Table 1. Inferred parameters over 100 independent realizations of respectively the Ornstein-Uhlenbeck, Ginzburg-Landau Double-Well and Lotka-Volterra dynamics. For every algorithm, we show the median $\pm$ one standard deviation.

Ground truth	NPSDE	VGPA	ESGF	AReS	MaRS
$\theta_0 = 0.5$	0.41 ± 0.11	0.53 ± 0.08	0.49 ± 0.07	$\mathbf{0.50 \pm 0.21}$	0.46 ± 0.06
$\theta_1 = 1$	0.71 ± 1.34	0.96 ± 0.31	0.96 ± 0.24	1.06 ± 0.93	$\mathbf{0.99 \pm 0.25}$
$H = 0.25$	0.00 ± 0.01	/	0.19 ± 0.06	$\mathbf{0.24 \pm 0.09}$	

(a) Ornstein-Uhlenbeck process

Ground truth	NPSDE	ESGF	AReS	MaRS
$\theta_0 = 2$	1.58 ± 0.71	2.04 ± 0.09	2.36 ± 0.18	$\mathbf{2.00 \pm 0.09}$
$\theta_1 = 1$	0.74 ± 0.31	1.02 ± 0.05	1.18 ± 0.9	$\mathbf{1.00 \pm 0.04}$
$\theta_2 = 4$	2.26 ± 1.51	3.87 ± 0.59	$\mathbf{3.97 \pm 0.63}$	3.70 ± 0.51
$\theta_3 = 1$	0.49 ± 0.35	0.96 ± 0.14	$\mathbf{0.98 \pm 0.18}$	0.91 ± 0.14
$\mathbf{H}_{1,1} = 0.05$	/	0.01 ± 0.03	$\mathbf{0.03 \pm 0.004}$	
$\mathbf{H}_{1,2} = 0.03$	/	0.01 ± 0.01	$\mathbf{0.02 \pm 0.01}$	
$\mathbf{H}_{2,1} = 0.03$	/	0.01 ± 0.01	$\mathbf{0.02 \pm 0.01}$	
$\mathbf{H}_{2,2} = 0.09$	/	0.03 ± 0.02	$\mathbf{0.09 \pm 0.03}$	

(b) Lotka-Volterra

Ground truth	NPSDE	VGPA	ESGF	AReS	MaRS
$\theta_0 = 0.1$	0.09 ± 7.00	0.05 ± 0.04	0.01 ± 0.03	0.09 ± 0.04	$\mathbf{0.10 \pm 0.05}$
$\theta_1 = 4$	3.36 ± 248.82	1.11 ± 0.66	0.11 ± 0.16	3.68 ± 1.34	$\mathbf{3.85 \pm 1.10}$
$H = 0.25$	0.00 ± 0.02	/	0.20 ± 0.05	$\mathbf{0.21 \pm 0.09}$	

(c) Double-Well potential

the same values. Due to space restrictions, the results for Lorenz '63 can be found in Table 2 of the appendix. As demonstrated by this experiment, AReS and MaRS can deal with locally linear systems, outperforming their competitors, especially in their estimates of the diffusion terms.

4.4. Non-Diagonal Diffusion

To investigate the effect of off-diagonal entries in $\mathbf{G}$, we use the Lotka-Volterra dynamics. Since NPSDE is unable to model non-diagonal diffusions, we provide it with the true $\mathbf{G}$ and only compare parameter estimates. As VGPA is already struggling in the lower dimensional cases, we omit it from this comparison due to limited computational resources. The results are shown in Table 1b. AReS and MaRS clearly outperform the other methods in terms of diffusion estimation, while ESGF is the only algorithm that yields drift parameter estimates of comparable quality.

4.5. Dealing with Multi-Modality

As a final challenge, we investigate the Ginzburg-Landau double well potential. Despite one-dimensional, its state distribution is multi-modal even if all parameters are known. As shown in Table 1c, this is definitely a challenge for all classical approaches. While the number of data-points is probably not enough for the non-parametric proxy for the drift function in NPSDE, the time-dependent Gaussianity assumptions in both VGPA and ESGF are problematic in this case. In our gradient matching framework, no such assumption is made. Thus, both AReS and MaRS are able to deal with the multimodality of the problem.

5. Conclusion

Parameter and diffusion estimations in stochastic systems arise in quantitative sciences and many fields of engineering. Current techniques based on Kalman filtering or Gaussian processes approximate the state distribution conditioned on the parameters and iteratively optimize the data likelihood. In this work, we propose to turn this procedure on its head by leveraging key ideas from gradient matching algorithms, originally designed for deterministic ODEs. By introducing a novel noise model for Gaussian process regression that leverages the Doss-Sussmann transformation, we are able to reliably estimate the parameters in the drift and the diffusion processes. Our algorithm can keep up with and occasionally outperform the state-of-the-art on the simpler benchmark systems, while it is also accurately estimating parameters for systems that exhibit multi-modal state densities, a case where traditional methods fail. While our approach is currently restricted to systems with a constant diffusion matrix $\mathbf{G}$, it would be interesting to see how it generalizes to other settings, perhaps using alternative or approximate bridge constructs. Unfortunately, this is outside of the scope of this work. We hope nevertheless that the publicly available code will facilitate future research in that direction.