigma^2_{v,\delta}+\sigma^2_{g,\delta})\cos^2\delta_v/a^4 + \\
&\quad \Delta(\delta_g-\delta_v)^2(\sigma^2_{v,\alpha}+\sigma^2_{g,\alpha})\cos^2\delta_v/a^4 - \\
&\quad 2\Delta(\alpha_g-\alpha_v)\Delta(\delta_g-\delta_v)\cdot \\
&\quad (\mathrm{Corr}_v\sigma_{v,\alpha}\sigma_{v,\delta}+\mathrm{Corr}_g\sigma_{g,\alpha}\sigma_{g,\delta})\cos^2\delta_v/a^4,
\end{aligned} \tag{1}$$

where Corr is correlation between right ascension and declination and uncertainties in right ascensions are assumed reported without $\cos\delta$ factor.

Figure 2 shows the distribution of the normalized arc-lengths a/σ_a among all the matches. The last bin contains 1067 sources with normalized arcs greater 5, or 11.4%. The number of sources with normalized arcs greater 4, what for this study we consider statistically significant, is 16.3%, or 1/6. The goal of our study is to explain these outliers.

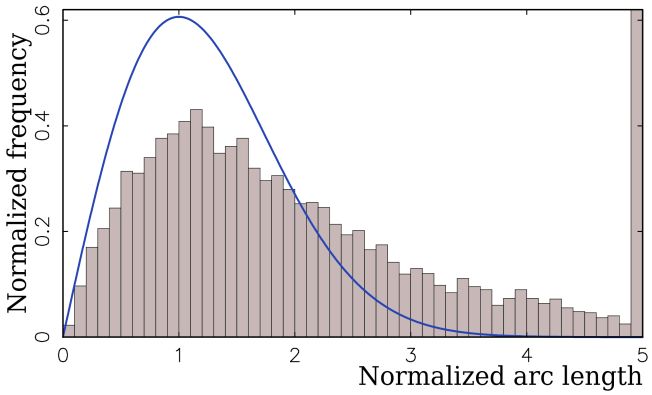

Figure 2. Distribution of normalized VLBI/Gaia arc-lengths over 9042 matching sources. The last bin that holds normalized arc lengths > 5 exceeds the plot bounding box. The blue smooth curve shows Rayleigh distribution with $\sigma = 1$.

We computed the histograms of the distribution of position angle offsets with respect to the jet direction determined from analysis of VLBI images of matches at milliarcsecond scales. We denote this quantity as ψ. Figure 3a shows such a histogram. Comparing this Figure with the upper left Figure 3 in Kovalev et al. (2017) shows the anisotropy is revealed even more clearly: the peaks became sharper and narrower. The height of the peak with respect to the background is 2.8 versus 1.7 and the full width at half maximum (FWHM) is 0.42 rad versus 0.62 rad. We confirmed that anisotropy of ψ angle is not an artifact of Gaia DR1, but the real phenomena, and more precise observations measured it more precisely. The prediction made in Petrov & Kovalev (2017b) has come true.

We should note that the empirical histogram of ψ that is sample estimate of the distribution is affected by measurement errors of ψ. Its error depends on a/σ_ϕ. Ignoring errors in determination of jet direction angle that at the moment we cannot precisely characterize, we assume $\sigma_\psi = \sigma_\phi$. At large a/σ_ϕ (say, more than 4) the distribution of a ψ errors for a given measurement converges to the normal distribution. At low a/σ_ϕ (say less than 0.25), the distribution is converging to the uniform. The analytic expression for ψ measurement errors can be found in page 233 of Thompson et al. (2017). Including measurements of ψ with large errors smears the histogram. In order to mitigate smearing, we filtered out matches with $\sigma_\psi > 0.3$ rad. We found empirically, that reducing the threshold further results in the histograms to degrade due to scarcity of remaining points, though not to change their shape noticeably.

Figure 3b shows the histogram of ψ angles for all the matches with $\sigma_\psi < 0.3$ rad. The peaks at 0° and 180° became much stronger. Further analysis revealed that the histograms are different for short and long arc distance between VLBI and Gaia positions. Figure 3c and 3d show the histogram for $\sigma_\psi < 0.3$ rad, for short and long arc lengths respectively.

To characterize histograms, we fitted it to a mathematical model. We have selected a model that is as simple as possible:

Table 1. Results of fitting the model in eq. 2 to the histograms in Figures 3a–d.

Case	α	FWHM$_1$ rad	β	FWHM$_2$ rad	$1-\alpha-2\beta$	# src
a	0.08	0.42	0.17	2.03	0.58	4015
b	0.23	0.40	0.22	1.48	0.33	985
c	0.07	0.35	0.17	1.01	0.47	423
d	0.24	0.40	0.17	1.84	0.28	565

$$f(\psi) = \alpha N(0,\sigma_1) \; + \; \beta N(0,\sigma_2) \; + \; \beta N(\pi,\sigma_2) + \frac{1-\alpha-2\beta}{2\pi}, \quad (2)$$

where $N(a,\sigma)$ is the normalized Gaussian function with first two moments a and σ. In the context of this study a choice of functions to fit the empirical distribution is irrelevant as far as the approximate the data. Parameter α describes the contribution of the main narrow peak, parameters β describes the contribution of the secondary wide peaks that has the maximum at both 0 and π, and the last term describes the contribution of the uniform component of the distribution. We noticed that the broad peaks at $\psi = 0$ and π has similar shape and fitting them separately with two additional parameters does not improve fit. Results of fitting this 4-parametric model to the histograms in Figures 3a–d is shown in Table 1.

We see that the main peak at $\psi = 0$ and FWHM around 0.4 rad that is rather insensitive to the way how a subsample is drawn. We tentatively conclude that this is the intrinsic width of the peak. The peak at $\psi = 0$ is formed by predominately matches with large position offsets. We surmise that this peak is related to optical jets and its width is roughly corresponds to the average jet opening angle.

Two secondary peaks are broad, with maximum at $\psi = 0$ and π. They are formed by matches almost exclusively with offsets shorter than 2–2.5 mas. This can be explained by a more turbulent environment or with clouds of obscuring matter that are more common at distances within 2.5 mas of the jet base. The share of these secondary peaks in the distribution is relatively insensitive to the way how the subsample is drawn, 0.17–0.22, but its FWHM varies. We interpret it as an indication that a simplistic 4-parameter model is too coarse to fully describe the empirical distribution which shape depends on the VLBI/Gaia offset length.

The fifth column in Table 1 shows the share of the component with the uniform distribution, i.e. the component that is not related to the core-jet morphology. This share is 0.58 for the histogram build using all the observations. The share is reduced to 0.33 for the subsample of observations with $\sigma_\psi < 0.3$ rad and to 0.25 for the subsample of observations with $\sigma_\psi < 0.2$ rad. This reduction occurs partly due to mitigation of the histogram smearing, partly due to the selection bias. Since σ_ψ depends on both uncertainties of position estimates and the arc-length, setting the upper limit for σ_ψ disproportionately favours the matches with long VLBI/Gaia offsets.

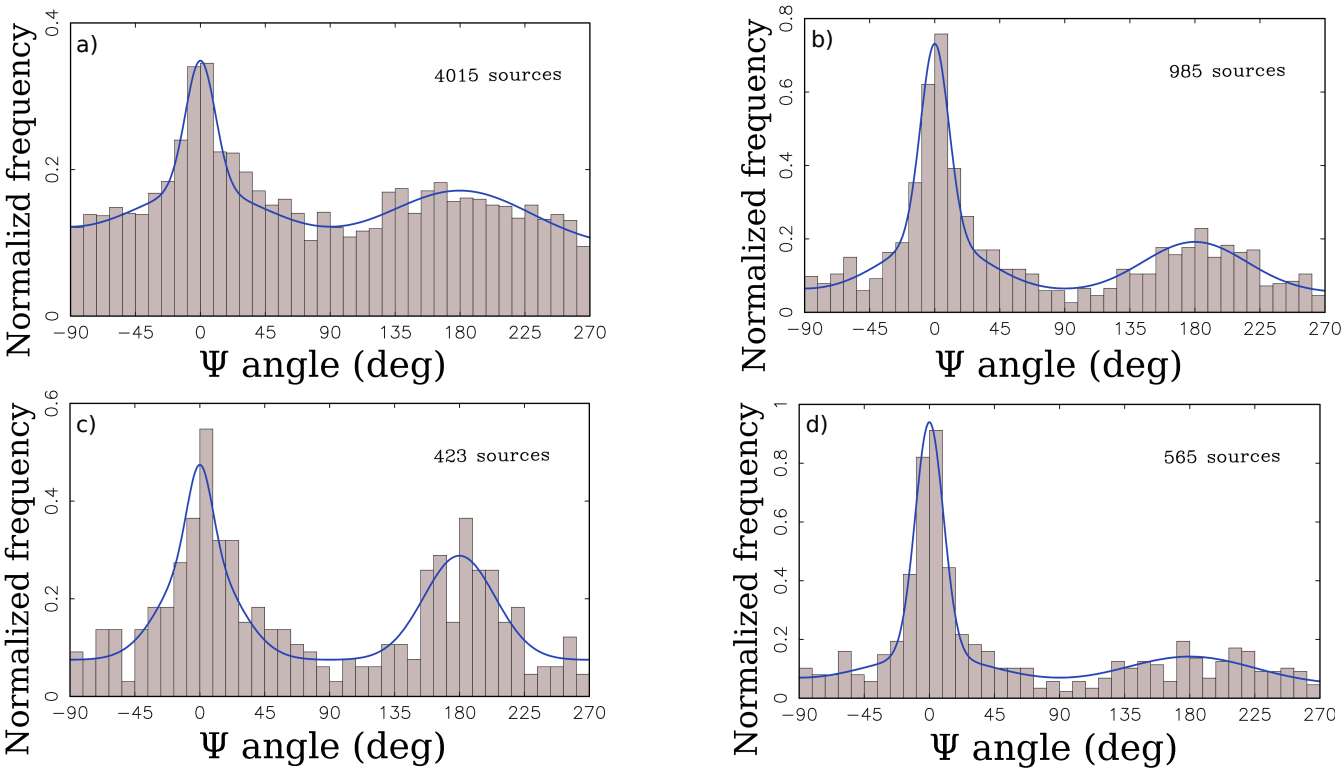

Figure 3. The histograms of the distribution of the position angle of Gaia offset with respect to VLBI position counted with respect to jet direction counter-clockwise. *Top left (a):* all the matches with known jet directions. *Top right (b):* the matches with $\sigma_\psi < 0.3$ rad. *Bottom left (c):* the matches with $\sigma_\psi < 0.3$ rad and arc-lengths < 2.5 mas. *Bottom right (d):* the matches with $\sigma_\psi < 0.3$ rad and arc-lengths > 2.5 mas. Blue curves are the best approximation of a three-component model.

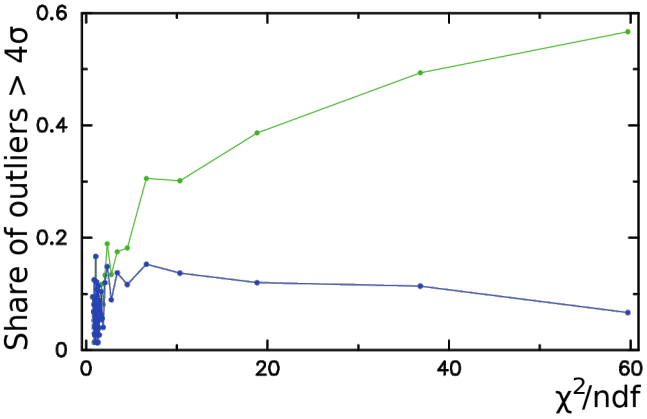

Figure 4. The share of outliers with normalized arc length of VLBI and Gaia matches > 4 for 1% percentiles of χ^2/ndf. The horizontal axis along the median value of χ^2/ndf within each 1% percentile. The upper green curve was computed using original Gaia position uncertainties. The low blue curve was computed using Gaia uncertainties multiplied by $\sqrt{\chi^2/\mathrm{ndf}}$ factor.

2.2 Rescaling VLBI and Gaia reported position uncertainties

The presence of strong peaks at histograms in Figures 3 means these matches are affected by systematic errors. These systematic errors also affect the distribution of normalized arc lengths shown in Figure 2. In order to mitigate their impact, we re-drew the histogram and excluded the sources with offset position angles with respect to jet direction within 0.5 rad of peaks at 0 and π. As a result, we got a clean sample that is not affected by the systematic errors due to the presence of optical jet. We used this clean sample for characterizing Gaia and VLBI reported position uncertainties. We wanted to answer the question how realsitic the uncertainties are.

We noticed that the number of outliers, i.e. the matches with the normalized arc > 4, grows with an increase of χ^2/ndf, where ndf is the number of degrees of freedom. χ^2 is provided in variable astrometric_chi2_al of the Gaia DR2 archive. The number of degrees of freedom was computed as the difference of the variables astrometric_n_good_obs_al and astrometric_params_solved. We split the dataset into 1% percentiles of χ^2/ndf and computed the share of outliers for each percentile. The dependence of the share of outliers as a function of the mean χ^2/ndf within a percentile is shown with a green curve in Figure 4. It grows approximately as $\sqrt{\chi^2/\mathrm{ndf}}$ when $\chi^2/\mathrm{ndf} > 1.5$–2. Since the number of degrees of freedom is the mathematical expectation of χ^2, in a case if all uncertainties of Gaia observables of a given source are underestimated by a common factor and in the absence of other systematic errors, multiplying the uncertainties in parameter estimates by $\sqrt{\chi^2/\mathrm{ndf}}$ corrects the impact of underestimation of measurements errors. The blue curve in Figure 4 demonstrates that after re-scaling Gaia position uncertainties, the dependence of the number of outliers as a function of χ^2/ndf has disappeared. Scaling position errors by χ^2/ndf makes them larger, which makes the normalized arc-lentght

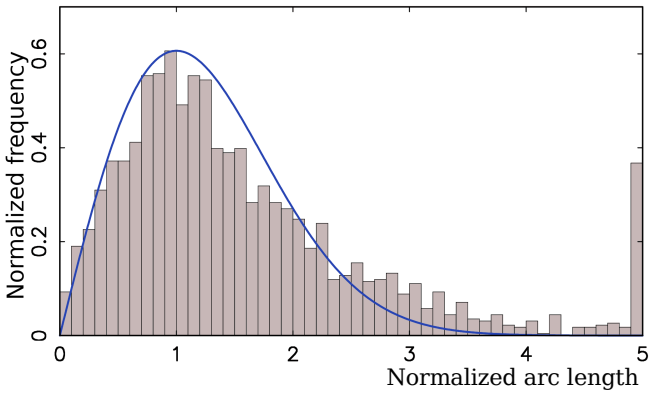

Figure 5. Distribution of normalized VLBI/Gaia arc-lengths over 2313 matching sources. The sample includes all the sources with known jet directions and excludes the sources with $\psi \in [-0.5, -0.5]$ and $\psi \in [\pi - 0.5, \pi + 0.5]$ rad. Scaling factors 1.05 and 1.30 were applied to Gaia and VLBI position uncertainties respectively. Gaia uncertainties were also multiplied by $\sqrt{\chi^2/\mathrm{ndf}}$. The blue smooth curve shows Rayleigh distribution with $\sigma = 1$.

Table 2. Table with the share of matches with normalized residuals > 4 for a number sub-samples in pro cents (column r). The last two rows show the sub-samples of matches with known jet directions. The second and fourth raw use a sub-sample of matches with VLBI semi-major error ellipse less than median among all matches and the matches with known jet directions respectively. Column "off-peak" excludes the sources with $\psi \in [-0.5, -0.5]$ and $\psi \in [\pi - 0.5, \pi + 0.5]$ rad. Column "on-peak" include the sources with ψ in these ranges and exclude everything else.

	all		off-peak		on-peak	
	r	# src	r	# src	r	# src
all	9.0	9042	6.6	7288	19.4	1702
$\sigma_v \leq 0.963$ mas	10.0	4496	5.9	3169	19.7	1323
all with known ψ	11.2	4015	5.4	2313	22.1	1702
$\sigma_v \leq 0.455$ mas	11.4	1997	4.3	1109	20.3	888

smaller. We argue that re-scaling Gaia postion errors makes them more realistic by accounting for the additional noise that increases χ^2/ndf.

In addition to source-dependent re-scaling that is based on χ^2/ndf statistics of a given source, we evaluated global scaling factors for both VLBI and Gaia that affect every source for accounting for both systematic errors and biases in reported uncertainties. This is the simplest way to mitigate the impact of systematic errors on uncertainties and make them more realistic without re-running a solution. Since the normalized arc lengths are affected by both uncertainties of VLBI and Gaia positions, we estimated the scaling factors of VLBI uncertainties by processing the subset of observations that have Gaia position uncertainties by a factor of 5 greater than VLBI uncertainties and vice versus: we estimated scaling factors for Gaia uncertainties (after scaling them by $\sqrt{\chi^2/\mathrm{ndf}}$) by processing the subset of observations with Gaia position uncertainties by a factor of 5 smaller than VLBI uncertainties. We adjusted scaling factors in such a way the distribution of normalized arc-lengths of the subsample be approximated with Rayleigh distribution $\sigma = 1$. The scaling factors are 1.06 for Gaia and 1.30 for VLBI. Applying scaling parameters to adjust uncertainties for accounting for the influence of systematic errors is a common technique. For instance, a scaling factor 1.5 was used to inflate source position uncertainties in the ICRF1 catalogue (Ma et al. 1998).

Since as we have established, the Gaia systematic errors caused by optical structure have a strong concentration towards $\psi = 0$ and $\psi = \pi$, we expect that removal of the matches with $\psi \in [-0.5, -0.5]$ and $\psi \in [\pi - 0.5, \pi + 0.5]$ rad and keeping only "off-peak" matches should affect the statistics of the number of outliers. We computed the share of matches with normalized residuals > 4 for for several subsamples. Since we applied error re-scaling, the number of outliers has reduced with respect to our initial estimate mentioned above. The first row of Table 2 shows that excluding the sources within the peaks of the distribution of ψ angle reduces the number of outliers by a factor of 1.36, but considering only the sources within 0.5 rad of the peaks doubles the number of outliers. Since the jet directions were determined only for 45% of the matches, these statistics underestimate the impact of the systematic errors caused by the presence of optical structure. If to count only the sources with known jet directions, excluding the sources within the peaks reduces the number of outliers by a factor of 2.07. Rows 2 and 4 of Table 2 shows also the statistics for the subsamples of low 50% percentile of VLBI re-scaled errors. The reduction of the number of outliers is 1.77 for the 50% percentile of the overall sample of matching sources and 2.65 for the sub-sample of the sources with known jet directions. The reduction of the number of outliers is greater for the lower 50% percentile because the sources with smaller position uncertainties have smaller errors in determining the ψ angle, what makes discrimination of the "on-peak" and "off-peak" sources more reliable.

Results in Table 2 show that the presence of optical structure aligned along the jet explains 62% VLBI/Gaia position offsets significant at 4σ level for a sub-sample of 23% VLBI/Gaia matches that have known jet direction and VLBI position error lower than the median. In order to generalize this result to the entire population of radio-loud AGN, we need assume that the significance of VLBI/Gaia does not depend on VLBI position error and does not depend on measurability of the radio jet direction. VLBI position errors above 0.2–0.3 mas level are limited by the thermal noise, and thus, the first assumption is valid. The validity of the second assumption is questionable. Detectability of parsec-scale radio jet depends on the jet brightness and the dynamic range of observations that in turn depends on the source flux density. Since the correlation between radio and optical fluxes is low, missing a jet just because the source was weak does not create a selection bias. However, if a jet direction for a given source was not detected because its radio jet is intrinsically weaker, missing such a source may create a selection bias, because a weak radio jet may imply a weak optical jet. A sub-sample of sources with determined jet direction may have a selection bias towards jets brighter in radio and optic with respect to the overall population.

Table 3. Estimates of rotation angles around axes 1,2,3 of the Gaia positions of matches with respect to VLBI positions of four sub-samples. Units are milliarcseconds.

all	9042	-0.030 ± 0.004	0.090 ± 0.004	-0.030 ± 0.005
with jets	4016	-0.010 ± 0.005	0.092 ± 0.005	-0.010 ± 0.006
off-peak	2647	-0.013 ± 0.006	0.095 ± 0.006	0.008 ± 0.007
on-peak	1369	-0.005 ± 0.008	0.091 ± 0.007	-0.037 ± 0.009

2.3 Impact of systematic errors on determination of the orientation of the Gaia catalogue with respect to the VLBI catalogue

Any source catalogue can be rotated at an arbitrary angle and the observables, f.e. group delays, remain the same. The orientation of a catalogue can be described by three angles. These angles for a given catalogue cannot be determined from observation and are *set* by imposing certain conditions. The orientation of the RFC catalogue is set to require the net rotation with respect to 212 so-called "defining" sources in the ICRF1 catalogue (Ma et al. 1998) be zero. Gaia DR2 catalogue was aligned with respect to 2843 matching source the ICRF3-prototype catalogue using the frame rotator technique described in detailed in Lindegren et al. (2012). The systematic differences caused by the optical structure affect the procedure for establishing the catalogue orientation. To provide a quantitative measure of sensitivity of the orientation angles to sampling, we computed the three angles of Gaia DR2 orientation with respect to the RFC VLBI catalogue (See Table 3). We see that selecting different samples, including those the most affected by systematic errors (on-peak) and least affected (off-peak), resulted in differences in orientation angles around 0.02 mas. A large value of the orientation angle around axis 2 is somewhat unexpected, but since the ICRF3-prototype catalogue used for alignment of the Gaia DR2 is not publicly available, the origin of this somewhat large value cannot be established.

3 ANALYSIS OF GAIA PROPER MOTIONS

The Gaia DR2 provides proper motions and parallaxes for 78% sources. Among 9081 matches proper motion estimates are available for 7728 sources. Since the AGNs are located at cosmological distances, their proper motions considered as a bulk tangential motion are supposed to be well below the Gaia detection limit. A flare at the accretion disk or jet will change position of the centroid. A flare will cause a shift in position of the centroid, and therefore, results in a non-zero estimate of proper motion. Such proper motion may be statistically significant even at Gaia level of accuracy. To check it, we made histograms for frequencies of proper motion as a function of the angle of the proper motion direction with respect to the jet direction denoted as $\bar{\psi}$. We analyzed the sample of 613 matching sources with $\sigma(\bar{\psi}) < 0.4$ rad. The histograms showed weak peaks. The peaks become much sharper when we split the sample into two subsets: the subset with χ^2/ndf less then the median 1.125 and the subset with χ^2/ndf greater than the median (See Figure 6).

We see that the subsample of matches with large χ^2/ndf shows two peaks at $\bar{\psi} = 0$ and $\bar{\psi} = \pi$ that are significant, while the subsample with χ^2/ndf below the median does not.

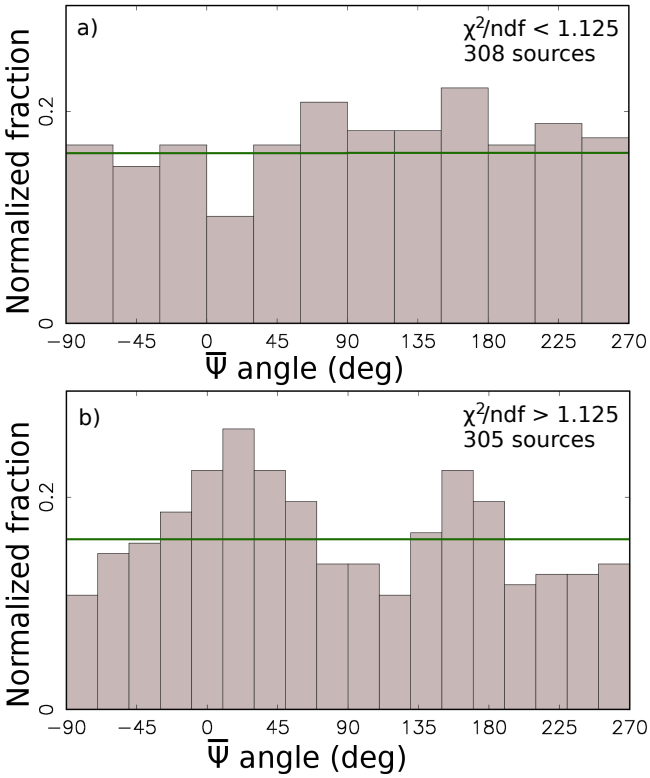

Figure 6. The histograms of the distribution of Gaia proper motion position angle with errors < 0.4 rad among the matches with different χ^2/ndf. *Left figure* uses the matches with χ^2/ndf less than the median in this subsample 1.125. *Right figure* uses the matches with χ^2/ndf greater than the median in this subsample 1.125. For comparison, green line shows the uniform distribution.

A non-linear motion is among the reasons why χ^2/ndf deviates from 1. The histogram at Figure 6-b tells us that among the sources with non-linear motion, the share of objects with proper motions along the jet or in the direction opposite the jet is disproportionately high. This dependence on angle $\bar{\psi}$ means that at least for a fraction of the sources the proper motion is caused by the activity along the jet direction: either at the accretion disk at the jet. Analysis of AGN kinematics from radio observations, for instance, MOJAVE program (Lister et al. 2009) over several decades shows that typically, a core emits components that moves along the jet, sometimes they become brighter at their journey, but later they are getting dimmer, and eventually disappear. Brightening a jet component shifts the centroid temporarily and irregularly. We call this behavior jitter and we predicted it in Petrov & Kovalev (2017b). Unlike to proper motions of stars, extending observations does not result in a convergence of a proper motion estimate to a some value with small uncertainty, but results in a slow convergence to zero. Peaks at 0 and π in the histogram of the number of sources with $\sigma(\bar{\psi})$ as a function of $\bar{\psi}$ at a sub-samples with high χ^2/ndf provides us the first evidence that predicted jitter takes place. We used here estimates of AGN proper motions and χ^2/ndf as a proxy for jitter evaluation.

We explored further the impact of a selection based on χ^2/ndf on the distribution of position offset angles with respect to jet direction. We did not find a noticeable impact of

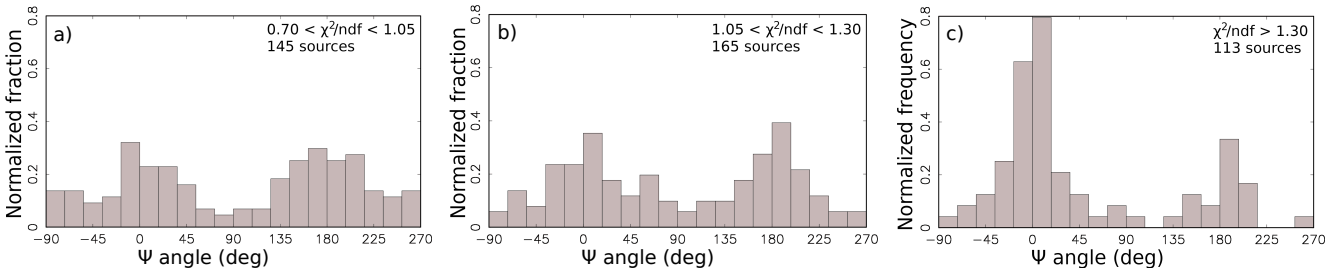

Figure 7. The histograms of the distribution of the position angle of Gaia offset with respect to VLBI position for matches with $\sigma_\psi < 0.3$ rad and arc-lengths < 2.5 mas and different ranges of χ^2/ndf.

χ^2/ndf for VLBI/Gaia offsets longer 2.5 mas, but we found this selection affects matches with VLBI/Gaia offsets shorter 2.5 mas. Figure 7 shows the distributions of ψ angles of matches with $\sigma(\psi) < 0.3$ rad divided into three sub-samples approximately equally distributed over χ^2/ndf. The peaks at $\psi = 0$ and $\psi = \pi$ are broad for the sub-sample of low χ^2/ndf. They are getting sharper for the sub-sample of intermediate χ^2/ndf. The sub-sample with large χ^2/ndf is strikingly different than the sub-sample with low χ^2/ndf: the histogram has a very strong peak at $\psi = 0$, i.e. along the jet direction and a smaller share of matches outside the main peaks.

Analysis of the connection of the Gaia DR2 proper motions with χ^2/ndf suggests that the matches with large χ^2/ndf are more prone to exhibit the jitter. This allows us to tentatively conclude that among the sub-sample of sources with VLBI/Gaia offsets shorter than 2.5 mas flares and jitter occur predominantly in objects that have Gaia offsets along the jet direction. That indicates the mechanism that causes an increase of χ^2/ndf may not work or at least is not dominating for sources with $\psi = \pi$. At the same time, Figure 6 suggests there is no strong preferable sign of the motion direction, either along or opposite to the jet. Such a pattern is consistent with a jitter caused by flares: depending when a flare has happened, at the beginning or the end of observing interval, the direction of the proper motion may be opposite.

4 SUMMARY AND CONCLUSIONS

Here we summarize our main results of our comparison of AGN positions and proper motions the Gaia DR2 against the most complete catalogue of VLBI positions to date, the RFC.

(i) The Gaia DR2 AGN position uncertainties of VLBI matching sources are a factor of two smaller than VLBI position uncertainties. The VLBI position catalogues have lost their superiority in accuracy with respect to space astrometry.

(ii) We have predicted in Petrov & Kovalev (2017b) that improvement in accuracy of VLBI and/or Gaia will not reconcile VLBI and Gaia positions, but will make these differences more significant. This prediction has come true. The share of outliers grew from 6 to 9%, the histogram of the position offset directions as a function of ψ angle became sharper.

(iii) We have established the main reason of statistically significant VLBI/Gaia position offset is the presence of optical structure. Among the sources with measured jet direction, 52–62% matching sources with the normalized arc length exceeding 4, i.e. *the majority*, is due to the presence of optical structure. Although this share may be somewhat lower for the entire population of matching AGNs, we got its firm lower limit: 27%. Other reasons mentioned by Mignard et al. (2018) can explain only a small fraction of outliers. Error in matching of VLBI and Gaia objects are easily controlled by computing the probability of false association based on source density in the vicinity of candidates to association. The cutoff of the probability of false association $2 \cdot 10^{-4}$ results in the mathematical expectation of the total number of false association to 2. The coarseness of the model of source density may increase the number of false associations, but very unlikely can increase their count by an order of magnitude. Double and/or lensed AGNs are easily identified in radio images and can be thoroughly accounted. The presence of emission from hosting galaxy within Gaia point spread function of an AGN may shift the centroid with respect to the nucleus if the galaxy central region structure is asymmetric or the AGN is dislodged with respect to the galaxy center of mass, but such a shift is independent on ψ angle. Table 1 provides the upper limit of the share of outliers which position offset does not depend on ψ, 1/3. It does not seem likely that all of these offsets are caused by the contribution of hosting galaxies, because the share of AGNs with discernible hosting galaxies is much less than that.

We found that scaling Gaia position uncertainties by $\sqrt{\chi^2/\mathrm{ndf}}$ eliminated the dependence of the share of the number of outliers on χ^2/ndf. Examining the subset of matches with dominating VLBI or Gaia errors allowed us to evaluate scaling factors for VLBI uncertainties, 1.30, and Gaia position uncertainties: $1.06\sqrt{\chi^2/\mathrm{ndf}}$. Eliminating the observations within 0.5 rad of $\psi = 0$ and $\psi = \pi$ and using re-scaled uncertainties, made the distribution of normalized VLBI/Gaia arc-lengths much closer to the Rayleigh distribution: compare Figures 2 and 5.

(iv) Influence of VLBI and/or Gaia systematic errors on estimates of the orientation angles of the Gaia DR2 catalogue with respect the VLBI catalogue does not exceed 0.02 mas.

(v) We have predicted in Petrov & Kovalev (2017b) that flares in AGNs will cause a jitter in AGN positions. Analysis of Gaia proper motions provided us an indirect confirmation of this prediction: the sources with excessive Gaia residuals, i.e. large χ^2/ndf, have proper motion direction aligned with the jet direction. Although AGNs proper motions should

not be interpreted as a bulk tangential motion, at the same time, these proper motions are not always an artifacts of Gaia data analysis, but some of them are real. The photocenters of at least some quasars are not unmovable points and the possibility of quasar proper motion should be taken into account in interpreting results of differential astrometry.

We do not claim that we have solved the problem of establishing the nature of *all* outliers. The distribution in Figure 5 still deviates from Rayleigh and we still did not uncover the nature of 1/3 outliers, but we made a quite substantial progress.

ACKNOWLEDGMENTS

We used un our work the Astrogeo VLBI FITS image database[2] that contains radio images contributed by A. Bertarini, L. Garcia, N. Corey, Y. Cui, L. Gurvits, X. He, Y. Y. Kovalev, S. Lee, R. Lico, E. Liuzzo, A. Marsher, S. Jorstad, C. Marvin, D. Homan, Y. Kovalev, M. Lister, A. Pushkarev, E. Ros, T. Savolainen, L. Petrov, A. Pushkarev, K. Sokolovski, G. Taylor, A. de Witt, M. Xu, and B. Zhang.

This project is supported by the Russian Science Foundation grant 16-12-10481. This work has made use of data from the European Space Agency (ESA) mission Gaia[3], processed by the Gaia Data Processing and Analysis Consortium (DPAC[4]). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement. We used in our work VLBA data provided by the Long Baseline Observatory that is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

REFERENCES

Frouard J., Johnson M. C., Fey A., Makarov V. V., Dorland B. N., 2018, AJ, 155, 229
Gabuzda D. C., Rastorgueva E. A., Smith P. S., O'Sullivan S. P., 2006, MNRAS, 369, 1596
Kovalev Y. Y., Petrov L., Plavin A. V., 2017, A&A, 598, L1
Lindegren L., Lammers U., Hobbs D., O'Mullane W., Bastian U., Hernández J., 2012, A&A, 538, A78
Lindegren L., et al., 2016, A&A, 595, A4
Lindegren L., et al., 2018, preprint, (arXiv:1804.09366)
Lister M. L., et al., 2009, AJ, 138, 1874
Liu J.-C., Zhu Z., Liu N., 2018a, AJ, 156, 13
Liu J.-C., Malkin Z., Zhu Z., 2018b, MNRAS, 474, 4477
Liu N., Zhu Z., Liu J.-C., 2018c, A&A, 609, A19
Ma C., et al., 1998, AJ, 116, 516
Makarov V. V., Frouard J., Berghea C. T., Rest A., Chambers K. C., Kaiser N., Kudritzki R.-P., Magnier E. A., 2017, ApJ, 835, L30
Meyer E. T., Petropoulou M., Georganopoulos M., Chiaberge M., Breiding P., Sparks W. B., 2018, ApJ, 860, 9
Mignard F., et al., 2016a, preprint, (arXiv:1609.07255)
Mignard F., et al., 2016b, A&A, 595, A5
Mignard F., Klioner S., Lindegren L., Hernandez J., Bastian U., Bombrun A., 2018, preprint, (arXiv:1804.09377)
Perlman E. S., et al., 2010, ApJ, 708, 171
Petrov L., Kovalev Y. Y., 2017a, MNRAS, 467, L71
Petrov L., Kovalev Y. Y., 2017b, MNRAS, 471, 3775
Thompson A. R., Moran J. M., Swenson Jr. G. W., 2017, Interferometry and Synthesis in Radio Astronomy, 3rd Edition. Springer, doi:10.1007/978-3-319-44431-4

This paper has been typeset from a TEX/LATEX file prepared by the author.

[2] Available at http://astrogeo.org/vlbi_images
[3] https://www.cosmos.esa.int/gaia
[4] https://www.cosmos.esa.int/web/gaia/dpac/consortium