Supplementary Material: Extensions and Monte Carlo simulation study

Holger Dette[a], Christophe Ley[b], and Francisco J. Rubio[c]

[a] Ruhr-Universität Bochum, Germany.

[b] Ghent University, Belgium.

[c] London School of Hygiene & Tropical Medicine, UK.

1 Extension to other distributions

We here briefly show how our method applies to other types of flexible distributions where a (skewness) parameter has a perturbation effect on the original distribution.

Log-skew-symmetric distributions

The proposed priors from our main paper have the same interpretation if they are used for the perturbation parameter in log-skew-symmetric distributions. Recall that a positive random variable Y is said to be distributed according to a log-skew-symmetric distribution if it is distributed according to (1) below. This sort of distributions have been used for modelling environmental, medical, biological, and financial data (see Marchenko and Genton, 2010 and the references therein). The pdf of Y is given by

$$s_l(y;\lambda) = \frac{2}{y} f(\log y) G(\lambda \log y), \quad y > 0. \tag{1}$$

It follows that the TV distance between (1) and the corresponding baseline log-symmetric density $f_l(y) = \frac{1}{y} f(\log y)$ satisfies

$$\mathrm{d_{TV}}(s_l, f_l|\lambda) = \frac{1}{2}\int_0^\infty |s_l(y;\lambda) - f_l(y)|\, dy = \frac{1}{2}\int_{-\infty}^\infty |s(x;\lambda) - f(x)| dx = \mathrm{d_{TV}}(s, f|\lambda).$$

Consequently, the priors proposed in Section 3 of the main paper for the skew-symmetric family coincide with those obtained for the log-skew-symmetric family. It is also clear that one could use any other increasing diffeomorphism from $\mathbb{R}_+$ to $\mathbb{R}$ instead of the logarithmic transformation.

Two-piece distributions

Consider the family of two-piece distributions with the following parameterisation (see Rubio and Steel, 2014 for a general overview):

$$s_{tp}(x;\gamma) = f\left(\frac{x}{1-\gamma}\right) I(x<0) + f\left(\frac{x}{1+\gamma}\right) I(x \geq 0), \;\; x \in \mathbb{R}, \tag{2}$$

where $\gamma \in (-1, 1)$, and f is a unimodal symmetric pdf with mode at 0. The parameter γ controls the mass cumulated on either side of the mode ($x = 0$) while preserving the tail behaviour of f. Density (2) is asymmetric for $\gamma \neq 0$ and it reduces to f for $\gamma = 0$. The TV distance between s_{tp} and the baseline pdf f is given by:

$$\mathrm{d_{TV}}(s_{tp}, f|\gamma) \;=\; \frac{1}{2}\int_{-\infty}^{0} \left| f\left(\frac{x}{1-\gamma}\right) - f(x) \right| dx + \frac{1}{2}\int_{0}^{\infty} \left| f\left(\frac{x}{1+\gamma}\right) - f(x) \right| dx = \frac{|\gamma|}{2}.$$

If we define the measure of perturbation $\mathrm{M_{TV}}(\gamma) = \gamma/2$, this coincides, up to a proportionality constant, with the AG measure of skewness proposed in Arnold and Groeneveld (1995) (see Rubio and Steel, 2014). Consequently, if we assume that $\mathrm{M_{TV}}(\gamma) = \gamma/2 \sim \mathrm{Beta}(\alpha, \beta)$ (using the notation in the main paper for the Beta distribution), we obtain the AG-Beta priors proposed in Rubio and Steel (2014) for this family of distributions.

2 Simulation study

This section is a complement to the Monte Carlo simulation study in Section 4.1 of the main paper. It is a performance comparison between noninformative priors built according to our new method and the Jeffreys prior.

We simulate $N = 1,000$ samples of sizes $n = 100$ and $n = 200$ from the skew-normal, skew-logistic and skew-Laplace distributions, in each case with location parameter $\mu = 0$, scale parameter $\sigma = 1$, and skewness parameter $\lambda = 0, 2.5, 5$. For each of these samples, we simulate a posterior sample of size $1,000$ from (μ, σ, λ) using the $\mathrm{BTV}(1,1)$, $\mathrm{BTV}(1/2, 1/2)$, and Jeffreys priors. We employ a self-adaptive MCMC sampler to obtain the posterior samples. For each posterior sample, we calculate the coverage proportions of the 95% credible intervals of each parameter (that is, the proportion of credible intervals that contain the true value of the parameter) as well as the 5%, 50% and 95% quantiles of the posterior medians and maximum *a posteriori* (MAP) estimators. In addition, we obtain the median of the Bayes factors (BFs) associated to the hypothesis $H_0 : \lambda = 0$. The Bayes factors are approximated using the Savage-Dickey density ratio. Results are reported in Tables 1S–6S. Overall, we observe that the Jeffreys and $\mathrm{BTV}(1/2, 1/2)$ priors exhibit the best, and very similar, performance.

References

B.C. Arnold and R.A. Groeneveld. Measuring skewness with respect to the mode. *The American Statistician*, 49:34–38, 1995.

Y.V. Marchenko and M.G. Genton. Multivariate log-skew-elliptical distributions with applications to precipitation data. *Environmetrics*, 21:318–340, 2010.

F.J. Rubio and M.F.J. Steel. Inference in two-piece location-scale models with Jeffreys priors. *Bayesian Analysis*, 9:1–22, 2014.

Prior	MAP			Median			Coverage	BF
	5%	50%	95%	5%	50%	95%		
$\lambda = 0$								
BTV(1/2,1/2)								
μ	-1.031	0.018	0.993	-0.797	0.009	0.775	0.983	–
σ	0.943	1.074	1.389	1.000	1.150	1.387	0.861	–
λ	-1.455	0.001	1.481	-1.244	-0.018	1.315	0.984	2.002
Jeffreys								
μ	-1.039	0.028	1.020	-0.864	0.011	0.800	0.982	–
σ	0.946	1.084	1.432	1.008	1.159	1.400	0.839	–
λ	-1.537	-0.021	1.482	-1.392	-0.026	1.401	0.982	2.204
BTV(1,1)								
μ	-0.997	-0.007	0.960	-0.708	0.015	0.683	0.992	–
σ	0.933	1.067	1.355	0.985	1.131	1.346	0.877	–
λ	-1.186	0.008	1.279	-0.992	-0.010	1.075	0.992	1.404
$\lambda = 2.5$								
BTV(1/2,1/2)								
μ	-0.188	0.014	0.646	-0.170	0.058	0.635	0.880	–
σ	0.650	0.955	1.177	0.704	0.947	1.175	0.898	–
λ	-0.029	2.111	4.318	0.212	2.139	5.127	0.879	0.386
Jeffreys								
μ	-0.202	0.009	0.478	-0.174	0.052	0.638	0.890	–
σ	0.651	0.959	1.180	0.708	0.952	1.177	0.919	–
λ	0.003	2.138	4.308	0.274	2.191	5.028	0.885	0.369
BTV(1,1)								
μ	-0.176	0.031	0.739	-0.148	0.088	0.667	0.870	–
σ	0.645	0.931	1.155	0.691	0.922	1.150	0.897	–
λ	-0.069	1.955	3.867	0.151	1.950	4.504	0.859	0.338
$\lambda = 5$								
BTV(1/2,1/2)								
μ	-0.142	-0.006	0.152	-0.120	0.004	0.185	0.922	–
σ	0.806	0.983	1.137	0.814	0.988	1.147	0.946	–
λ	1.724	4.245	10.966	2.289	4.995	14.257	0.931	0.004
Jeffreys								
μ	-0.134	-0.003	0.150	-0.123	0.004	0.184	0.925	–
σ	0.810	0.986	1.147	0.817	0.991	1.147	0.944	–
λ	1.893	4.336	10.033	2.409	5.044	15.041	0.927	0.004
BTV(1,1)								
μ	-0.117	0.011	0.177	-0.103	0.019	0.231	0.922	–
σ	0.788	0.975	1.133	0.780	0.975	1.131	0.935	–
λ	1.767	4.044	8.088	1.945	4.582	10.286	0.932	0.003

Table 1S: Skew-normal data for noninformative priors: $\mu = 0, \sigma = 1, n = 100$.

Prior	MAP			Median			Coverage	BF
	5%	50%	95%	5%	50%	95%		
$\lambda = 0$								
BTV(1/2,1/2)								
μ	-0.886	-0.035	0.869	-0.682	-0.009	0.692	0.994	–
σ	0.960	1.054	1.330	1.008	1.108	1.296	0.846	–
λ	-1.286	0.028	1.295	-1.028	-0.004	1.037	0.993	2.306
Jeffreys								
μ	-0.890	0.032	0.890	-0.718	0.006	0.720	0.991	–
σ	0.961	1.058	1.351	1.009	1.115	1.319	0.844	–
λ	-1.335	-0.002	1.297	-1.082	-0.019	1.052	0.991	2.557
BTV(1,1)								
μ	-0.857	0.002	0.860	-0.602	-0.007	0.652	0.995	–
σ	0.952	1.048	1.285	0.999	1.098	1.275	0.862	–
λ	-1.207	-0.001	1.180	-0.992	-0.004	0.851	0.995	1.570
$\lambda = 2.5$								
BTV(1/2,1/2)								
μ	-0.124	0.008	0.211	-0.119	0.021	0.317	0.905	–
σ	0.699	0.981	1.123	0.760	0.980	1.124	0.909	–
λ	1.158	2.283	3.671	0.924	2.362	3.814	0.907	0.012
Jeffreys								
μ	-0.127	0.006	0.202	-0.119	0.019	0.321	0.914	–
σ	0.709	0.984	1.120	0.772	0.982	1.125	0.911	–
λ	1.191	2.319	3.645	0.922	2.369	3.886	0.916	0.006
BTV(1,1)								
μ	-0.123	0.015	0.225	-0.110	0.031	0.401	0.898	–
σ	0.687	0.976	1.123	0.743	0.972	1.117	0.902	–
λ	0.787	2.231	3.541	0.722	2.294	3.666	0.897	0.014
$\lambda = 5$								
BTV(1/2,1/2)								
μ	-0.090	0.000	0.097	-0.082	0.003	0.098	0.938	–
σ	0.874	0.992	1.108	0.874	0.995	1.107	0.935	–
λ	2.948	4.603	7.795	3.129	4.939	8.662	0.923	5×10^{-11}
Jeffreys								
μ	-0.087	-0.001	0.095	-0.081	0.004	0.098	0.940	–
σ	0.873	0.989	1.099	0.874	0.995	1.108	0.937	–
λ	2.922	4.615	7.907	3.134	4.947	8.943	0.926	4×10^{-11}
BTV(1,1)								
μ	-0.080	0.004	0.098	-0.076	0.008	0.106	0.945	–
σ	0.867	0.987	1.096	0.869	0.990	1.101	0.944	–
λ	2.853	4.457	7.475	3.048	4.748	8.275	0.929	5×10^{-11}

Table 2S: Skew-normal data for noninformative priors: $\mu = 0, \sigma = 1, n = 200$.

Prior	MAP			Median			Coverage	BF
	5%	50%	95%	5%	50%	95%		
$\lambda = 0$								
BTV(1/2,1/2)								
μ	-1.139	0.048	1.191	-0.994	0.055	1.046	0.957	–
σ	0.906	1.060	1.244	0.936	1.099	1.282	0.911	–
λ	-0.682	-0.014	0.641	-0.842	-0.030	0.820	0.952	2.794
Jeffreys								
μ	-1.190	0.047	1.227	-1.044	0.059	1.097	0.952	–
σ	0.913	1.070	1.254	0.944	1.106	1.295	0.910	–
λ	-0.783	-0.014	0.719	-0.899	-0.044	0.865	0.946	2.913
BTV(1,1)								
μ	-1.030	0.029	1.082	-0.932	0.051	0.965	0.965	–
σ	0.901	1.054	1.230	0.938	1.092	1.270	0.916	–
λ	-0.614	-0.016	0.608	-0.766	-0.035	0.739	0.961	1.901
$\lambda = 2.5$								
BTV(1/2,1/2)								
μ	-0.312	0.037	0.546	-0.278	0.090	0.602	0.922	–
σ	0.706	0.947	1.196	0.734	0.950	1.203	0.923	–
λ	0.540	1.929	4.478	0.766	2.141	5.002	0.904	0.122
Jeffreys								
μ	-0.305	0.031	0.535	-0.277	0.089	0.604	0.925	–
σ	0.704	0.949	1.203	0.737	0.955	1.204	0.921	–
λ	0.565	1.951	4.205	0.742	2.137	5.028	0.911	0.115
BTV(1,1)								
μ	-0.264	0.074	0.642	-0.237	0.134	0.676	0.918	–
σ	0.685	0.917	1.174	0.718	0.926	1.180	0.913	–
λ	0.496	1.725	3.919	0.675	1.917	4.491	0.894	0.099
$\lambda = 5$								
BTV(1/2,1/2)								
μ	-0.222	-0.003	0.260	-0.198	0.018	0.294	0.931	–
σ	0.780	0.973	1.168	0.792	0.980	1.175	0.942	–
λ	1.649	4.077	10.223	2.174	4.733	13.117	0.916	0.006
Jeffreys								
μ	-0.210	0.001	0.264	-0.200	0.015	0.287	0.933	–
σ	0.788	0.975	1.168	0.796	0.981	1.172	0.943	–
λ	1.507	4.109	9.910	2.178	4.766	13.401	0.907	0.006
BTV(1,1)								
μ	-0.182	0.020	0.303	-0.170	0.040	0.330	0.930	–
σ	0.761	0.961	1.156	0.774	0.967	1.155	0.932	–
λ	1.601	3.792	7.933	1.964	4.338	10.000	0.909	0.005

Table 3S: Skew-logistic data for noninformative priors: $\mu = 0, \sigma = 1, n = 100$.

Prior	MAP			Median			Coverage	BF
	5%	50%	95%	5%	50%	95%		
$\lambda = 0$								
BTV(1/2,1/2)								
μ	-0.907	-0.022	0.904	-0.824	-0.010	0.807	0.938	–
σ	0.931	1.037	1.168	0.950	1.059	1.187	0.913	–
λ	-0.529	-0.002	0.540	-0.594	-0.002	0.632	0.938	3.838
Jeffreys								
μ	-0.913	-0.012	0.902	-0.826	-0.007	0.820	0.936	–
σ	0.932	1.038	1.165	0.949	1.061	1.191	0.915	–
λ	-0.531	0.011	0.566	-0.606	0.002	0.662	0.931	4.186
BTV(1,1)								
μ	-0.879	-0.013	0.837	-0.780	-0.012	0.755	0.945	–
σ	0.926	1.035	1.154	0.946	1.058	1.181	0.921	–
λ	-0.481	0.007	0.542	-0.531	0.003	0.618	0.942	2.574
$\lambda = 2.5$								
BTV(1/2,1/2)								
μ	-0.225	0.011	0.363	-0.207	0.037	0.382	0.930	–
σ	0.787	0.975	1.142	0.804	0.975	1.142	0.930	–
λ	1.069	2.233	3.760	1.187	2.330	3.999	0.931	0.001
Jeffreys								
μ	-0.228	0.014	0.334	-0.211	0.034	0.392	0.924	–
σ	0.790	0.974	1.143	0.804	0.977	1.145	0.926	–
λ	1.042	2.253	3.745	1.188	2.349	3.936	0.927	0.001
BTV(1,1)								
μ	-0.207	0.029	0.370	-0.194	0.052	0.427	0.918	–
σ	0.777	0.966	1.128	0.790	0.968	1.128	0.918	–
λ	0.954	2.174	3.617	1.088	2.236	3.800	0.911	0.002
$\lambda = 5$								
BTV(1/2,1/2)								
μ	-0.141	0.003	0.157	-0.132	0.009	0.172	0.934	–
σ	0.859	0.989	1.125	0.859	0.992	1.128	0.940	–
λ	2.773	4.640	7.651	3.013	4.942	8.602	0.942	9×10^{-9}
Jeffreys								
μ	-0.141	0.004	0.169	-0.138	0.010	0.174	0.940	–
σ	0.854	0.987	1.126	0.859	0.992	1.132	0.933	–
λ	2.811	4.615	7.605	3.023	4.957	8.638	0.945	8×10^{-9}
BTV(1,1)								
μ	-0.131	0.012	0.179	-0.123	0.018	0.184	0.936	–
σ	0.851	0.983	1.114	0.851	0.985	1.123	0.932	–
λ	2.704	4.429	7.301	2.883	4.754	8.134	0.939	8×10^{-9}

Table 4S: Skew-logistic data for noninformative priors: $\mu = 0, \sigma = 1, n = 200$.

Prior	MAP			Median			Coverage	BF
	5%	50%	95%	5%	50%	95%		
$\lambda = 0$								
BTV(1/2,1/2)								
μ	-0.342	-0.007	0.347	-0.341	0.006	0.356	0.936	–
σ	0.850	1.021	1.197	0.868	1.038	1.213	0.950	–
λ	-0.240	0.001	0.226	-0.335	0.001	0.304	0.934	6.549
Jeffreys								
μ	-0.343	-0.007	0.350	-0.347	0.006	0.349	0.939	–
σ	0.851	1.021	1.199	0.869	1.038	1.210	0.953	–
λ	-0.244	0.001	0.221	-0.330	0.004	0.303	0.936	6.333
BTV(1,1)								
μ	-0.354	-0.003	0.328	-0.338	0.003	0.333	0.944	–
σ	0.853	1.021	1.200	0.868	1.035	1.212	0.952	–
λ	-0.243	0.001	0.214	-0.319	0.001	0.292	0.934	4.277
$\lambda = 2.5$								
BTV(1/2,1/2)								
μ	-0.173	0.002	0.237	-0.157	0.014	0.240	0.945	–
σ	0.781	0.969	1.195	0.799	0.983	1.216	0.952	–
λ	0.834	2.039	4.492	1.068	2.377	5.605	0.944	0.011
Jeffreys								
μ	-0.173	0.001	0.249	-0.155	0.012	0.238	0.938	–
σ	0.771	0.971	1.203	0.788	0.985	1.217	0.953	–
λ	0.841	2.065	4.552	1.060	2.386	5.598	0.940	0.010
BTV(1,1)								
μ	-0.163	0.012	0.257	-0.145	0.026	0.251	0.938	–
σ	0.772	0.958	1.192	0.788	0.974	1.200	0.944	–
λ	0.778	1.951	4.391	0.991	2.241	5.094	0.946	0.008
$\lambda = 5$								
BTV(1/2,1/2)								
μ	-0.120	0.001	0.147	-0.113	0.008	0.157	0.938	–
σ	0.786	0.974	1.187	0.801	0.983	1.194	0.947	–
λ	1.943	4.172	9.410	2.427	4.877	12.860	0.936	0.002
Jeffreys								
μ	-0.128	-0.000	0.153	-0.117	0.009	0.159	0.943	–
σ	0.787	0.973	1.182	0.798	0.985	1.194	0.945	–
λ	1.950	4.128	10.446	2.381	4.868	12.906	0.936	0.003
BTV(1,1)								
μ	-0.109	0.009	0.164	-0.099	0.017	0.177	0.947	–
σ	0.772	0.962	1.159	0.787	0.974	1.182	0.947	–
λ	1.852	3.848	8.546	2.224	4.479	10.951	0.942	0.002

Table 5S: Skew-Laplace data for noninformative priors: $\mu = 0, \sigma = 1, n = 100$.

Prior	MAP			Median			Coverage	BF
	5%	50%	95%	5%	50%	95%		
$\lambda = 0$								
BTV(1/2,1/2)								
μ	-0.204	-0.002	0.225	-0.209	-0.002	0.223	0.941	–
σ	0.889	1.008	1.136	0.898	1.013	1.142	0.942	–
λ	-0.147	-0.000	0.149	-0.183	0.000	0.185	0.947	10.585
Jeffreys								
μ	-0.206	-0.001	0.224	-0.211	-0.000	0.214	0.939	–
σ	0.884	1.007	1.133	0.895	1.014	1.141	0.939	–
λ	-0.148	0.001	0.154	-0.180	0.001	0.190	0.945	10.440
BTV(1,1)								
μ	-0.208	-0.002	0.211	-0.206	0.001	0.210	0.941	–
σ	0.887	1.007	1.132	0.898	1.014	1.140	0.942	–
λ	-0.138	0.000	0.148	-0.180	-0.000	0.189	0.945	6.852
$\lambda = 2.5$								
BTV(1/2,1/2)								
μ	-0.126	-0.004	0.155	-0.119	0.003	0.155	0.945	–
σ	0.840	0.987	1.147	0.845	0.994	1.159	0.936	–
λ	1.307	2.311	3.849	1.429	2.469	4.217	0.941	$7{\times}10^{-6}$
Jeffreys								
μ	-0.127	-0.001	0.157	-0.118	0.004	0.161	0.940	–
σ	0.836	0.982	1.148	0.851	0.992	1.154	0.939	–
λ	1.289	2.291	3.830	1.403	2.460	4.176	0.944	$6{\times}10^{-6}$
BTV(1,1)								
μ	-0.119	0.004	0.159	-0.110	0.009	0.160	0.945	–
σ	0.828	0.978	1.140	0.843	0.988	1.145	0.938	–
λ	1.263	2.210	3.738	1.382	2.385	4.045	0.945	$5{\times}10^{-6}$
$\lambda = 5$								
BTV(1/2,1/2)								
μ	-0.084	-0.001	0.097	-0.080	0.002	0.096	0.951	–
σ	0.849	0.984	1.137	0.855	0.991	1.142	0.939	–
λ	2.817	4.596	7.956	3.036	4.900	8.621	0.952	$7{\times}10^{-8}$
Jeffreys								
μ	-0.085	-0.002	0.095	-0.079	0.002	0.097	0.951	–
σ	0.847	0.987	1.142	0.853	0.990	1.142	0.940	–
λ	2.784	4.567	7.767	3.033	4.879	8.693	0.945	$1{\times}10^{-7}$
BTV(1,1)								
μ	-0.080	0.004	0.099	-0.074	0.006	0.102	0.951	–
σ	0.840	0.978	1.128	0.849	0.986	1.135	0.932	–
λ	2.615	4.427	7.547	2.859	4.744	8.132	0.943	$7{\times}10^{-8}$

Table 6S: Skew-Laplace data for noninformative priors: $\mu = 0, \sigma = 1, n = 200$.