Action-based Milky Way Disk Modelling with *RoadMapping* and our imperfect Knowledge of the "Real World"

W. Trick[1,2], J. Bovy[3,4], and H.-W. Rix[1]

trick@mpia.de

0.1. Effect of measurement errors on recovery of potential?

[TO DO]

Collection tests and plots (tests are still running on the cluster)

- *Plot 1:* number of MC samples needed for the error convolution vs. maximum velocity error inside the observed volume, such that a given accuracy in potential and qDF parameters is reached. Similar to what I had on the poster. However, we still haven't tested, if this plot depends on: hotness of stars and or umber of stars.
- *Plot 2:* 2 columns of panels (one row for each parameter), bias vs. standard error. First column: only proper motion and vlos errors $\longrightarrow$ shows that our error convolution works and should be bias free, plus, when knowing the errors perfectly we can get a perfect deconvolution and tight constraints. Second column: proper motion, vlos and distance modulus errors $\longrightarrow$ shows that for too large proper motion and distance errors our approximation for the error convolution does not work anymore.

Convergence of the error integral. In §**??** we introduced how we convolve the model probability with the measurement errors. In the absence of distance errors the accuracy of the parameter recovery is limited by an insufficient MC sampling of the convolution integral in Equation (**??**). Test 6 in Table 3 and Figure 1 investigate how many MC samples are needed, given the size of the velocity error, for the integral to be accurate within certain

[1]Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

[2]Correspondence should be addressed to trick@mpia.de.

[3]University of Toronto [TO DO: What is Jo's current address???]

[4]Hubble fellow

limits:
For each $\delta\mu \in [2,3,4,5]$mas yr^{-1} we set up four mock data sets and evaluated the likelihood for different N_{error}. We used $N_{\mathrm{conv}} := 800$ and 1200 MC samples to calculate the numerically converged likelihood for proper motion errors $\delta\mu \leq 3$mas yr^{-1} and $\delta\mu > 3$mas yr^{-1}, respectively. We determined the mean bias

$$\mathrm{BIAS}(N_{\mathrm{error}}, \delta\mu) \equiv \frac{1}{4}\sum_{j=1}^{4} \left[\langle p_i \rangle (N_{\mathrm{error}}, \delta\mu)\right]_j - \left[\langle p_i \rangle (N_{\mathrm{conv}}, \delta\mu)\right]_j ,$$

where $\left[\langle p_i \rangle (N_{\mathrm{error}}, \delta\mu)\right]_j$ is the best estimate for the i-th model parameter $p_i \in p_M$ from the analysis of the j-th mock data realisation with $\delta\mu$ using N_{error} MC samples. From this we then generated the curves $N_{\mathrm{error},i}(\delta v_{\mathrm{max}}, \mathrm{BIAS})$ in Figure 1 by linear interplolation, that show how many MC samples are needed for parameter p_i given a velocity error and a systematic bias in units of the standard error (SE) of the estimate. The proper motion error $\delta\mu$ translates to a velocity error according to

$$\delta v_{\mathrm{max}}[\mathrm{km\ s}^{-1}] \equiv 4.74047 \cdot r_{\mathrm{max}}[\mathrm{kpc}] \cdot \delta\mu[\mathrm{mas\ yr}^{-1}], \tag{1}$$

where r_{max} is the maximum distance of stars. We find in Figure 1 the relation

$$N_{\mathrm{error},i}(\delta v_{\mathrm{max}}, \mathrm{BIAS}) \propto (\delta v_{\mathrm{max}})^2 .$$

Figure 1 also demonstrates that different model parameters do not have the same sensitivity to the numerical inaccuracies introduced by insufficient sampling.

Underestimation of the proper motion error. We found that in case we perfectly knew the measurement errors (and the distance error is negligible), the convolution of the model probability with the measurement errors gives precise and accurate constraints on the model parameters - even if the error itself is quite large. Now we investigate what would happen if the quoted measurement errors, e.g. the proper motion errors, were actually smaller than the true errors. Figure 3 shows the case for two different stellar populations and an error underestimation of 10% and 50%.
Overall the parameter recovery gets worse the larger the proper motion error and the stronger the underestimation. The relation between the bias due to error misjudgment and the size of the proper motion error seems to be linear.
For the recovery of the isochrone potential scale length b the hotness of the population does not matter (see lower left panel in Figure 3). The circular velocity $v_{\mathrm{circ}}(R_\odot)$ is, as always, better measured by cooler than by hotter populations (see upper left panel in Figure 3).
We find that the recovery of the qDF parameters on the other hand is more strongly affected

by the misjudgment of the velocity error for *cooler* stellar popluations. The measured velocity dispersion is the convolution of the intrinsic dispersion with the measurement errors. If the proper motion error is underestimated, the deconvolved velocity dispersion is larger than the intrinsic velocity dispersion and the relative difference is bigger for a cooler population (see upper right panel for σ_z in Figure 3). The intrinsic velocity dispersion is also cooler at larger radii than at smaller radii, therefore the deconvolved dispersion is overestimated more strongly at large R and the velocity dispersion scale length will be overestimated as well (see lower left panel for h_{σ_z} in Figure 3). We get analogous results for the qDF parameters σ_R and h_{σ_R}. The recovery of the tracer density scale length h_R is not affected by the misjudgment of velocity errors.

The most important and encouraging result from Figure 3 is, that for an underestimation of 10% the bias is still $\lesssim 2\sigma$ - even for proper motion errors of 3 mas/yr.

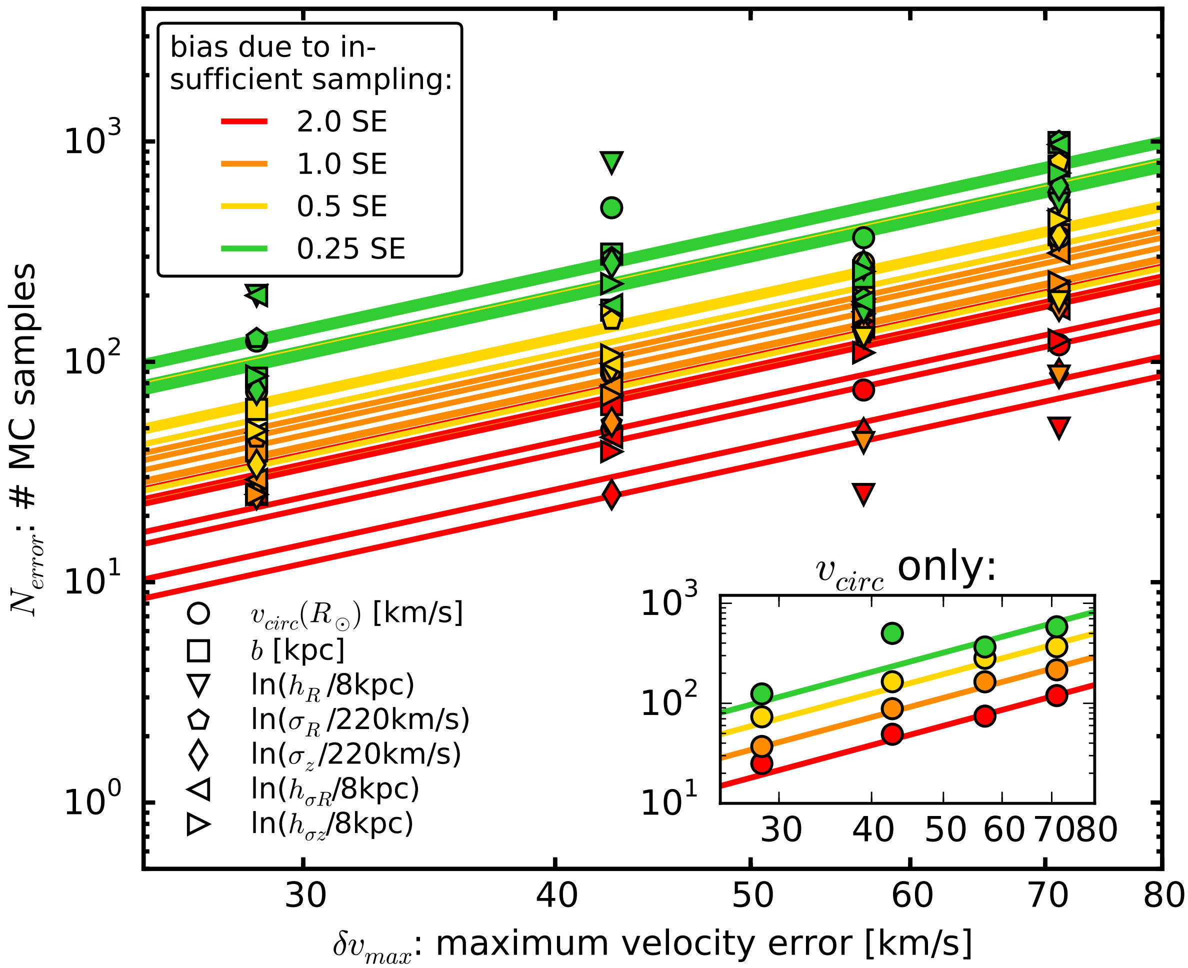

Fig. 1.— Number of Monte Carlo (MC) samples N_{error} needed for the numerical error convolution in Equation (**??**), given the maximum velocity error $\delta v_{\max}$ in the sample to reach a given accuracy. An insufficient sampling of the convolution integral leads to systematic biases in the reconstruction of the true model parameters. The size of the bias is color coded as indicated in the legend and is given in units of the standard error (SE). The model parameters, marked by different symbols, have different sensitivities to the numerical inaccuracy of the error convolution, therefore the range in N_{error} for the same given bias. Here we assume that the distance error is zero and the proper motion error $\delta\mu$ translates to a velocity error according to Equation (1) and $\delta v_{\text{los}} \ll \delta v_{\max}$. All model parameters are listed in Table 3 as Test ⑥. The number of MC samples needed increases with the velocity error as $N_{\text{error}} \propto (\delta v_{\max})^2$, as can be seen especially well in the inset figure for the potential parameter $v_{\text{circ}}(R_\odot)$. All lines are fits of this functional form to each four points derived for a given model parameter (symbol) and bias (color). The large scatter in the points comes from low number statistics and errors introduced by linear interpolation of the bias vs. N_{error} relation found from the analyses.

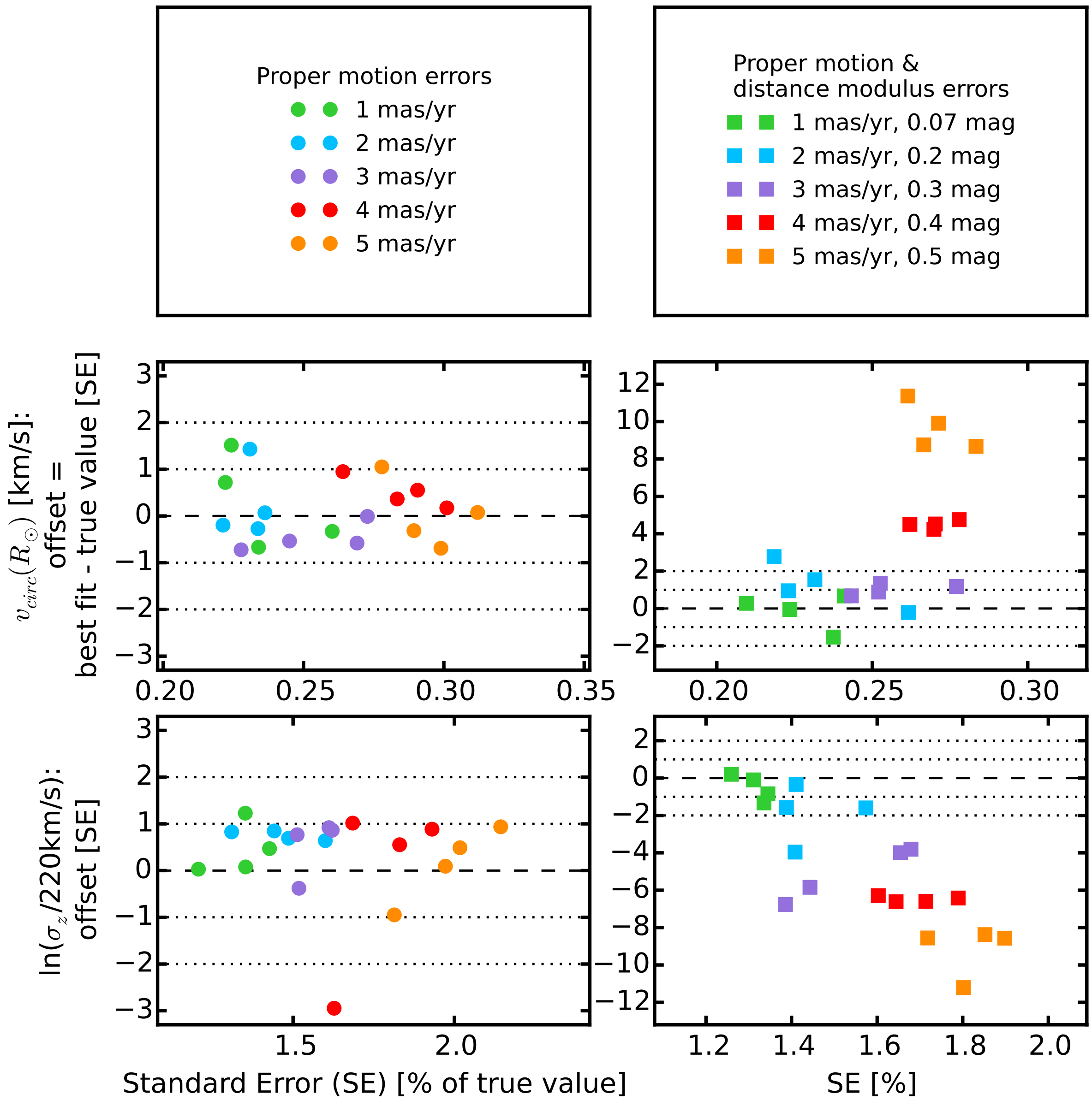

Fig. 2.— [TO DO: Caption]

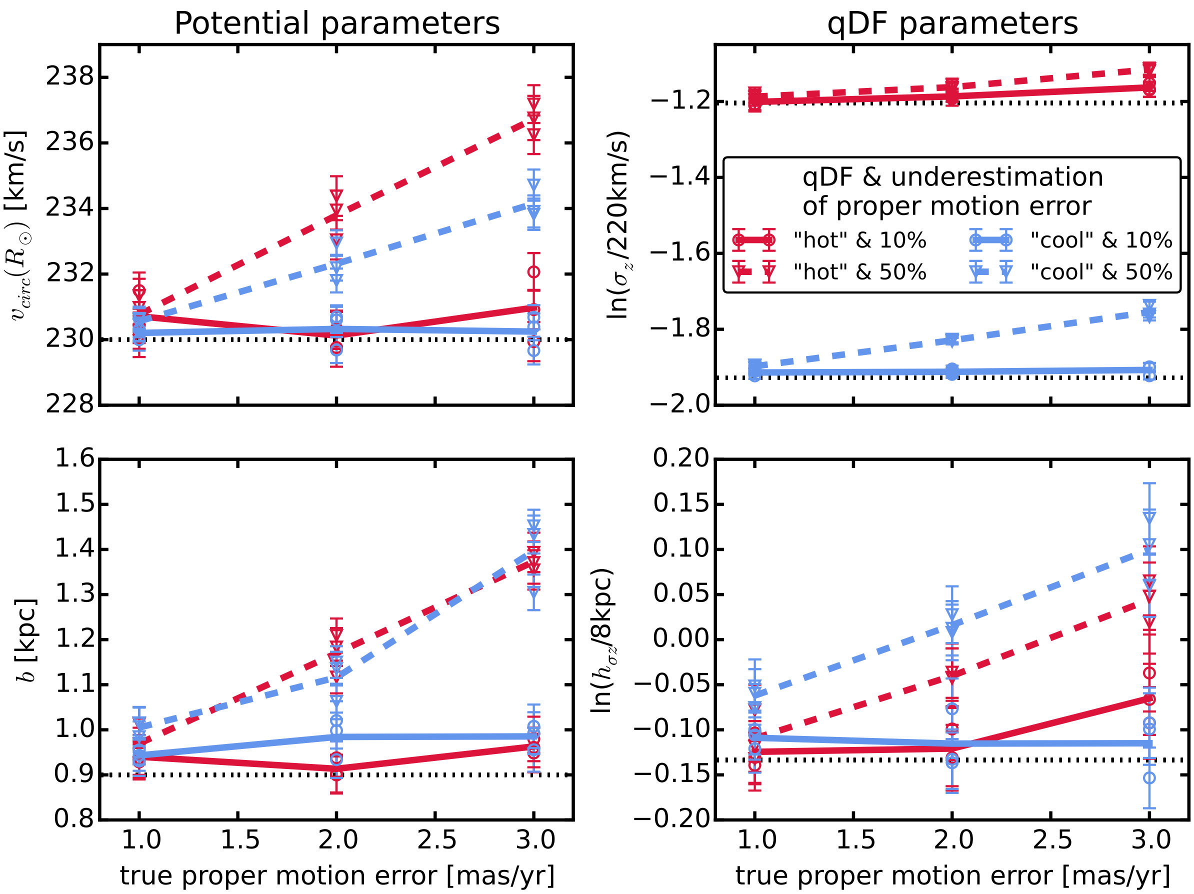

Fig. 3.— Effect of an systematic underestimation of proper motion errors in the recovery of the model parameters. The true model parameters used to create the mock data are summarized as Test (1) in Table 3, four of them are given on the y-axes and the true values are indicated as black dashed lines. The velocities of the mock data were perturbed according to Gaussian errors in the α and δ proper motions as indicated on the x-axis. The circles and triangles are the best fit parameters of several mock data set assuming the proper motion error, with which the model probability was convolved, was underestimated in the analysis by 10% or 50%, respectively. The error bars correspond to 1σ confidence. The lines connect the mean of each two data realisations and are just guides to the eyes.

Table 1. Gravitational potentials of the reference galaxies used troughout this work and the respective ways to calculate actions in these potentials. All four potentials are axisymmetric. The potential parameters are fixed for the mock data creation at the values given in this table. In the subsequent analyses we aim to recover these potential parameters again. The parameters of "MW13-Pot" and "KKS-Pot" were found as direct fits to the "MW14-Pot".

name	potential type	potential parameters p_Φ		action calculation	reference for potential type
"Iso-Pot"	isochrone potential	circular velocity at the sun isochrone scale length	$v_{\text{circ}} = 230$ km s^{-1} $b = 0.9$ kpc	***analytical and exact*** J_r, J_ϑ, L_z; use $J_r \rightarrow J_R, J_\vartheta \rightarrow J_z$ in Equation (**??**)	**?**
"KKS-Pot"	2-component Kuzmin-Kutuzov- Stäckel potential (disk + halo) (analytic potential)	circular velocity at the sun focal distance of coordinate system[a] axis ratio of the coordinate surfaces[a] ... ...of the disk component ...of the halo component relative contribution of the disk mass to the total mass	$v_{\text{circ}} = 230$ km s^{-1} $\Delta = 0.3$ $\left(\frac{a}{c}\right)_{\text{Disk}} = 20$ $\left(\frac{a}{c}\right)_{\text{Halo}} = 1.07$ $k = 0.28$	***exact*** J_R, J_z, L_z using "Stäckel Fudge" (**?**) and interpolation on action grid[b] (**?**)	**?**
"MW13-Pot"	MW-like potential with Hernquist bulge, 2 exponential disks (stars + gas), spherical power-law halo (interpolated potential)	circular velocity at the sun stellar disk scale length stellar disk scale height relative halo contribution to $v^2_{\text{circ}}(R_\odot)$ "flatness" of rotation curve	$v_{\text{circ}} = 230$ km s^{-1} $R_d = 3$ kpc $z_h = 0.4$ kpc $f_h = 0.5$ $\frac{\mathrm{d}\ln(v_{\text{circ}}(R_\odot))}{\mathrm{d}\ln(R)} = 0$	***approximate*** J_R, J_z, L_z using "Stäckel Fudge" (**?**) and interpolation on action grid[a] (**?**)	**?**
"MW14-Pot"	MW-like potential with cut-off power-law bulge, Miyamoto-Nagai stellar disk, NFW halo	-	-	***approximate*** J_R, J_z, L_z (see "MW13-Pot")	**?**

[a]The coordinate system of each of the two Stäckel-potential components is $\frac{R^2}{\tau_{i,p}+\alpha_p} + \frac{z^2}{\tau_{i,p}+\gamma_p} = 1$ with $p \in \{\text{Disk}, /\text{Halo}\}$ and $\tau_{i,p} \in \{\lambda_p, \nu_p\}$. Both components have the same focal distance $\Delta = \sqrt{\gamma_p - \alpha_p}$, to make sure that the superposition of the two components itself is still a Stäckel potential. The axis ratio of the coordinate surfaces $\left(\frac{a}{c}\right)_p := \sqrt{\frac{\alpha_p}{\gamma_p}}$ describes the flattness of the corresponding Stäckel component.

[b]We use a finely spaced action interpolation grid with $R_{\max} = 10$ [TO DO: What's that??? units???] and 50 grid points in E and ψ [TO DO: Find out what's that???], and 60 grid points in L_z. [TO DO: more details?]

Table 2. Reference distribution function parameters for the qDF in Equations (**??**)-(**??**). These qDFs describe the phase-space distribution of stellar *MAP*s for which mock data is created and analysed throughout this work for testing purposes. The parameters of the "cooler" & "colder" ("hotter" & "warmer") *MAP*s were chosen such, that the they have the same σ_R/σ_z ratio as the "hot" ("cool") *MAP* . The "colder" and "warmer" *MAP*s have a free parameter X that governs how much colder/warmer they are then the reference "hot" and "cool" qDFs. Hotter populations have shorter tracer scale lengths (**?**) and the velocity dispersion scale lengths were fixed according to **?**.

name of *MAP*	qDF parameters p_{DF}				
	h_R [kpc]	σ_R [km s^{-1}]	σ_z [km s^{-1}]	h_{σ_R} [kpc]	h_{σ_z} [kpc]
"hot"	2	55	66	8	7
"cool"	3.5	42	32	8	7
"cooler"	2 +50%	55-50%	66-50%	8	7
"hotter"	3.5-50%	42+50%	32+50%	8	7
"colder"	2 +X%	55-X%	66-X%	8	7
"warmer"	3.5-X%	42+X%	32+X%	8	7

Table 3. Summary of test suites in this work: The first column indicates the test suite, the second column the potential, DF and selection function model etc. used for the mock data creation, the third model the corresponding model assumed in the analysis, and the last column lists the figures belonging to the test suite. Parameters that are not left free in the analyis, are always fixed to their true value. Unless otherwise stated we calculate the likelihood by the nested-grid and MCMC approach outlined in §**??** and use $N_{\text{spatial}} = 16$, $N_{\text{velocity}} = 24$, $N_{\text{sigma}} = 5$ as numerical accuracy for the likelihood normalisation in Equations (**??**) and (**??**). [TO DO: Change encircled numbers to proper order. Make sure the plot references are the right ones.]

Test		Model for Mock Data	Model in Analysis	Figures
①	*Potential:*	"KKS-Pot"	-	Mock data:
Influence of	*MAP :*	2 *MAP*s "hot" or "cold" qDF		Figure **??**
survey volume on	*Survey volume:*	a) $R \in [4, 12]$ kpc,$z \in [-4, 4]$ kpc, $\phi \in [-20°, 20°]$.		
mock data distribution,		b) $R \in [6, 10]$ kpc,$z \in [1, 5]$ kpc, $\phi \in [-20°, 20°]$.		
also in action space	*# stars per data set:*	20,000		
	# data sets:	4 (= 2×2 models)		
⑨	*Potential:*	"Iso-Pot", "MW13-Pot" & "KKS-Pot"	-	Convergence
Numerical accuracy	*MAP :*	"hot" qDF		of normalisation:
in calculation	*Survey volume:*	sphere around sun, $r_{\max} = 0.2, 1, 2, 3$ or 4 kpc		Figure **??**
of the likelihood	*Numerical accuracy:*	$N_{\text{spatial}} \in [5, 20]$, $N_{\text{velocity}} \in [6, 40]$, $N_{\text{sigma}} \in [3.5, 7]$		
normalisation				
⑩	*Potential:*	"Iso-Pot"	"Iso-Pot", all parameters free	Figure **??**
pdf is a	*MAP :*	"hot" qDF	qDF, all parameters free	
multivariate	*Survey Volume:*	sphere around sun, $r_{\max} = 2$ kpc	(fixed & known)	
Gaussian	*# stars per data set:*	20,000		
for large data sets.	*# data sets:*	5 (only one is shown)		
	Numerical accuracy:		$N_{\text{velocity}} = 20$ and $N_{\text{sigma}} = 4$	
②	*Potential:*	"Iso-Pot"	"Iso-Pot", free parameter: b	Figure **??**
Width of the	*MAP :*	"hot" qDF	"hot" qDF, free parameters:	
likelihood scales			$\ln\left(\frac{h_R}{8\text{kpc}}\right), \ln\left(\frac{\sigma_R}{230\text{km s}^{-1}}\right), \ln\left(\frac{h_{\sigma,R}}{8\text{kpc}}\right)$	
with number of stars	*Survey volume:*	sphere around sun, $r_{\max} = 3$ kpc	(fixed & known)	
by $\propto 1/\sqrt{N}$.	*# stars per data set:*	between 100 and 40,000		
	# data sets:	132		
	Analysis method:		likelihood on grid	
	Numerical accuracy:		$N_{\text{velocity}} = 20$ and $N_{\text{sigma}} = 4$ (for speed)	
③	*Potential:*	2 "Iso-Pot" with	"Iso-Pot", free parameter: b	Figure **??**
Parameter estimates		$b = 0.8$ kpc or $b = 1.5$ kpc		
are unbiased.	*MAP :*	2 *MAP*s , "hot" or "cool" qDF	"hot"/"cool" qDF, free parameters:	
			$\ln\left(\frac{h_R}{8\text{kpc}}\right), \ln\left(\frac{\sigma_R}{230\text{km s}^{-1}}\right), \ln\left(\frac{h_{\sigma,R}}{8\text{kpc}}\right)$	
	Survey volume:	5 spheres around sun, $r_{\max} = 0.2, 1, 2, 3$ or 4 kpc	(fixed & known)	
	# stars per data set:	20,000		
	# data sets:	640 (= $2 \times 2 \times 5$ models $\times 32$ realisations)		
	Analysis method:		likelihood on grid	
	Numerical accuracy:		$N_{\text{velocity}} = 20$ and $N_{\text{sigma}} = 4$ (for speed)	
④	*Potential:*	i) "Iso-Pot", ii) "MW13-Pot" or iii) "KKS-Pot"	i) "Iso-Pot", all parameters free	Figure **??**

Table 3—Continued

Test		Model for Mock Data	Model in Analysis	Figures
Influence of			ii) ”MW13-Pot”, R_d and f_h free	
position & shape			iii) ”KKS-Pot”, all free except $v_{\mathrm{circ}}(R_\odot)$	
of survey volume	*MAP :*	”hot” qDF	i) & iii) qDF, all parameters free	
on parameter recovery			ii) qDF, only h_R, $\sigma_{z,0}$ and h_{σ_R} free	
	Survey volume:	4 different wedges, see Figure **??**, upper right panel	(fixed & known)	
	# of stars per data set:	20,000		
	# data sets:	48 (= 4×3 models $\times 4$ realisations)		
	Analysis method:		i) & ii) MCMC, iii) likelihood on grid	
	Action calculation:	ii) & iii) low accuracy	(same as mock data creation)	
		”Stäckel Fudge” grid (**?**) for speed		
		(# grid points: 25 in each E and ψ,		
		30 in L_z, $R_{\max} = 5$		
		[TO DO: What is psi and Rmax (units)?])		
5	*Potential:*	”Iso-Pot”	”Iso-Pot”, all parameters free	Illustration & mock data:
Influence of	*MAP :*	2 *MAPs* , a) ”hot” or b) ”cool” qDF	qDF, all parameters free	Figures **??** & **??**
wrong assumptions	*Survey volume:*	sphere around sun, $r_{\max} = 3$ kpc	(fixed & known)	Analysis results:
about the data set	*Completeness:*	*Example 1:* radial incompleteness,	data set complete,	Figures **??** & **??**
(in-)completeness		completeness$(r) = 1 - \epsilon_r \frac{r}{r_{\max}}$, twenty $\epsilon_r \in [0, 0.7]$	completeness$(r) = 1$, $\epsilon_r = 0$	Analysis results:
on parameter recovery		$r \equiv$ distance from sun,		when not using v_T data:
		Example 2: planar incompleteness,	data set complete,	Figure **??**
		completeness$(z) = 1 - \epsilon_z \frac{\lvert z \rvert}{r_{\max}}$, $\epsilon_r \in [0, 0.7]$,	completeness$(r) = 1$, twenty $\epsilon_z = 0$	
		$z \equiv$ distance from Gal. plane.		
	# stars per data set:	20,000		
	# data sets:	40 (= $2 \times 2 \times 20$)		
6	*Potential:*	”Iso-Pot”	”Iso-Pot, all parameters free”	Figure 1
Numerical convergence	*MAP :*	”hot” qDF	qDF, all parameters free	
of convolution	*Survey Volume:*	sphere around sun, $r_{\max} = 3$ kpc	(fixed & known)	
with measurement	*Errors:*	$\delta\mathrm{RA} = \delta\mathrm{DEC} = \delta(m - M) = 0$	Convolution with	
errors		$\delta v_{\mathrm{los}} = 2$ km/s	perfectly known errors	
		$\delta\mu_{\mathrm{RA}} = \delta\mu_{\mathrm{DEC}} =$ 2,3,4 or 5 mas/yr		
	Numerical Accuracy:		convolution using MC integration	
			with between 25 and 1200 MC samples	
	# stars per data set:	10,000		
	# data sets:	16 (= 4×4 realisations)		
12 Testing the	*Potential:*	”Iso-Pot”	”Iso-Pot, all parameters free”	Figure 2

Table 3—Continued

Test		Model for Mock Data	Model in Analysis	Figure
convolution	*MAP :*	"hot" qDF	qDF, all parameters free	
with measurement	*Survey Volume:*	sphere around sun, $r_{\max} = 3$ kpc	(fixed & known)	
errors with & without	*Errors:*	$\delta\mathrm{RA} = \delta\mathrm{DEC} = 0$	Convolution with errors,	
distance errors		$\delta v_{\mathrm{los}} = 2$ km/s	ignoring distance errors in position (see §[TO DO: CHECK???])	
		$\delta\mu_{\mathrm{RA}} = \delta\mu_{\mathrm{DEC}} = 1$, 2,3,4 or 5 mas/yr		
		a) $\delta(m-M) = 0$, b) $\delta(m-M) \neq 0$ (see Figure 2)		
	Numerical Accuracy:		800 or 1200 MC samples	
	# stars per data set:	10,000		
	# data sets:	40 ($= 2 \times 5 \times 4$ realisations)		
11	*Potential:*	"Iso-Pot"	"Iso-Pot", all parameters free	Figure 3
Underestimation	*MAP :*	"hot" or "cool" qDF	qDF, all parameters free	
of proper motion	*Survey volume:*	sphere around sun, $r_{\max} = 3$ kpc [TO DO: CHECK]	(fixed & known)	
errors	*Errors:*	only proper motion errors	Convolution with proper motion errors	
		1, 2 or 3 mas/yr	10% or 50% underestimated	
	# stars per data set:	10,000		
	# data sets:	24 ($= 2 \times 2 \times 3 \times 3$ realisations)		
7	*Potential:*	"Iso-Pot"	"Iso-Pot", all parameters free	mock data:
Deviations in the	*MAP :*	mix of two qDFs	single qDF, all parameters free	Figure **??**
assumed DF		*Example 1:* with fixed qDF parameters,		Analysis res
from the		but 20 different mixing rates:		**??** & Figure
star's true DF		a) "hot" & "cooler" qDF or b) "cool" & "hotter" qDF		
		Example 2: 20 fixed 50/50 mixtures,		
		with varying qDF parameters (by X%):		
		a) "hot" & "colder" qDF or b) "cool" & "warmer" qDF		
	Survey volume:	sphere around sun, $r_{\max} = 2$ kpc	(fixed & known)	
	# stars per data set:	20,000		
	# data sets:	40 ($= 2 \times 2 \times 20$)		
8	*Potential:*	"MW14-Pot"	"KKS-Pot", all parameters free,	potential co
Deviations of the			only $v_{\mathrm{circ}}(R_\odot) = 230$km s^{-1} fixed	Figure **??**
assumed potential model	*MAP :*	"hot" or "cool" qDF	qDF, all parameters free	qDF recover
from the star's	*Survey volume:*	sphere around sun, $r_{\max} = 4$ kpc	(fixed & known)	Figure **??**
true potential	*# stars per data set:*	20,000		
	# data sets:	2		