## Supplementary Material

Conditional independence testing based on a nearest-neighbor estimator of conditional mutual information

Jakob Runge

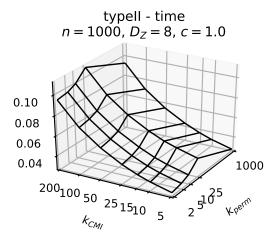


Figure S1: Runtime per estimate [in s] for the same setup as in Fig. 3 but with n = 1,000. For  $k_{\text{perm}} = n$  a computationally cheaper full permutation scheme was used.

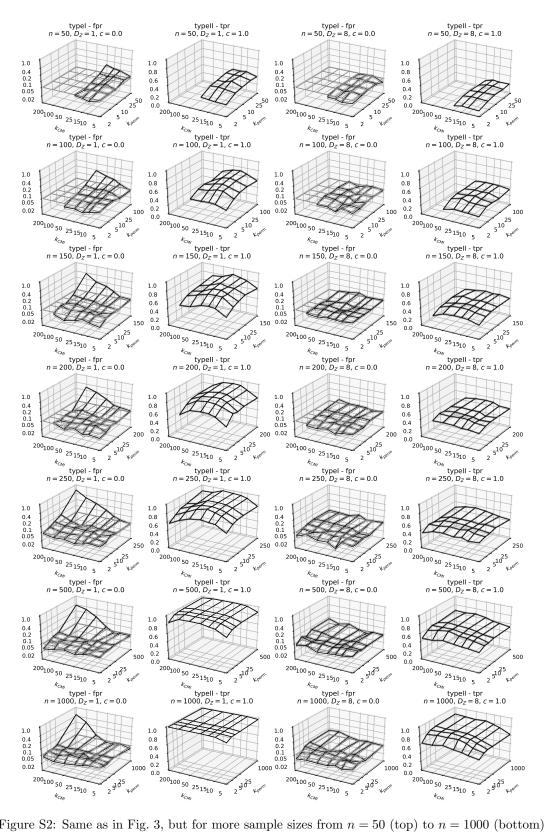


Figure S2: Same as in Fig. 3, but for more sample sizes from n = 50 (top) to n = 1000 (bottom).

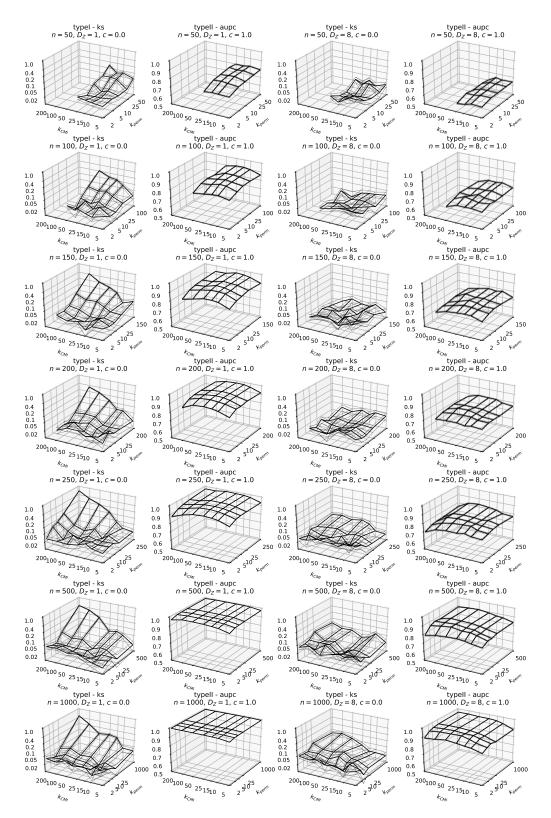


Figure S3: Same as in Fig. 3, but for KS and AUPC metrics and more sample sizes from n = 50 (top) to n = 1000 (bottom).

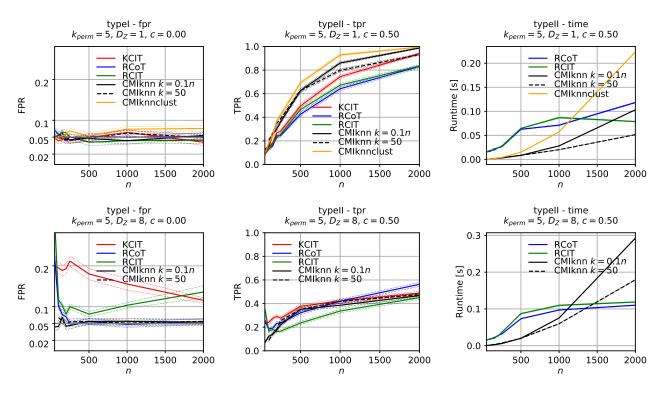


Figure S4: As in Fig. 4 but for false positive rates (FPR) and true positive rates (TPR) at an  $\alpha = 0.05$  significance level.

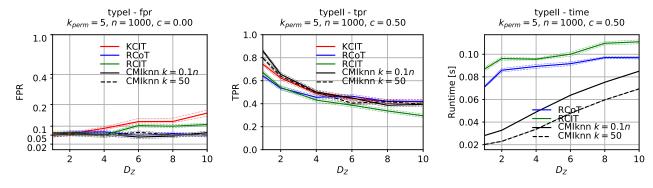


Figure S5: As in Fig. 5 but for false positive rates (FPR) and true positive rates (TPR) at an  $\alpha=0.05$  significance level.

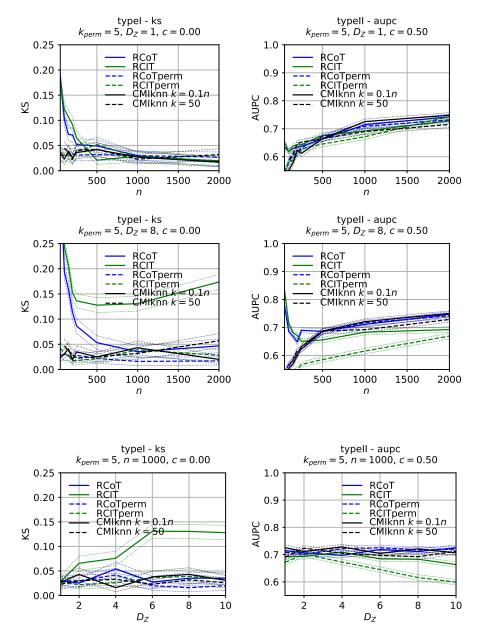


Figure S6: Numerical experiments for a version of model (6) where Z is independent of X and Y when X and Y are dependent under  $H_1$ , that is  $X = g_X(c\epsilon_b + \epsilon_X)$  and  $Y = g_Y(c\epsilon_b + \epsilon_Y)$ , which is the setup studied in Zhang et al. (2011); Strobl et al. (2017). Additionally, we show results for the kernel measures combined with the proposed nearest-neighbor permutation test with  $k_{\text{perm}} = 5$ . The upper two rows depict KS and AUPC for different sample sizes and  $D_Z = 1, 8$ . The two bottom panels show different dimensions  $D_Z$  for n = 1000.

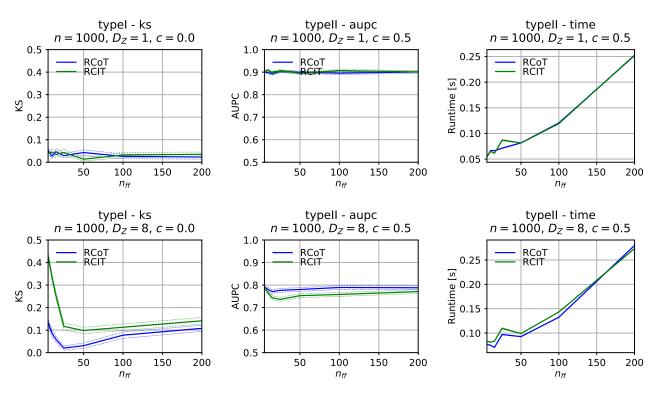


Figure S7: Choice of number of fourier features  $(n_{\rm ff})$  for random fourier-feature based kernel-measures for model (6). Shown are KS (left column), AUPC (center column), and runtime (right column) for a sample size experiment with  $D_Z=1$  (top row) and  $D_Z=8$  (bottom row).  $n_{\rm ff}$  corresponds to the number of features in subspace Z, the number of fourier features in subspaces X and Y is fixed to 5 as implemented in https://github.com/ericstrobl/RCIT. While for  $D_Z=1$   $n_{\rm ff}>10$  yields similar results, for  $D_Z=8$  both the KS and AUPC metrics are more sensitive to the choice of  $n_{\rm ff}$ . The default  $n_{\rm ff}=25$  here gives the most well-calibrated result, but this calibration is quite unstable for smaller or larger  $n_{\rm ff}$  values. The runtime of RCIT and RCoT scales roughly quadratically in the number of fourier features.

Table S1: Results from Wang et al. (2015) together with results from RCoT and the CMIknn test. The experiments are described in Wang et al. (2015). Examples 1–4 correspond to conditional independence showing false positives and Examples 5–8 to dependent cases showing true positives at the 5% significance level. CMIknn was run with  $k_{\rm CMI}=0.2n$  and  $k_{\rm perm}=5,10$ . The numbers 50..250 denote the sample size.

•	Example 1					Example 2				
Test	50	100	150	200	250	50	100	150	200	250
CDIT	0.035	0.034	0.05	0.057	0.048	0.046	0.053	0.055	0.048	0.058
CI.test	0.041	0.051	0.037	0.054	0.041	0.062	0.046	0.044	0.045	0.039
KCI.test	0.039	0.043	0.041	0.04	0.046	0.035	0.004	0.037	0.047	0.05
Rule-of-thumb	0.017	0.027	0.028	0.033	0.033	0.034	0.052	0.044	0.042	0.045
RCoT	0.074	0.059	0.055	0.043	0.050	0.056	0.056	0.069	0.055	0.073
CMIknn $(k_{\text{perm}} = 5)$	0.064	0.055	0.050	0.053	0.045	0.076	0.060	0.074	0.061	0.065
CMIknn $(k_{\text{perm}} = 10)$	0.058	0.061	0.057	0.058	0.046	0.075	0.066	0.053	0.057	0.071
	Example 3					Example 4				
Test	50	100	150	200	250	50	100	150	200	250
CDIT	0.035	0.048	0.055	0.053	0.043	0.049	0.054	0.051	0.058	$\frac{250}{0.053}$
CI.test	0.039	0.363	0.482	0.603	0.677	0.043	0.064	0.066	0.050	0.053
KCI.test	0.058	0.047	0.057	0.061	0.054	0.037	0.035	0.058	0.039	0.049
Rule-of-thumb	0.019	0.038	0.032	0.039	0.039	0.037	0.04	0.055	0.059	0.053
RCoT	0.074	0.047	0.046	0.053	0.054	0.115	0.072	0.066	0.061	0.053
CMIknn $(k_{\text{perm}} = 5)$	0.044	0.043	0.046	0.046	0.054	0.084	0.071	0.067	0.079	0.070
CMIknn $(k_{\text{perm}} = 10)$	0.063	0.065	0.061	0.076	0.067	0.101	0.113	0.106	0.098	0.084
	E1. F									
		_	, ,	_						
T	F0		Example		250	F0.		xample (		250
Test	50	100	150	200	250	50	100	150	200	250
$\overline{ ext{CDIT}}$	0.898	100 0.993	150	200	1	0.752	100 0.995	150	200	1
CDIT CI.test	0.898 0.978	100 0.993 1	150 1 1	200 1 1	1 1	0.752 0.468	100 0.995 0.434	150 1 0.467	200 1 0.476	1 0.474
CDIT CI.test KCI.test	0.898 0.978 0.158	100 0.993 1 0.481	150 1 1 0.557	200 1 1 0.602	1 1 0.742	0.752 0.468 0.296	100 0.995 0.434 0.862	150 1 0.467 0.995	200 1 0.476 1	1 0.474 1
CDIT CI.test KCI.test Rule-of-thumb	0.898 0.978 0.158 0.368	100 0.993 1 0.481 0.793	150 1 1 0.557 0.927	200 1 1 0.602 0.983	1 1 0.742 0.994	0.752 0.468 0.296 1	100 0.995 0.434 0.862 1	150 1 0.467 0.995 1	200 1 0.476 1 1	1 0.474 1 1
CDIT CI.test KCI.test Rule-of-thumb RCoT	0.898 0.978 0.158 0.368 0.817	100 0.993 1 0.481 0.793 0.986	150 1 1 0.557 0.927 0.998	200 1 1 0.602 0.983 1	1 1 0.742 0.994 1	0.752 0.468 0.296 1 0.301	100 0.995 0.434 0.862 1 0.533	150 1 0.467 0.995 1 0.679	200 1 0.476 1 1 0.807	1 0.474 1 1 0.860
CDIT CI.test KCI.test Rule-of-thumb RCoT CMIknn $(k_{perm} = 5)$	0.898 0.978 0.158 0.368 0.817 0.782	100 0.993 1 0.481 0.793 0.986 0.981	150 1 1 0.557 0.927 0.998 0.998	200 1 1 0.602 0.983 1 1	1 1 0.742 0.994 1	0.752 0.468 0.296 1 0.301 0.806	100 0.995 0.434 0.862 1 0.533 0.997	150 1 0.467 0.995 1 0.679 0.999	200 1 0.476 1 1 0.807 1	1 0.474 1 1 0.860 1
CDIT CI.test KCI.test Rule-of-thumb RCoT	0.898 0.978 0.158 0.368 0.817	100 0.993 1 0.481 0.793 0.986	150 1 1 0.557 0.927 0.998	200 1 1 0.602 0.983 1	1 1 0.742 0.994 1	0.752 0.468 0.296 1 0.301	100 0.995 0.434 0.862 1 0.533	150 1 0.467 0.995 1 0.679	200 1 0.476 1 1 0.807	1 0.474 1 1 0.860
CDIT CI.test KCI.test Rule-of-thumb RCoT CMIknn $(k_{perm} = 5)$	0.898 0.978 0.158 0.368 0.817 0.782	0.993 1 0.481 0.793 0.986 0.981 0.995	150 1 1 0.557 0.927 0.998 0.998	200 1 1 0.602 0.983 1 1	1 1 0.742 0.994 1	0.752 0.468 0.296 1 0.301 0.806	100 0.995 0.434 0.862 1 0.533 0.997 0.995	150 1 0.467 0.995 1 0.679 0.999	200 1 0.476 1 1 0.807 1	1 0.474 1 1 0.860 1
CDIT CI.test KCI.test Rule-of-thumb RCoT CMIknn $(k_{perm} = 5)$	0.898 0.978 0.158 0.368 0.817 0.782	0.993 1 0.481 0.793 0.986 0.981 0.995	150 1 1 0.557 0.927 0.998 0.998 1	200 1 1 0.602 0.983 1 1	1 1 0.742 0.994 1	0.752 0.468 0.296 1 0.301 0.806	100 0.995 0.434 0.862 1 0.533 0.997 0.995	150 1 0.467 0.995 1 0.679 0.999 1	200 1 0.476 1 1 0.807 1	1 0.474 1 1 0.860 1
CDIT CI.test KCI.test Rule-of-thumb RCoT CMIknn $(k_{perm} = 5)$ CMIknn $(k_{perm} = 10)$	0.898 0.978 0.158 0.368 0.817 0.782 0.855	100 0.993 1 0.481 0.793 0.986 0.981 0.995	150 1 1 0.557 0.927 0.998 0.998 1	200 1 1 0.602 0.983 1 1 1	1 1 0.742 0.994 1 1	0.752 0.468 0.296 1 0.301 0.806 0.805	100 0.995 0.434 0.862 1 0.533 0.997 0.995	150 1 0.467 0.995 1 0.679 0.999 1	200 1 0.476 1 1 0.807 1 1	1 0.474 1 1 0.860 1
CDIT CI.test KCI.test Rule-of-thumb RCoT CMIknn $(k_{perm} = 5)$ CMIknn $(k_{perm} = 10)$ Test CDIT CI.test	0.898 0.978 0.158 0.368 0.817 0.782 0.855	100 0.993 1 0.481 0.793 0.986 0.981 0.995 E 100 0.998 0.984	150 1 1 0.557 0.927 0.998 0.998 1 Example 150	200 1 1 0.602 0.983 1 1 7 200	1 1 0.742 0.994 1 1 1	0.752 0.468 0.296 1 0.301 0.806 0.805	100 0.995 0.434 0.862 1 0.533 0.997 0.995	150 1 0.467 0.995 1 0.679 0.999 1 example 8	200 1 0.476 1 1 0.807 1 1 8	1 0.474 1 1 0.860 1 1 250 0.994 0.485
CDIT CI.test KCI.test Rule-of-thumb RCoT CMIknn $(k_{perm} = 5)$ CMIknn $(k_{perm} = 10)$ Test CDIT	0.898 0.978 0.158 0.368 0.817 0.782 0.855 50 0.918 0.953 0.574	100 0.993 1 0.481 0.793 0.986 0.981 0.995 F 100 0.998 0.984 0.947	150 1 1 0.557 0.927 0.998 0.998 1 Example 150 1	200 1 1 0.602 0.983 1 1 1 7 200 1	1 1 0.742 0.994 1 1 1 250	0.752 0.468 0.296 1 0.301 0.806 0.805	100 0.995 0.434 0.862 1 0.533 0.997 0.995 Example 100 0.731 0.476 0.401	150 1 0.467 0.995 1 0.679 0.999 1 example 8 150 0.949	200 1 0.476 1 1 0.807 1 1 8 200 0.977	1 0.474 1 1 0.860 1 1 250 0.994 0.485 1
CDIT CI.test KCI.test Rule-of-thumb RCoT CMIknn $(k_{perm} = 5)$ CMIknn $(k_{perm} = 10)$ Test CDIT CI.test KCI.test Rule-of-thumb	0.898 0.978 0.158 0.368 0.817 0.782 0.855 50 0.918 0.953	100 0.993 1 0.481 0.793 0.986 0.981 0.995 F 100 0.998 0.984 0.947 0.302	150 1 1 0.557 0.927 0.998 0.998 1 0.983 0.998 0.385	200 1 1 0.602 0.983 1 1 1 7 200 1 0.995 1 0.514	1 1 0.742 0.994 1 1 1 1 250 1 0.987 1 0.515	0.752 0.468 0.296 1 0.301 0.806 0.805 50 0.361 0.456 0.089 0.043	100 0.995 0.434 0.862 1 0.533 0.997 0.995 E: 100 0.731 0.476 0.401 0.233	150 1 0.467 0.995 1 0.679 0.999 1 example 3 150 0.949 0.464 0.685 0.551	200 1 0.476 1 1 0.807 1 1 8 200 0.977 0.461 1 0.851	1 0.474 1 1 0.860 1 1 250 0.994 0.485 1 0.972
CDIT CI.test KCI.test Rule-of-thumb RCoT CMIknn $(k_{perm} = 5)$ CMIknn $(k_{perm} = 10)$ Test CDIT CI.test KCI.test Rule-of-thumb RCoT	0.898 0.978 0.158 0.368 0.817 0.782 0.855 50 0.918 0.953 0.574 0.073 0.594	100 0.993 1 0.481 0.793 0.986 0.981 0.995 E 100 0.998 0.984 0.947 0.302 0.880	150 1 1 0.557 0.927 0.998 0.998 1 0.983 0.998 0.385 0.962	200 1 1 0.602 0.983 1 1 1 7 200 1 0.995 1 0.514 0.985	1 1 0.742 0.994 1 1 1 250 1 0.987	0.752 0.468 0.296 1 0.301 0.806 0.805 50 0.361 0.456 0.089	100 0.995 0.434 0.862 1 0.533 0.997 0.995 E: 100 0.731 0.476 0.401 0.233 0.392	150 1 0.467 0.995 1 0.679 0.999 1 example 8 150 0.949 0.464 0.685 0.551 0.470	200 1 0.476 1 1 0.807 1 1 8 200 0.977 0.461 1	1 0.474 1 1 0.860 1 1 250 0.994 0.485 1 0.972 0.654
CDIT CI.test KCI.test Rule-of-thumb RCoT CMIknn $(k_{perm} = 5)$ CMIknn $(k_{perm} = 10)$ Test CDIT CI.test KCI.test Rule-of-thumb	0.898 0.978 0.158 0.368 0.817 0.782 0.855 50 0.918 0.953 0.574 0.073	100 0.993 1 0.481 0.793 0.986 0.981 0.995 F 100 0.998 0.984 0.947 0.302	150 1 1 0.557 0.927 0.998 0.998 1 0.983 0.998 0.385	200 1 1 0.602 0.983 1 1 1 7 200 1 0.995 1 0.514	1 1 0.742 0.994 1 1 1 1 250 1 0.987 1 0.515	0.752 0.468 0.296 1 0.301 0.806 0.805 50 0.361 0.456 0.089 0.043	100 0.995 0.434 0.862 1 0.533 0.997 0.995 E: 100 0.731 0.476 0.401 0.233	150 1 0.467 0.995 1 0.679 0.999 1 example 3 150 0.949 0.464 0.685 0.551	200 1 0.476 1 1 0.807 1 1 8 200 0.977 0.461 1 0.851	1 0.474 1 1 0.860 1 1 250 0.994 0.485 1 0.972