Gaussian Process Model for the Local Stellar Velocity Field from Gaia Data Release 2

Patrick Nelson$^1$ ♠ and Lawrence M. Widrow$^1$ †

$^1$Department of Physics, Engineering Physics, and Astronomy, Queen’s University, Kingston, K7L 3N6, Canada

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We model the local stellar velocity field using position and velocity measurements for 4M stars from the second data release of Gaia. We determine the components of the mean or bulk velocity in ~ 27k spatially-defined bins. Our assumption is that these quantities constitute a Gaussian process where the correlation between the bulk velocity at different locations is described by a simple covariance function or kernel. We use a sparse Gaussian process algorithm based on inducing points to construct a non-parametric, smooth, and differentiable model for the underlying velocity field. We estimate the Oort constants $A$, $B$, $C$, and $K$ and find values in excellent agreement with previous results. Maps of the velocity field within 2 kpc of the Sun reveal complicated substructures, which provide clear evidence that the local disk is in a state of disequilibrium. We present the first 3D map of the divergence of the stellar velocity field and identify regions of the disk that may be undergoing compression and rarefaction.

Key words: Galaxy: kinematics and dynamics – Galaxy: structure

1 INTRODUCTION

A common strategy for understanding the dynamics of the Milky way is to assume that phenomena such as the bar, spiral arms, and warp can be understood as perturbative departures from an equilibrium state (See, for example, Binney (2013); Sellwood (2013)). A further assumption is that stars orbit in the mean gravitational field of gas, dark matter, and the other stars. The stellar components of the Galaxy are then described by a phase space distribution function (DF), $f(x, v, t)$, which obeys the collisionless Boltzmann equation (CBE) coupled to gravity via Poisson’s equation. The equilibrium/perturbation split then carries over to the DF and gravitational potential $\Psi$. For a disk galaxy such as the Milky Way, the equilibrium state exhibits symmetry about both the its midplane and rotation axis. Any departure from these symmetries therefore signals a departure from equilibrium.

The second Gaia data release (Gaia DR2, Gaia Collaboration et al. (2018)) includes measurements of positions and velocities for over seven million stars thereby vastly increasing the number of stars for which all phase space coordinates are known. This data provides an estimate of the present-day stellar DF near the Sun in the sense that each star can be represented as a phase space probability distribution function $P_t$ that reflects observational uncertainties in its position and velocity:

$$f(x, v, t) = \sum_i P_t(x - x_i, v - v_i).$$

In the limit of perfect observations, $P_t$ reduces to a six-dimensional delta function. In general, interpretation of the DF involves the construction of a set of observables

$$O = \int d^3x' d^3v' f(x', v', t) F(x', v')$$

where $F$ is designed to pick out particular features of the DF and also account for the selection function of the survey. For example, an estimate of the number density $n(x)$ is obtained by setting $F = W(x - x')$ where $W$ is a localized window function:

$$n(x) = \int d^3x' d^3v' f(x', v', t) W(x - x').$$

In this paper, we focus on the first velocity moment of the DF:

$$V(x) = n^{-1} \int d^3x' d^3v' f(x', v', t) W(x - x') v.$$

For a system in equilibrium, $V$ is a function of Galactocentric radius $R$ and distance from the midplane $|z|$. Studies of the bulk velocity in pre-Gaia surveys such as the Sloan Extension for Galactic Understanding and Exploration (SEGUE) (Yanny et al. 2009), the Radial Velocity Experiment (RAVE) (Steinmetz et al. 2006), and Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) (Cui et al. 2012) revealed bulk motions that signaled a Galaxy in disequilibrium. For example, Widrow et al. (2012) determined the bulk vertical velocity $V_z$ and its dispersion $\sigma_z$ as functions of $z$ for SEGUE stars and found that they showed features asymmetric about $z = 0$. Williams et al. (2013) determined the full 3D velocity field from some 70k red clump stars in the RAVE survey and found vertical motions that varied with Galactocentric radius $R$ as well as evidence for radial flows. Similarly, Carlin et al. (2013) and Pearl et al. (2017) used LAMOST spectroscopic velocities with proper motions from the PPMXL catalog (Roester et al. 2006).
functions in the GP prior. In GP regression, or kernel density estimation, where one produces a smooth model of the likelihood function, defined below, over the space of hyperparameters that control the strength and scale length of correlations in the velocity field. Thus, although the velocity field and the inputs are different positions within the Galaxy. The covariance matrix is constructed from a kernel function, which depends of a set of hyperparameters that control the strength and scale length of correlations in the velocity field. Moreover, the Gaia maps are somewhere between a projection, where one integrates out one spatial dimension, and a two-dimensional slice, where one takes a narrow range in one of the dimensions to get the velocity field on a two-dimensional surface. By contrast, our GP model can be queried to give the inferred velocity field with uncertainties at any point in the sample region. Furthermore, since the model is differentiable, we can use it to infer quantities such as the Oort constants at the position of the Sun. In addition, derivatives of V can be used to make statements about the dynamical state of the Galaxy. This feature allows us to map out variations in Oort functions in the vicinity of the Sun. It also allows us to construct a 3D map of the divergence, which is proportional to the total time derivative of the stellar number density. In so doing, we are able to identify regions near the Sun where the number density of stars may be undergoing compression and rarefaction.

An outline of our paper is as follows. In Section 2 we describe the sample used in our analysis as well as our binning strategy. In Section 3 we provide a brief introduction to Gaussian processes and outline our sparse GP algorithm. We also summarize results of tests performed with mock data. The results of our analysis for Gaia DR2 are discussed in Sections 4 and 5. We present estimates for the generalized Oort constants and maps of the velocity field as well as a three-dimensional map of the divergence of the velocity field. We discuss possible extensions of this work in Section 6 and provide a brief summary of our results and some concluding remarks in Section 7.

2 PRELIMINARIES

2.1 6D phase space catalog

The Gaia DR2 Radial Velocity survey contains 6D phase-space measurements for over 7 million stars (Gaia Collaboration et al. 2018; Katz, D. et al. 2019). In this work, we use the gaiaRvdelpepsdelsp43 catalog, which was constructed to correct for systematic errors in the Gaia parallaxes and uncertainties (Schönrich et al. 2019). Following Katz, D. et al. (2019) we use both Cartesian coordinates (X, Y, Z, U, V, W) and Galactic cylindrical coordinates (R, φ, Z, VR, Vφ, VZ). The Cartesian coordinate system has the Galactic centre at the origin and the Sun on the −X-axis; U, V and W are the usual velocity components where positive values at the position of the Sun indicate motion toward the Galactic centre, the direction of Galactic rotation, and the North Galactic Pole, respectively. We use 8.27 kpc for the Sun's distance.
to the Galactic centre and 20 pc for its distance to the Galactic midplane (see Schönrich et al. (2019) and references therein) so that \( (X_{\odot}, Y_{\odot}, Z_{\odot}) = (-8.27, 0, 0.02) \) kpc. We further set \( (U_{\odot}, V_{\odot}, W_{\odot}) = (11.1, 250, 7.24) \) km s\(^{-1}\) (Schönrich 2012). Our cylindrical coordinate system also has the Galactic centre at the origin, the Sun at \( \phi = 0^\circ \), and \( \phi \) increasing in the direction of Galactic rotation. This system is left-handed in the sense that increasing \( \phi \) corresponds to the direction toward the South rather than North Galactic pole. We note that at the position of the Sun \( V_R = -U \) and \( V_\phi = V \); \( V_Z = W \).

We calculate the \( U, V, W \) velocity components and their errors from astrometric observations using the method described in Johnson & Soderblom (1987). We then convert to \( V_R, V_\phi \) and \( V_Z \). The velocity components, but not their errors, were provided with the \textit{gaiaRV}delpepsdelsp43 catalog. Note that in that catalog, \( V_R, V_\phi, W \) are roughly Maxwellian. Thus, the stellar velocities are drawn from a Gaussian process than the stellar velocities themselves. Stars in the Galactic disk or the stellar halo. The velocity distributions for these components are not Gaussian but the Galactic rotation. This system is left-handed in the sense that increasing \( \phi \) corresponds to the direction toward the South rather than North Galactic pole. We note that at the position of the Sun \( V_R = -U \) and \( V_\phi = V \); \( V_Z = W \).

We apply the following restrictions to the sample:

- color: \( G_{bp} - G_{rp} < 1.5 \)
- magnitude: \( G > 14.5 \)
- parallax signal to noise: \( p/\sigma_p > 4 \)
- parallax uncertainty: \( \sigma_p < 0.1 \) mas with \( \sigma_p \) given by the Gaia pipeline
- excess \( B - R \) flux: \( 1.172 < E_{bp,rp} < 1.3 \)

These quality cuts are recommended in Schönrich et al. (2019) to ensure minimal systematic biases in derived kinematic quantities. They leave 4584106 stars from the original catalog of 6606247. Finally, we apply the following kinematic cuts to remove high-velocity outliers:

- proper motion: \( \mu_\alpha, \mu_\delta < 400 \text{ mas yr}^{-1} \)
- proper motion error: \( \sigma_{\mu_\alpha}, \sigma_{\mu_\delta} < 20 \text{ mas yr}^{-1} \)
- Galactocentric speed: \( |V| < 600 \text{ km s}^{-1} \)

These cuts are similar to those implemented by Williams et al. (2013) with the modification that our velocity cut is in terms of the speed in the local frame of rest whereas they apply a cut on the radial velocity. These cuts eliminate only about 400 stars.

### 2.2 Binning

As discussed above, computation time and RAM requirements make it unfeasible to apply GP regression directly to data sets with much more than \( 10^4 \) entries. For this reason, we bin stars so that input data for our GP analysis are the mean velocity components in cells. In a sense, velocity components averaged over stars in a cell are closer to Gaussian process than the stellar velocities themselves. Stars in the region near the Sun can come from the thin disk, the thick disk, or the stellar halo. The velocity distributions for these components are roughly Maxwellian. Thus, the stellar velocities are drawn from what might be better described as a mixture of Gaussians. On the other hand, the average velocity of some large number of stars in a cell will be approximately Gaussian due to the central limit theorem.

We set up a Cartesian grid of cells with size \( \Delta X, \Delta Y, \Delta Z = (125, 125, 50) \) pc for the region \( 4 < X (\text{kpc}) < 12, -4 < Y (\text{kpc}) < 4, \) and \( -2 < Z (\text{kpc}) < 2 \) and keep only those cells with more than 20 stars. Mean values and uncertainties for the \( V_R, V_\phi, V_Z \) components are then calculated by a standard least squares algorithm. In the end, we are left with mean velocities for 27305 cells representing the observations of 3972825 stars.

### 3 GAUSSIAN PROCESS REGRESSION

#### 3.1 Overview of Gaussian processes

We begin with a brief review of GP regression. More thorough discussions can be found in numerous resources such as the excellent book by (Rasmussen & Williams 2006). Suppose we have observations of a real scalar process \( f \) such that

\[
y_i = f(x_i) + \epsilon_i \quad i = \{ 1, 2, \ldots, N \}
\]

where \( x_i \) is the input vector for the \( i \)th observation, \( y_i \) is the scalar output and \( \epsilon_i \) is additive noise for that observation. For the case at hand, \( x_i = (X_i, Y_i, Z_i) \) is the position vector of the \( i \)th cell and \( f \) is \( V_R, V_\phi, \) or \( V_Z \). If \( f \) is a Gaussian Process then it is completely specified by a mean function \( m(x) \) and covariance function \( k(x, x') \) such that

\[
m(x) = \mathbb{E} \left[ f(x) \right]
\]

\[
k(x, x') = \mathbb{E} \left[ (f(x) - m(x))(f(x') - m(x')) \right]
\]

where \( \mathbb{E} \) denotes expectation value. Furthermore, the probability distribution function of \( f \) is given by

\[
P(f | X) = N(f | \mu, K)
\]

where \( X, \mu, \) and \( K \) are aggregate vectors of the input vectors \( x_i \), the mean functions \( m_i = m(x_i) \), and the kernel functions \( K_{ij} = k(x_i, x_j) \), respectively. As usual, \( N \) denotes a multivariate Gaussian. Equation 8 constitutes a Gaussian process prior \( f \). For simplicity, we assume \( m = 0 \).

The goal of GP regression is to infer \( f_i \), at new inputs \( X \), given the data \( \{X, y\} \). If the noise \( \epsilon \) is identical, independent, and Gaussian with dispersion \( \sigma_n \), then the joint probability for \( y \) and \( f_i \) is itself Gaussian

\[
P(y, f_i | X, x_i) = N \left( 0, \begin{bmatrix} K + \sigma_n^2 I & \mathbf{K}_x \\ \mathbf{K}_x^T & \mathbf{K}_{xx} \end{bmatrix} \right)
\]

(9)

where \( \mathbf{K} = K(X, X), \mathbf{K}_x = K(X, x_i) \) and \( \mathbf{K}_{xx} = K(x_i, x_i) \). The quantity of interest is the conditional probability

\[
P(f_i | y, X, x_i) = \frac{P(y, f_i | X, x_i)}{P(y)} = N \left( \bar{f}_i, C \right)
\]

(10)

where

\[
\bar{f}_i = \mathbf{K}_x^T \left( \mathbf{K} + \sigma^2 I \right)^{-1} y
\]

(12)

\[
C = \mathbf{K}_{xx} - \mathbf{K}_x^T \left( \mathbf{K} + \sigma^2 I \right)^{-1} \mathbf{K}_x
\]

(13)

(see Rasmussen & Williams (2006), section 2.2 and appendix A).

#### 3.2 Kernel function

A key ingredient of GP regression is the kernel or covariance function. Though there is considerable flexibility in choosing a kernel, for most problems it is constructed by taking sums and/or products of standard kernels, which themselves are constructed from elementary functions of the input variables.

In the present analysis, we allow for different kernels for \( V_R, V_\phi, \) and \( V_Z \) since they are modeled as independent scalar functions. Each of the kernels uses a three-dimensional radial basis function (RBF)
plus a term proportional to the identity matrix that accounts for unknown noise
\[ k(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 e^{-\frac{\ell^2}{2}} + \sigma_n^2 \delta_{ij} \]  (14)
where \( \sigma_f^2 \) is the signal variance and
\[ \ell_{ij} \equiv \frac{(X_i - X_j)^2}{l_x^2} + \frac{(Y_i - Y_j)^2}{l_y^2} + \frac{(Z_i - Z_j)^2}{l_z^2} \]  (15)
is a dimensionless pseudo-distance between \( \mathbf{x}_i \) and \( \mathbf{x}_j \). The RBF kernel is positive definite, differentiable, and maximal for \( \mathbf{x}_i = \mathbf{x}_j \). These are all features that lead to realistic models for the bulk velocity field of the Galactic disk. Through trial and error, we find that the fit for \( \nu_\phi \) is significantly improved by including in \( k \) the product of two linear kernels. The linear kernel is given by
\[ k_{\text{lin}}(\mathbf{x}, \mathbf{x}') = \sum_{i=x,y,z} \mu_i x_i x'_i. \]  (16)
It expands the space of priors on \( f \) to include linear functions of the inputs \( \mathbf{x} \). Unlike the RBF kernel, it is non-stationary in the sense that it depends on the absolute positions of the data points rather than the distance between pairs of data points. For \( \nu_\phi \), we add
\[ k_{\text{lin2}} = \left( \sum_{i=x,y,z} \mu_i x_i x'_i \right) \left( \sum_{j=x,y,z} y_j x'_j x'_i \right) \]  (17)
to the RBF and noise kernels in Equation 14.

The kernels defined above depend on a number of free parameters commonly referred to as hyperparameters, since they parameterize the covariance function rather than \( f \) itself. The hyperparameters determine characteristics of functions in the prior of \( f \). For example, \( l_x \), \( l_y \), and \( l_z \) control the correlation lengths along the three coordinate axes. Proper choice of hyperparameters is crucial for obtaining a good model of the data. If the length scales are too small, then the model will tend to overfit the data, thereby attributing small-scale bumps and wiggles to \( f \) rather than noise. Conversely, if the length scales are too large, then the model will tend to miss important features represented in the data.

In very simple problems, one can find suitable hyperparameters by trial and error as illustrated in Chapter 2 of Rasmussen & Williams (2006). A principled approach, and one that is suitable for more complex problems, is to maximize the marginal likelihood function \( p(y|\mathbf{X}, \theta) \) where
\[ \log(p(y|\mathbf{X}, \theta)) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log|K + \sigma_n^2 I| - \frac{1}{2} \mathbf{y}^T (K + \sigma_n^{-2} I)^{-1} \mathbf{y}. \]  (18)
Here, \( \theta \) represents the hyperparameters.

### 3.3 Sparse GP regression via inducing points

In any optimization scheme, one must evaluate the marginal likelihood function in Equation 18 a large number of times for different choices of the hyperparameters \( \theta \). Each likelihood call involves inversion of the \( N \times N \) matrix \( K \), which is an \( O(N^3) \) operation that requires \( O(N^2) \) RAM. This process becomes unfeasible for \( N \) much greater than \( 10^4 \). Fortunately, there are a number of algorithms that allow one to estimate the marginal likelihood using \( M < N \) inputs. These algorithms, generally referred to as sparse GP regression, reduce the computational complexity to \( O(NM^2) \) and the RAM requirement to \( O(NM) \). In this work, we follow the method described in (Bauer et al. 2016). We denote the full covariance matrix (\( \mathbf{K} \) in Equation 8) as \( \mathbf{K}_{f} \), the covariance matrix for the inducing points as \( \mathbf{K}_{u} \), and the cross covariance matrix as \( \mathbf{K}_{f u} \). In sparse GP regression, one replaces \( \mathbf{K}_{f} \) in Equation 18 with \( \mathbf{K}_{f} \mathbf{K}_{u}^{-1} \) and includes an additional term given by \( \text{Tr}(\mathbf{K}_{u}^{-1} - \mathbf{K}_{u})/(2\sigma_n^2) \) (Titsias 2009; Bauer et al. 2016). Note that optimization requires that we compute the gradient of the marginal likelihood with respect to the hyperparameters.

We implement our GP algorithm using GPy, a Gaussian process package written in Python by the University of Sheffield machine learning group (GPy 2012). We apply the sparse GP module with a heteroscedastic Gaussian likelihood, which allows us to incorporate uncertainties in the mean velocity components for individual stars. Optimization is carried out by applying the VarDTC inference method (Titsias 2009) using the GPy optimization module and the LBFGS-B algorithm from the software package SciPy (Jorge Nocedal 2006; Virtanen et al. 2020).

### 3.4 Mock data tests

Before turning to Gaia DR2, we test our algorithm on a mock data sample. This sample is constructed by replacing measured velocities in the gaiaRVdelpedsels43 catalog with velocities drawn from an analytic function that is chosen so that the resulting maps are qualitatively similar to ones found with the real data. For brevity, we present results for a single generic component with velocity field. The analytic function is given by
\[ V(X, Y, Z) = 2 \left( \frac{11 \cos g_1 - 3 \cos g_2 - 2 \cos g_3 - Z^2}{10} \right) \]  (19)
where \( g_1 = Z + (X - 8)/6, g_2 = 0.52(Y - 2), \) and \( g_3 = X - 1.3Y - 11.3 \). The velocity field for \( Z = 0 \) is shown in the upper left panel of Figure 1.

The mock data is analysed using the same procedure that will be applied to the real data. We first determine mean values for \( V \) and errors about the mean for the 27k cells. We then optimize the likelihood function over the hyperparameters using the inducing point algorithm described above with \( M = 3000 \) and the RBF plus noise kernel function. Armed with optimal hyperparameters, we determine
the posterior of the model velocity field \( \tilde{V} \) using Equation 12. The result for the \( XY \)-plane is shown in the lower left panel of Figure 1. We also determine the covariance matrix of the posterior using Equation 13. The square root of its diagonal elements provides an estimate of the dispersion, \( \sigma_V \), which is shown in the upper right panel of Figure 1. As expected, the dispersion rises sharply toward the edges of the region where we have data. Finally, we show the normalized residual \( \delta = (V - \tilde{V})/\sigma_V \) in the \( XY \) plane. In Figure 2, we plot a histogram of \( \delta \) over the sample volume. If the agreement between model and mock data were perfect, we would expect the histogram to be well-approximated by the normal distribution \( N(0,1) \). The red dashed line is the normal distribution fit \( N(0.11,0.79) \).

### Table 1. Optimal hyperparameters from our sparse GP analysis of \( Gaia \) data.

| \( \mu \) or \( \nu \) | \( VR \) | \( V\phi \) | \( VZ \) |
|---------------------|-------|-------|-------|
| Kernel              | RBF   | RBF   | RBF   |
| \( l_X \)           | 1.06  | 1.13  | 1.89  |
| \( l_Y \)           | 2.72  | 1.75  | 4.18  |
| \( l_Z \)           | 0.536 | 0.242 | 0.484 |
| \( \sigma_f \)      | 12.4  | 10.4  | 3.31  |
| \( \mu_{x}, \nu_{x} \) | 0.236 | 0.228 |
| \( \mu_{y}, \nu_{y} \) | 0.207 | 0.209 |
| \( \mu_{z}, \nu_{z} \) | 3.86  | 3.71  |

that the radial length scale of the disk is a factor of \( 5-10 \) times larger than the scale height and that variations in the disk tend to be stronger in the radial direction than the azimuthal direction.

### 4.1 Velocity field in components

Figures 7 and 6 show model predictions for the velocity field in the \( z = 0 \) and \( \phi = 0^\circ \) planes. These figures are produced by querying the model at the appropriate points using Equation 12. By and large, the maps are in good agreement with those found in Katz, D. et al. (2019) (See their Figures 10 and 11) though they are clearly much smoother. The GP analysis tends to pick out features on the length scales characterized by \( l_X, l_Y, \) and \( l_Z \). We also show the statistical uncertainties in the mean for the three components, \( \epsilon_i \), as calculated from Equation 13. In general the uncertainties are less than \( 1 \) km s\(^{-1} \) within the sample volume, but rise rapidly as one approaches the edge of the volume. Our maps can also be compared to those in Khanna et al. (2022) who combine data from \( Gaia \) EDR3 with radial velocity measurements from a number of spectroscopic surveys. They construct parametric models of the velocity field in heliocentric coordinates so a direct comparison is difficult.

The upper left panels of Figures 3 and 4 point to several regions of radial bulk flows. In particular, there is inward bulk motion just beyond the Solar circle that is approximately independent of \( \phi \) across the sample volume. This motion is primarily in the region \( |z| < 500 \) pc and is thus likely associated with the thin disk. On the other hand, the outward radial flow just inside the Solar circle is mainly found at positive \( \phi \) but across the full range in \( z \) of the survey. There also appears to be another outward flow at \( R > 10 \) kpc, \( \phi \approx -10^\circ \) and positive \( z \).

The middle-left panels of Figures 3 and 4 show \( V_\phi \). The dominant flow here, and indeed for \( V \) in general, is the motion about the Galactic centre. We see that contours of constant \( V_\phi \) roughly follow contours

![Figure 2. Histogram of the normalized residual \( \delta \) as defined in the text. The histogram includes residuals from all cells used in fitting the mock data and is itself normalized so that the area under the curve is unity. The black dashed line is the normal distribution \( N(0,1) \). The red dashed line is the normal distribution fit \( N(0.11,0.79) \).](image)

![Figure 3. Left: Components of the velocity field \( V_R, V_\phi, \) and \( V_Z \) in the \( z = 0 \) plane. Right: One-sigma confidence regions of the GP models.](image)
of constant $R$ in Figure 7. As discussed below, the average of $V_\phi$ over $\phi$ at $z = 0$ yields the rotation curve near the Solar circle. The main feature in Figure 4 is a decrease of $V_\phi$ as one moves away from the midplane of the Galaxy. This trend is easily explained by an increase in asymmetric drift due to a larger contribution from dynamically warmer stars in the thick disk. As with $V_R$, we see that $V_\phi$ is not perfectly symmetric about the midplane. Consider, for example, the ridge in $V_\phi$ between $X = -6$ and $X = -8$ in Figure 7, which coincides with the peak seen in Figure 6 at $R = 7$ kpc. The peak and its outer slopes are shifted by a small, but non-negligible, amount above the midplane.

The bottom left panels of Figures 3 and 4 show our results for the vertical bulk motion, $V_Z$. In the midplane we see a clear trend of increasing $V_Z$ as one moves across the local patch of the Galaxy in the direction of increasing $R$ and $\phi$. The view in the $RZ$ plane shows a mix of bending and breathing motions. The disk is moving downward inside the Solar circle and upward outside the Solar circle and in addition it appears to be expanding away from the midplane.

The generation of bending and breathing waves through internal disk dynamics and interactions with satellites such as the Sagittarius dwarf have been studied in N-body simulations by numerous authors including (Gómez et al. 2013; Chequers et al. 2018; Laporte et al. 2019; Poggio et al. 2021; Bennett et al. 2022) and Thulasidharan et al. (2022). Observational studies using data from various surveys in Widrow et al. (2012); Williams et al. (2013); Carlin et al. (2013); Carrillo et al. (2018); Katz, D. et al. (2019); Wang et al. (2019); López-Corredoira et al. (2020).

### 4.2 Velocity vectors

To aid our understanding of the velocity field we present velocity vector maps in three projections. Similar vector field maps were presented in Pearl et al. (2017) using data from LAMOST and, most recently, by Fedorov et al. (2022) using data from Gaia DR2. We begin with Figure 5, which presents the vector field in the $\phi Z$ plane at $R = R_0$, that is, on a curved cylindrical surface that includes a part of the Solar circle. To help visualize the velocity field, we have subtracted off the vector $220 \text{ km s}^{-1} \hat{\phi}$. Radial velocities are shown as a color map. Thus, all three components of the velocity field are represented in the figure.

As already noted in the $V_\phi$ panel of Figure 6, the dominant feature of the map is the decrease in $V_\phi$ due to asymmetric drift as one moves away from the midplane. Velocities in the midplane are about $10 \text{ km s}^{-1}$ higher than they are at $|Z| = 500 \text{ pc}$. Though the dominant flow is in the azimuthal direction, we do see a clear downward motion for $\phi < 0^\circ$, in agreement with Figure 7.

In Figure 6 we show the velocity field in the $RZ$-plane at $\phi = 0^\circ$, which passes through the position of the Sun. We see that there are three distinct regions defined primarily by the sign of $V_R$. Specifically, we find an inward flow centered on $(R, z) = (8.5, -0.2) \text{kpc}$, an outward and downward flow inside the Solar circle, and an outward
and upward flow for $R > 10$ kpc and $z \sim 500$ pc. The results are in good agreement with Figure 10 of Pearl et al. (2017) for regions where the samples overlap.

### 4.2.1 XY vectorfield

In Figure 7 we show the velocity field in the midplane with vertical bulk motions are shown as the background color map. The map paints an image of several flows in the vicinity of the Sun, which can be described by a combination of expansions or compression and shear. For example, roughly along the Solar circle, we find radial compression and azimuthal shear — the dominant motion inside the Solar circle is outward and toward positive $\phi$ while the motion just beyond the Solar circle is inward and toward negative $\phi$.

It is tempting to attribute flows in the midplane to the presence of spiral arms. With this in mind, we overlay the spiral arm model of Reid et al. (2019). In general, one expects to see motion toward the Galactic centre on the leading side of a spiral arm and motion away from the Galactic centre on the trailing side (Kawata et al. 2018; Grand et al. 2015; Hunt et al. 2015). This pattern is indeed seen along the Local Arm in Figure 7 though the connection between flows in our map and the other arms is more tenuous. For example, roughly along the Solar circle, we find radial compression and azimuthal shear — the dominant motion inside the Solar circle is outward and toward positive $\phi$ while the motion just beyond the Solar circle is inward and toward negative $\phi$.

### 5 VELOCITY GRADIENT

In the previous section, we showed that our GP model could be queried to give the bulk velocity field $V$ at any position in the sample region. Since the GP solution is continuous and differentiable, we can also use it to infer the gradient of the velocity field, a $3 \times 3$-tensor whose components are $\partial V_i / \partial X_j$. These components can be computed either by replacing the kernel functions in Equations 12 and 13 with their gradients or by a finite difference scheme. We have confirmed that the two methods agree for the RBF kernel. In what follows we use finite differencing. Similar results are presented in Fedorov et al. (2021).

#### 5.1 Oort constants

Gradients of the velocity field played a central role in Oort’s seminal work on the rotation of the Galaxy Oort (1927). The basic idea was to expand the velocity field near the Sun as a Taylor series in position. Though Oort considered only axisymmetric flows in the $X-Y$ plane, the method was extended to general three dimensional flows by Ogrodnikoff (1932); Milne (1935); Chandrasekhar (1942) and Ogorodnikov (1965).

To linear order in position, the velocity field near the Sun can be written as a Taylor series

$$ V = V_0 + H \cdot x + O(x^2) . $$

If we restrict ourselves to the projection of the velocity field onto the midplane of the Galaxy, then $H$ reduces to the $2 \times 2$ matrix

$$ H = \begin{pmatrix} \partial v_x / \partial x & \partial v_x / \partial \phi \\ \partial v_y / \partial x & \partial v_y / \partial \phi \end{pmatrix} \equiv \begin{pmatrix} K + C & A - B \\ A + B & K - C \end{pmatrix} , $$

where the second equality defines the Oort constants $A, B, C,$ and $K$. These constants measure, respectively, the azimuthal shear, vorticity, radial shear, and divergence of the velocity field in the midplane of the disk. In Galactocentric polar coordinates $(R, \phi)$ the Oort constants are given by

$$ 2A = \frac{v_\phi}{R} - \frac{\partial v_\phi}{\partial R} - \frac{1}{R} \frac{\partial v_R}{\partial \phi} , $$

$$ 2B = -\frac{v_\phi}{R} - \frac{\partial v_\phi}{\partial R} + \frac{1}{R} \frac{\partial v_R}{\partial \phi} , $$

$$ 2C = -\frac{v_R}{R} + \frac{\partial v_R}{\partial R} - \frac{1}{R} \frac{\partial v_\phi}{\partial \phi} , $$

$$ 2K = \frac{v_R}{R} + \frac{\partial v_R}{\partial R} + \frac{1}{R} \frac{\partial v_\phi}{\partial \phi} , $$

Chandrasekhar (1942).

The usual method for determining the Oort constants from astrometric data is based on the observation that the proper motion of a star in the direction of Galactic longitude, $\mu_l$, and the line-of-sight velocity, $v_{los}$, can be expressed in terms of sine and cosine functions of $l$. The Oort constants appear as coefficients in this truncated Fourier series and can therefore be determined by standard statistical methods. (See, for example, Feast & Whitelock (1997); Olling & Dehnen (2003); Bovy (2017); Vityazev et al. (2018); Wang et al. (2021).

In this paper, we compute the Oort constants directly from our GP model for the velocity field using the above equations. Our results along with a selection of values from the literature are given in Table 2. We find excellent agreement with recent measurements from Bovy (2017); Vityazev et al. (2018); Li et al. (2019) and Wang et al. (2021). As noted above, the values from the literature all determine the Oort constants by fitting $\mu_l$ and $v_{los}$ to a low-order Fourier series in $l$. They do, however, use data from different surveys and with different geometric selection functions and sample sizes. For example, both Bovy (2017) and Li et al. (2019) consider large samples of stars within 500 pc of the Sun. Bovy (2017) considers main sequence stars from the Tycho–Gaia Astrometric (TGAS) catalog (Michalik, Daniel et al. 2015) while Li et al. (2019) considers all stars from Gaia DR2.
providetheinferredvaluesattheexactpositionoftheSunandaveragevalues
overasphericalvolumeofradius500pc.

| Origin          | A     | B     | C     | K
|------------------|-------|-------|-------|-------|
| Oort (1927)      | ~19   | ~24   | 0     | 0     |
| Feast (1997)     | 14.8 ± 0.8 | -12.4 ± 0.6 | -9.8 ± 2 | 0     |
| Olling (2003)    | 15.9 ± 2 | -16.9 ± 2 | -9.8 ± 2 | 0     |
| Boyy (2017)      | 15.3 ± 0.4 | -11.9 ± 0.4 | -3.2 ± 0.4 | -3.3 ± 0.6 |
| Vityazev (2018)  | 16.3 ± 0.1 | -11.9 ± 0.1 | -3.0 ± 0.1 | -4.0 ± 0.2 |
| Li (2019)        | 15.1 ± 0.1 | -13.4 ± 0.1 | -2.7 ± 0.1 | -1.7 ± 0.2 |
| Wang (2021)      | 16.3 ± 0.9 | -12.0 ± 0.8 | -3.1 ± 0.5 | -1.3 ± 1.0 |
| This work (p)    | 16.2 ± 0.2 | -11.7 ± 0.2 | -3.1 ± 0.2 | -3.0 ± 0.2 |
| This work (v)    | 15.2 ± 0.8 | -12.4 ± 0.9 | -2.9 ± 0.8 | -2.3 ± 0.5 |

Wang et al. (2021) uses a relatively small sample of A-stars from LAMOST with a similar range in distance from the Sun. Finally, Vityazev et al. (2018) uses stars from TGAS but with the larger reach of 1.5 kpc.

The model allows us to predict the Oort constants at a single point, namely the position of the Sun. These values are given in the second to last line in Table 2. To allow for a closer comparison to literature values, we also include values for the Oort constants averaged over a spherical volume of radius 500 pc.

5.2 Oort functions

In Oort’s 1927 original work, stars are assumed to follow circular orbits in the midplane of the Galaxy. Under this assumption, C and K are identically zero and A and B can be written in terms of circular speed and its gradient at the position of the Sun. They therefore provide a direct probe of the gravitational potential and hence matter distribution in the Galaxy. Olling and Merrifield (1998) extended the idea of Oort constants to Oort functions with

\[
A(R) = \frac{1}{2} \left( \frac{\nu_c(R)}{R} - \frac{\partial \nu_c}{\partial R} \right)
\]

(26)

\[
B(R) = \frac{1}{2} \left( \frac{\nu_c(R)}{R} + \frac{\partial \nu_c}{\partial R} \right)
\]

(27)

\[
M^\pm = \frac{1}{2} \left( \frac{\partial \nu_c}{\partial x_i} \pm \frac{\partial \nu_c}{\partial x_k} \right)
\]

(28)

where as before, the components are evaluated at the position of the Sun (Ogorodnikoff 1932; Milne 1935; Ogorodnikov 1965; Tsvertex & Amosov 2019; Fedorov et al. 2021). Note that the coefficients of \(M^\pm\) include the Oort constants. For example, \(A = M^+_{12}\) and \(B = M^-_{21}\).
where the trace of the dispersion tensor would allow one to study the first moments of the CBE, namely the Jeans equations. In contrast with the continuity equation, these equations involve gradients of the gravitational potential. Thus, we will typically have two unknown quantities, one involving time-derivatives of the moments and the other involving the potential. Nevertheless, one should be able to infer something about the dynamics of the stellar disk, and in particular, departures from equilibrium by modelling moments of the DF via GP regression.

These considerations suggest an extension of the present work where we treat the output as a single three-components vector rather than three independent scalars. Doing so would allow for correlations between the different components of the velocity field. The framework for extending GP regression to vector outputs is already well-established (Alvarez et al. 2011).

7 CONCLUSIONS

In this paper, we present a GP model for the mean or bulk stellar velocity field in the vicinity of the Sun using astrometric measurements from Gaia DR2. The model is nonparametric in the sense that there are no prior constraints on the functional form of the velocity field. Instead, one specifies the functional form of the kernel function, which through a set of hyperparameters, controls properties of the prior on the velocity field such as its coherence length.

The main challenge in applying GP regression to Gaia data comes from the large-N requirements in both computing time and RAM. Fortunately, there is a large effort within the machine learning community on addressing these problems (See, for example, Hensman et al. (2013)). In this work, we have just started to exploit methods developed in that field.

Our model provides smooth, differentiable versions of the velocity field maps found in Katz, D. et al. (2019) and elsewhere. The Gaia maps used different binning schemes to compute different projections of the velocity field. In our case, we compute a single GP model (or more precisely, three independent models for each of the cylindrical velocity field components) from which properties of the velocity field could be derived. As a check of the model, we confirm that the values for the Oort constants derived by differentiating the velocity field agree extremely well with recent determinations in the literature.

A GP model for the velocity field can thus provide a starting point for making contact between astrometric data and dynamics via the continuity and Jeans equations. Already, our results for the divergence of the velocity field have provided a unique perspective on departures in the disk from an equilibrium state.
ACKNOWLEDGEMENTS

It is a pleasure to thank Dan Foreman-Mackey, Hamish Silverwood, and Lauren Anderson for helpful conversations on Gaussian processes. We acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada. LMW is grateful to the Kavli Institute for Theoretical Physics at the University of California, Santa Barbara for providing a stimulating environment during a 2019 program on galactic dynamics. His research at the KITP was supported by the National Science Foundation under Grant No. NSF PHY-1748958.

DATA AVAILABILITY

Data used in this paper is available through Zenodo (https://doi.org/10.5281/zenodo.2557803)

REFERENCES

Alvarez M. A., Rosasco L., Lawrence N. D., 2011, arXiv e-prints, p. arXiv:1106.6251
Bauer M., van der Wilk M., Rasmussen C. E., 2016, arXiv e-prints, p. arXiv:1606.04820
Benjamin R. A., et al., 2005, ApJ, 630, L149
Bennett M., Bovy J., Hunt J. A. S., 2022, ApJ, 927, 131
Binney J., 2013, New Astron. Rev., 57, 29
Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press
Bobylev V. V., Bajkova A. T., 2021, Astronomy Letters, 47, 607
Bovy J., 2017, MNRAS:Letters, 468, L63
Carlin J. L., et al., 2013, ApJ, 777, L5
Carrillo I., et al., 2018, MNRAS, 475, 2679
Chandrasekhar S., 1942, Principles of stellar dynamics. Dover Pheonix Editions
Chequers M. H., Widrow L. M., Darling K., 2018, MNRAS, 480, 4244
Churchwell E., et al., 2009, PASP, 121, 213
Cui X.-Q., et al., 2012, Research in Astronomy and Astrophysics, 12, 1197
Churchwell E., et al., 2009, PASP, 121, 213
Cui X.-Q., et al., 2012, Research in Astronomy and Astrophysics, 12, 1197
Feautrier P., Whitelock P., 1997, MNRAS, 291, 683
Feautrier P., Whitelock P., 1997, MNRAS, 291, 683
Fedorov P. N., Akhmetov V. S., Velichko A. B., Dmytrenko A. M., Denischenko S. I., 2021, MNRAS, 508, 3055–3067
Fedorov P. N., Akhmetov V. S., Velichko A. B., Dmytrenko A. M., Denischenko S. I., 2022, arXiv e-prints, p. arXiv:2205.09381
Friske J. K. S., Schönrich R., 2019, MNRAS, 490, 5414
GPY since 2012, GPY: A Gaussian process framework in python, http://github.com/SheffieldML/GPy
Gaia Collaboration et al., 2018, A&A, 616, A1
Gómez F. A., Minchev I., O’Shea B. W., Beers T. C., Bullock J. S., Purcell C. W., 2013, MNRAS, 429, 159
Grand R. J. J., Bovy J., Kawata D., Hunt J. A. S., Famaey B., Siebert A., Monari G., Cropper M., 2015, MNRAS, 453, 1867
Hensman J., Fusi N., Lawrence N. D., 2013, Proceedings of the Twenty-Ninth Conference on Uncertainty in Artificial Intelligence, pp 282 – 290
Hunt J. A. S., Kawata D., Grand R. J. J., Minchev I., Pasetto S., Cropper M., 2015, MNRAS, 450, 2132
Johnson D. R. H., Soderblom D. R., 1987, AJ, 93, 864
Jorge Nocedal S. J. W., 2006, Numerical Optimization. Springer, https://doi.org/10.1007/978-0-387-40065-5
Katz, D. et al., 2019, A&A, 622, A205
Kawata D., Hunt J. A. S., Grand R. J. J., Pasetto S., Cropper M., 2014, MNRAS, 443, 2757
Li C., Zhao G., Yang C., 2019, ApJ, 872, 205
López-Corredoira M., Garzón F., Wang H. F., Sylos Labini F., Nagy R., Chrobakóvá Z., Chang J., Villarroel B., 2020, A&A, 634, A66
Mihalik, Daniel Lindegren, Lennart Hobbs, David 2015, A&A, 574, A115
Milne E. A., 1935, MNRAS, 95, 560
Ogordnikov K. F., 1965, Dynamics of stellar systems. Pergamon Press
Olling R. P., Dehnen W., 2003, ApJ, 599, 275–296
Olling R. P., Merrifield M. R., 1998, MNRAS, 297, 943
Pearl A. N., Newberg H. J., Carlin J. L., Smith R. F., 2017, ApJ, 847, 123
Poggio E., Laporte C. F. P., Johnston K. V., D’Onghia E., Drimmel R., Grier Filho D., 2021, MNRAS, 508, 541
Rasmussen C. E., Williams C. K., 2006, Gaussian Processes for Machine Learning, MIT Press, http://www.gaussianprocess.org/gpml/chapters/RW.pdf
Steinmetz M., et al., 2006, AJ, 132, 1645
Thulasiadharan L., D’Onghia E., Poggio E., Gallagher J. S. I., Swiggum C., Benjamin R. A., Alves J., 2022, A&A, 660, L12
Titsias M., 2009, Journal of Machine Learning Research - Proceedings Track, 5, 567
Tsvetkov A. S., Amosov F. A., 2019, Astronomy Letters, 45, 462
Virtanen P., et al., 2020, Nature Methods, 17, 261
Vityazev V. V., Popov A. V., Tsvetkov A. S., Petrov S. D., Trofimov D. A., Kiyaei V. I., 2018, Astronomy Letters, 44, 236
Wang H.-F., et al., 2019, MNRAS, 491, 2104
Wang F., Zhang H.-W., Huang Y., Chen B.-Q., Wang H.-F., Wang C., 2021, MNRAS, 504, 199
Williams M. E. K., et al., 2013, MNRAS, 436, 101
Yanny B., et al., 2009, AJ, 137, 4377

This paper has been typeset from a LaTeX file prepared by the author.