The following is the supplementary Appendix for the paper; *Probabilistic Entity Representation Model for Reasoning over Knowledge Graphs*. All the references given in the following sections are made in context of the main paper.

A Derivation for Product of Multivariate Gaussians

The following sections provide the proof for the product of Gaussians for both the univariate case and multivariate case (used in Eqs. (6) and (10)).

A.1 Univariate Case

$$\begin{split} \mathcal{N}(\mu,\sigma) &= \exp\left(\left(\frac{x-\mu}{\sigma}\right)^2\right) \\ P(\theta) &= P(\theta_1)P(\theta_2) = \exp\left(\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right) \cdot \exp\left(\left(\frac{x-\mu_2}{\sigma_2}\right)^2\right) \\ \log P(\theta) &= \left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x-\mu_2}{\sigma_2}\right)^2 \\ &= \frac{(\sigma_2^2 + \sigma_1^2)x^2 - 2(\sigma_1^2\mu_2 + \sigma_2^2\mu_1)x + (\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2)}{\sigma_1^2\sigma_2^2} \\ &= \frac{x^2 - 2\frac{(\sigma_1^2\mu_2 + \sigma_2^2\mu_1)}{\sigma_2^2 + \sigma_1^2}x + \frac{\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2}{\sigma_2^2 + \sigma_1^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}} \\ &= \left(\frac{x - \frac{(\sigma_1^2\mu_2 + \sigma_2^2\mu_1)}{\sigma_2^2 + \sigma_1^2}}{\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)^2}\right)^2 + K, \text{where } K = \frac{\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2}{\sigma_1^2\sigma_2^2} - \left(\frac{\sigma_1^2\mu_2 + \sigma_2^2\mu_1}{\sigma_1^2\sigma_2^2}\right)^2 \\ P(\theta) \propto \exp\left(\left(\frac{x - \frac{(\sigma_1^2\mu_2 + \sigma_2^2\mu_1)}{\sigma_2^2 + \sigma_1^2}}{\left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}\right)^{-2}}\right)^2\right) \approx \mathcal{N}\left(\frac{(\sigma_1^2\mu_2 + \sigma_2^2\mu_1)}{\sigma_2^2 + \sigma_1^2}, (\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2})^{-2}\right) \end{split}$$

A.2 Multivariate Case

$$\begin{split} \mathcal{N}(\mu, \Sigma) &= \exp\left((x-\mu)^T \Sigma^{-1} (x-\mu)\right) \\ P(\theta) &= P(\theta_1) P(\theta_2) = \exp\left((x-\mu_1)^T \Sigma_1^{-1} (x-\mu_1)\right) . \exp\left((x-\mu_2)^T \Sigma_2^{-1} (x-\mu_2)\right) \\ \log(P(\theta)) &= (x-\mu_1)^T \Sigma_1^{-1} (x-\mu_1) + (x-\mu_2)^T \Sigma_2^{-1} (x-\mu_2) \\ &= x^T \Sigma_1^{-1} x - \mu_1^T \Sigma_1^{-1} x - x^T \Sigma_1^{-1} \mu_1 - \mu_1^T \Sigma_1^{-1} \mu_1 + x^T \Sigma_2^{-1} x - \mu_2^T \Sigma_2^{-1} x - x^T \Sigma_2^{-1} \mu_2 - \mu_2^T \Sigma_2^{-1} \mu_2 \\ &= x^T (\Sigma_1^{-1} + \Sigma_2^{-1}) x - (\mu_1^T \Sigma_1^{-1} + \mu_2^T \Sigma_2^{-1}) x - x^T (\Sigma_1^{-1} \mu_1 + \Sigma_2^{-1} \mu_2) - (\mu_1^T \Sigma_1^{-1} \mu_1 + \mu_2^T \Sigma_2^{-1} \mu_2) \\ Let's assume P(\theta) \propto \mathcal{N}(\mu_3, \Sigma_3), \text{ then,} \end{split}$$

 $log(P(\theta)) = (x - \mu_3)^T \Sigma_3^{-1} (x - \mu_3) + K$ = $x^T \Sigma_3^{-1} x - x^T \Sigma_3^{-1} \mu_3 - \mu_3^T \Sigma_3^{-1} x + \mu_3^T \Sigma_3^{-1} \mu_3 + K$

Comparing coefficients,

$$\begin{split} \Sigma_3^{-1} &= \Sigma_1^{-1} + \Sigma_2^{-1} \\ \Sigma_3^{-1} \mu_3 &= \Sigma_1^{-1} \mu_1 + \Sigma_2^{-1} \mu_2 \\ \implies \mu_3 &= \Sigma_3 (\Sigma_1^{-1} \mu_1 + \Sigma_2^{-1} \mu_2) \\ \mu_3 &= (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1} (\Sigma_1^{-1} \mu_1 + \Sigma_2^{-1} \mu_2) \end{split}$$

Notice that we need Σ_3 while calculation μ_3 . However, to save computational memory, we only store the inverses of covariances, i.e., Σ_1^{-1} , Σ_2^{-1} and Σ_3^{-1} . So, to solve for μ_3 and avoid the computationally

expensive process of matrix inversion, we use the linear solver torch.solve on the equation $\Sigma_3^{-1}\mu_3 = \Sigma_1^{-1}\mu_1 + \Sigma_2^{-1}\mu_2$.

B Algorithm for KG Reasoning with PERM

Algorithm 1 provides an outline of PERM's overall framework to learn representations of entities $e \in E$ and relations $r \in R$. The algorithm describes the training from FOE operations of translation (lines 4.7), intersection (lines 8.11), and union (lines 12.15).

Algorithm 1: PERM training algorithm

Input: Training data D_t, D_{\cap}, D_{\cup} , which are set of all (query (Q), result (V)) for translation, intersection, and union, respectively; **Output:** Entity E and Relation R gaussian density functions; 1 Randomly initialize $e = \mathcal{N}(\mu_e, \Sigma_e) \in E$ and $r = \mathcal{N}(\mu_r, \Sigma_r) \in R$; 2 for number of epochs; until convergence do l = 0; # Initialize loss 3 for $\{(e, r, V_t) \in D_t\}$ do 4 $q_t = \mathcal{N}(\mu_e + \mu_r, (\Sigma_e^{-1} + \Sigma_r^{-1})^{-1})$ from Eq. (5) # Update loss for translation queries 5 $l = l + \sum_{v_t \in V_t} d_{\mathcal{N}}(v_t, q_t)$ 6 7 end for $\{(Q_{\cap}, V_{\cap}) \in D_{\cap}\}$ do 8 $q_{\cap} = \mathcal{N}(\mu_3, \Sigma_3)$, from Eq. (6) 9 # Update loss for intersection queries $l = l + \sum_{v_{\cap} \in V_{\cap}} d_{\mathcal{N}}(v_{\cap}, q_{\cap})$ 10 11 end for $\{(Q_{\cup}, V_{\cup}) \in D_{\cup}\}$ do $q_{\cup} = \sum_{i=1}^{n} \phi_i \mathcal{N}(\mu_{e_i}, \Sigma_{e_i})$ from Eq. (7) # Update loss for union queries $l = l + \sum_{v_{\cup} \in V_{\cup}} \sum_{i=1}^{n} \phi_i d_{\mathcal{N}}(v_{\cup}, \mathcal{N}(\mu_{e_i}, \Sigma_{e_i}))$ 12 13 14 end 15 # Update E and R with backpropagation $E \leftarrow E - \Delta_E l$ 16 $R \leftarrow R - \Delta_R l$ 18 19 end 20 return E,R

C MRR metrics for Reasoning over KGs

Table 6 provides the Mean Reciprocal Rank (MRR) results for the reasoning over KGs experiment, given in section 5

D Finer Evaluation of Ablation Study

Table 7 provides finer results of our ablation study.

Table 6: Performance comparison of PERM against the baselines to study the efficacy of the query representations. The columns present the different query structures and the overall average performance. The last two rows presents the Average Relative Improvement (%) of PERM compared to Q2B and CQD over all datasets across query types. Best results for each dataset are shown in bold.

Metrics					Mea	n Recip	rocal F	Rank			
Dataset	Model	1t	2t	3t	$2\cap$	$3\cap$	$2\cup$	$\cap t$	$t\cap$	$\cup t$	Avg
FB15k-237	GQE	.346	.191	.144	.258	.361	.144	.087	.164	.149	.205
	BQE	.390	.109	.100	.228	.425	.124	.224	.126	.097	.203
	Q2B	.400	.225	.173	.275	.378	.198	.105	.180	.178	.235
	CQD	.439	.270	.206	.299	.381	.235	.271	.415	.112	.292
	PERM	.445	.268	.201	.306	.409	.253	.269	.353	.220	.303
NELL995	GQE	.311	.193	.175	.273	.399	.159	.078	.168	.130	.210
	BQE	.530	.130	.114	.376	.475	.122	.241	.143	.085	.246
	Q2B	.413	.227	.208	.288	.414	.266	.125	.193	.155	.254
	CQD	.442	.251	.226	.304	.441	.348	.124	.212	.104	.273
	PERM	.432	.244	.217	.296	.438	.332	.122	.178	.190	.272
DBPedia	GQE	.502	.005	N.A.	.749	.773	.320	.154	.597	0.00	.388
	BQE	.657	.006	N.A.	.964	.966	.306	.419	.527	0.00	.481
	Q2B	.619	.006	N.A.	.840	.863	.468	.212	.779	0.00	.473
	CQD	.648	.006	N.A.	.840	.863	.485	.206	.716	0.00	.471
	PERM	.706	.006	N.A.	.841	.862	.564	.219	.869	0.00	.508
DRKG	GQE	.313	.182	.132	.232	.360	.144	.097	.166	.163	.199
	BQE	.413	.118	.106	.298	.451	.147	.270	.154	.116	.230
	Q2B	.371	.225	.178	.283	.422	.205	.064	.122	.223	.233
	CQD	.413	.277	.213	.310	.427	.246	.174	.282	.143	.276
	PERM	.420	.276	.211	.325	.465	.271	.179	.249	.282	.298
PERM vs Q		11.1	16.3	12.5	4.90	4.70	24.9	55.9	29.4	24.5	20.5
PERM vs C	QD (%)	4.20	-1.2	-2.5	1.40	4.40	10.6	1.80	1.50	92.8	12.6

Madad	1	-				TTT	503				
Metrics	Varianta	1t	2t	3t	20		$\frac{8@3}{2\cup}$	<u></u>	+0	1.14	Ava
Dataset	Variants				2∩ 282	<u>3</u> ∩		$\cap t$	$t \cap$	$\cup t$	Avg
FB15k-237	1t	.516	.179	.119	.282	.360	.302	.071	.134	.133	.233
	translations	.516	.231	.167	.318	.413	.304	.096	.160	.185	.266
	single	.511	.282	.212	.359	.486	.296	.126	.207	.235	.302
	average	.499	.282	.209	.360	.482	.282	.119	.201	.234	.296
	MLP	.510	.285	.212	.363	.488	.293	.125	.208	.238	.302
	(final)	.520	.286	.216	.361	.490	.305	.128	.212	.239	.306
NELL995	lt	.576	.179	.134	.275	.373	.456	.072	.123	.111	.255
	translations	.576	.231	.188	.310	.428	.458	.097	.147	.155	.288
	single	.571	.282	.239	.350	.504	.446	.127	.190	.197	.323
	average	.558	.282	.235	.351	.500	.425	.120	.185	.196	.317
	MLP	.570	.285	.239	.354	.506	.442	.126	.191	.199	.324
	(final)	.581	.286	.243	.352	.508	.460	.129	.195	.200	.328
DBPedia	1t	.942	.004	N.A.	.781	.734	.775	.129	.600	0.00	.496
	translations	.942	.006	N.A.	.881	.843	.779	.174	.718	0.00	.543
	single	.934	.007	N.A.	1.00	1.00	.758	.228	.928	0.00	.607
	average	.912	.007	N.A.	.997	.984	.723	.216	.903	0.00	.593
	MLP	.932	.007	N.A.	.996	.992	.751	.227	.932	0.00	.605
	(final)	.950	.007	N.A.	1.00	1.00	.782	.232	.952	0.00	.615
DRKG	lt	.560	.202	.130	.302	.396	.373	.106	.172	.165	.267
	translations	.560	.260	.183	.341	.455	.374	.143	.206	.230	.306
	single	.555	.317	.232	.385	.536	.365	.187	.266	.293	.348
	average	.543	.317	.228	.386	.531	.347	.177	.259	.291	.342
	MLP	.554	.321	.232	.389	.538	.361	.186	.267	.296	.349
	(final)	.565	.322	.236	.387	.540	.376	.190	.273	.297	.354
Metrics	()		.522	.230					.270		
Metrics Dataset		1t	2t	3t		n Recip 3∩			t∩	$\cup t$	
Dataset	Variants	1t	2t	3t	Mea 2∩	n Recip 3∩	rocal R $2 \cup$	ank $\cap t$	$t\cap$		Avg
	Variants PERM-1t	1t .410	2t .180	3t .122	Mea 2∩ .217	n Recip 3∩ .274	rocal R 2∪ .209	ank ∩t .085	<i>t</i> ∩ .127	<i>∪t</i> .145	Avg .197
Dataset	Variants PERM-1t translations	1t .410 .410	2t .180 .232	3t .122 .171	Mean 2∩ .217 .245	n Recip 3∩ .274 .314	rocal R 2∪ .209 .210	ank ∩t .085 .115	<i>t</i> ∩ .127 .152	<i>∪t</i> .145 .202	Avg .197 .228
Dataset	Variants PERM-1t translations single	1t .410 .410 .406	2t .180 .232 .283	3t .122 .171 .217	Mea 2∩ .217 .245 .277	n Recip 3∩ .274 .314 .370	rocal R 2∪ .209 .210 .204	ank ∩t .085 .115 .151	<i>t</i> ∩ .127 .152 .197	∪t .145 .202 .257	Avg .197 .228 .262
Dataset	Variants PERM-1t translations single average	1t .410 .410 .406 .396	2t .180 .232 .283 .283	3t .122 .171 .217 .214	Mear 2∩ .217 .245 .277 .278	n Recip 3∩ .274 .314 .370 .367	rocal R 2∪ .209 .210 .204 .194	Aank ∩t .085 .115 .151 .143	<i>t</i> ∩ .127 .152 .197 .191	Ut .145 .202 .257 .256	Avg .197 .228 .262 .258
Dataset	Variants PERM-1t translations single average MLP	1t .410 .410 .406 .396 .405	2t .180 .232 .283 .283 .286	3t .122 .171 .217 .214 .217	Mea 2∩ .217 .245 .277 .278 .280	n Recip 3∩ .274 .314 .370 .367 .372	rocal R <u>2</u> ∪ .209 .210 .204 .194 .202	cank ∩t .085 .115 .151 .143 .150	<i>t</i> ∩ .127 .152 .197 .191 .198	$\cup t$.145 .202 .257 .256 .260	Avg .197 .228 .262 .258 .263
Dataset FB15k-237	Variants PERM-1t translations single average MLP (final)	1t .410 .406 .396 .405 .445	2t .180 .232 .283 .283 .286 .268	3t .122 .171 .217 .214 .217 .201	Mea 2∩ .217 .245 .277 .278 .280 .306	n Recip 3∩ .274 .314 .370 .367 .372 .409	rocal R <u>2</u> ∪ .209 .210 .204 .194 .202 .253	Image: constraint of the second se	<i>t</i> ∩ .127 .152 .197 .191 .198 .353	Ut .145 .202 .257 .256 .260 .220	Avg .197 .228 .262 .258 .263 .303
Dataset	Variants PERM-1t translations single average MLP (final) 1t	1t .410 .410 .406 .396 .405 .445 .432	2t .180 .232 .283 .283 .283 .286 .268 .191	3t .122 .171 .217 .214 .217 .201 .160	Mea 2∩ .217 .245 .277 .278 .280 .306 .234	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275	rocal R <u>2</u> ∪ .209 .210 .204 .194 .202 .253 .332	Image: constraint of the system .085 .115 .151 .143 .150 .269 .094	<i>t</i> ∩ .127 .152 .197 .191 .198 .353 .162	Ut .145 .202 .257 .256 .260 .220 .125	Avg .197 .228 .262 .258 .263 .303 .223
Dataset FB15k-237	Variants PERM-1t translations single average MLP (final) 1t translations	1t .410 .406 .396 .405 .445 .432 .428	2t .180 .232 .283 .283 .286 .268 .191 .197	3t .122 .171 .217 .214 .217 .201 .160 .168	Mean 2∩ .217 .245 .277 .278 .280 .306 .234 .261	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369	rocal R 2 ∪ .209 .210 .204 .194 .202 .253 .332 .331	ank ∩t .085 .115 .151 .143 .150 .269 .094 .092	<i>t</i> ∩ .127 .152 .197 .191 .198 .353 .162 .134	Ut .145 .202 .257 .256 .260 .220 .125 .147	Avg .197 .228 .262 .258 .263 .303 .223 .236
Dataset FB15k-237	Variants PERM-1t translations single average MLP (final) 1t translations single	It .410 .406 .396 .405 .445 .432 .428 .425	2t .180 .232 .283 .283 .286 .268 .191 .197 .241	3t .122 .171 .217 .214 .217 .201 .160 .168 .213	Mear 2∩ .217 .245 .277 .278 .280 .306 .234 .261 .294	n Recip <u>3</u> ∩ .274 .314 .370 .367 .372 .409 .275 .369 .435	rocal R 2∪ .209 .210 .204 .194 .202 .253 .332 .331 .322	ank ∩t .085 .115 .151 .143 .150 .269 .094 .092 .120	$t \cap$.127 .152 .197 .191 .198 .353 .162 .134 .173	$\begin{array}{c} \cup t \\ .145 \\ .202 \\ .257 \\ .256 \\ .260 \\ .220 \\ .125 \\ .147 \\ .187 \end{array}$	Avg .197 .228 .262 .258 .263 .303 .223 .236 .268
Dataset FB15k-237	Variants PERM-1t translations single average MLP (final) 1t translations single average	It .410 .410 .406 .396 .405 .445 .432 .428 .425 .415	2t .180 .232 .283 .283 .286 .268 .191 .197 .241 .241	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210	Mean 2∩ .217 .245 .277 .278 .280 .306 .234 .261 .294 .295	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431	rocal R 2∪ .209 .210 .204 .194 .202 .253 .332 .331 .322 .307	ank ∩t .085 .115 .151 .143 .150 .269 .094 .092 .120 .113	$\begin{array}{c} t \cap \\ .127 \\ .152 \\ .197 \\ .191 \\ .198 \\ .353 \\ .162 \\ .134 \\ .173 \\ .169 \end{array}$	Ut .145 .202 .257 .256 .260 .220 .125 .147 .187 .186	Avg .197 .228 .262 .258 .263 .303 .223 .236 .268 .268 .263
Dataset FB15k-237	Variants PERM-1t translations single average MLP (final) 1t translations single average MLP	It .410 .410 .406 .396 .405 .445 .432 .428 .425 .415 .424	2t .180 .232 .283 .283 .286 .268 .191 .197 .241 .241 .243	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210 .213	Mean 2∩ .217 .245 .277 .278 .280 .306 .234 .261 .294 .295 .298	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431 .436	rocal R 2∪ .209 .210 .204 .194 .202 .253 .332 .331 .322 .307 .319	ank ∩t .085 .115 .151 .143 .150 .269 .094 .092 .120 .113 .119	$\begin{array}{c} t \cap \\ .127 \\ .152 \\ .197 \\ .191 \\ .198 \\ .353 \\ .162 \\ .134 \\ .173 \\ .169 \\ .174 \end{array}$	Ut .145 .202 .257 .256 .260 .220 .125 .147 .187 .186 .189	Avg .197 .228 .262 .258 .263 .303 .223 .236 .268 .268 .263 .268
Dataset FB15k-237 NELL995	Variants PERM-1t translations single average MLP (final) 1t translations single average MLP (final)	It .410 .406 .396 .405 .445 .428 .425 .415 .424 .432	2t .180 .232 .283 .283 .286 .268 .191 .197 .241 .241 .243 .244	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210 .213 .217	Mean 2∩ .217 .245 .277 .278 .280 .306 .234 .261 .294 .295 .298 .296	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431 .436 .438	Procal R 2∪ .209 .210 .204 .194 .202 .332 .331 .322 .307 .319 .332	ank ∩t .085 .115 .151 .143 .150 .269 .094 .092 .120 .113 .119 .122	t∩ .127 .152 .197 .191 .198 .353 .162 .134 .173 .169 .174 .178	Ut .145 .202 .257 .256 .260 .220 .125 .147 .187 .186 .189 .190	Avg .197 .228 .262 .258 .263 .303 .223 .236 .268 .268 .263 .268 .272
Dataset FB15k-237	VariantsPERM-1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1t	It .410 .406 .396 .405 .445 .425 .425 .415 .422 .425 .425 .425 .425 .425 .425 .425 .425	2t .180 .232 .283 .283 .286 .268 .191 .197 .241 .241 .243 .244 .005	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210 .213 .217 N.A.	Mean 2∩ .217 .245 .277 .278 .280 .306 .234 .261 .294 .295 .298 .296	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431 .436 .438 .541	Procal R 2∪ .209 .210 .204 .194 .202 .332 .331 .322 .307 .319 .332 .564	ank ∩t .085 .115 .151 .143 .150 .269 .094 .092 .120 .113 .119 .122 .169	<i>t</i> ∩ .127 .152 .197 .191 .198 .353 .162 .134 .173 .169 .174 .178 .791	Ut .145 .202 .257 .256 .260 .220 .125 .147 .187 .186 .189 .190 0.00	Avg .197 .228 .262 .258 .263 .303 .223 .236 .268 .263 .268 .263 .268 .272 .430
Dataset FB15k-237 NELL995	VariantsPERM-1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationsingleaverageMLP(final)1ttranslations	It .410 .406 .396 .405 .445 .428 .425 .415 .424 .432 .706 .700	2t .180 .232 .283 .283 .286 .268 .191 .197 .241 .241 .243 .244 .005 .005	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210 .213 .217 N.A. N.A.	Mean 2∩ .217 .245 .277 .278 .280 .306 .234 .261 .294 .295 .298 .296 .665 .741	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431 .436 .438 .541 .727	Procal R 2∪ .209 .210 .204 .194 .202 .332 .331 .322 .307 .319 .332 .564 .562	ank ∩t .085 .115 .151 .143 .150 .269 .094 .092 .120 .113 .119 .122 .169 .164	$t \cap$.127 .152 .197 .191 .198 .353 .162 .134 .173 .169 .174 .178 .791 .655	Ut .145 .202 .257 .256 .260 .220 .125 .147 .187 .186 .189 .190 0.00 0.00	Avg .197 .228 .262 .258 .263 .303 .223 .236 .268 .268 .263 .268 .272 .430 .444
Dataset FB15k-237 NELL995	VariantsPERM-1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingle	It .410 .406 .396 .405 .445 .425 .425 .415 .422 .706 .700 .694	2t .180 .232 .283 .283 .286 .268 .191 .197 .241 .241 .243 .244 .005 .005 .006	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210 .213 .217 N.A. N.A. N.A. N.A.	Mean 2∩ .217 .245 .277 .278 .280 .306 .234 .261 .294 .295 .298 .296 .665 .741 .841	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431 .436 .438 .541 .727 .862	Procal R 2∪ .209 .210 .204 .194 .202 .332 .331 .322 .307 .319 .332 .564 .562 .547	ank ∩t .085 .115 .151 .143 .150 .269 .094 .092 .120 .113 .119 .122 .169 .164 .215	<i>t</i> ∩ .127 .152 .197 .191 .198 .353 .162 .134 .173 .169 .174 .178 .791 .655 .847	Ut .145 .202 .257 .256 .260 .220 .125 .147 .187 .186 .189 .190 0.00 0.00 0.00 0.00	Avg .197 .228 .262 .258 .263 .303 .223 .236 .268 .263 .268 .272 .430 .444 .502
Dataset FB15k-237 NELL995	VariantsPERM-1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverage	It .410 .406 .396 .405 .445 .428 .425 .415 .424 .432 .706 .700 .694 .678	2t .180 .232 .283 .283 .286 .268 .191 .197 .241 .241 .243 .244 .005 .005 .006 .006	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210 .213 .217 N.A. N.A. N.A. N.A. N.A.	Mean 2∩ .217 .245 .277 .278 .280 .306 .234 .261 .294 .295 .298 .296 .665 .741 .838	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431 .436 .438 .541 .727 .862 .848	Procal R 2∪ .209 .210 .204 .194 .202 .253 .332 .331 .322 .307 .319 .332 .564 .562 .547 .521	ank ∩t .085 .115 .151 .143 .150 .269 .094 .092 .120 .113 .119 .122 .169 .164 .215 .204	<i>t</i> ∩ .127 .152 .197 .191 .198 .353 .162 .134 .173 .169 .174 .178 .791 .655 .847 .824	Ut .145 .202 .257 .256 .260 .220 .125 .147 .187 .186 .189 .190 0.00 0.00 0.00 0.00 0.00	Avg .197 .228 .262 .258 .263 .303 .223 .236 .268 .263 .268 .272 .430 .444 .502 .490
Dataset FB15k-237 NELL995	Variants PERM-1t translations single average MLP (final) 1t translations single average MLP MLP	It .410 .406 .396 .405 .445 .425 .425 .415 .422 .706 .700 .694 .678 .693	2t .180 .232 .283 .283 .286 .268 .191 .197 .241 .241 .243 .244 .005 .005 .006 .006	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210 .213 .217 N.A. N.A. N.A. N.A. N.A. N.A. N.A.	Mean 2∩ .217 .245 .277 .278 .280 .306 .234 .261 .294 .295 .298 .296 .665 .741 .838 .838	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431 .436 .438 .541 .727 .862 .848 .855	Procal R 2∪ .209 .210 .204 .194 .202 .253 .332 .331 .322 .307 .319 .332 .564 .562 .547 .521 .542	ank ∩t .085 .115 .151 .143 .150 .269 .094 .092 .120 .113 .119 .122 .169 .164 .215 .204 .214	$t \cap$.127 .152 .197 .191 .198 .353 .162 .134 .173 .169 .174 .178 .791 .655 .847 .824 .851	Ut .145 .202 .257 .256 .260 .220 .125 .147 .187 .186 .189 .190 0.00 0.00 0.00 0.00 0.00 0.00	Avg .197 .228 .262 .258 .263 .303 .223 .236 .268 .263 .268 .272 .430 .444 .502 .490 .500
Dataset FB15k-237 NELL995 DBPedia	VariantsPERM-1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)	It .410 .406 .396 .405 .445 .428 .425 .415 .424 .432 .706 .700 .694 .678 .693 .706	2t .180 .232 .283 .283 .286 .268 .191 .197 .241 .241 .243 .244 .005 .005 .006 .006 .006 .006	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210 .213 .217 N.A. N.A. N.A. N.A. N.A. N.A. N.A. N.A	Mean 2∩ .217 .245 .277 .278 .280 .306 .234 .261 .294 .295 .298 .296 .665 .741 .838 .838 .841	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431 .436 .438 .541 .727 .862 .848 .855 .862	Procal R 2∪ .209 .210 .204 .194 .202 .253 .332 .331 .322 .307 .319 .332 .564 .547 .521 .542 .564	ank ∩t .085 .115 .151 .153 .150 .269 .094 .092 .120 .113 .119 .122 .169 .164 .215 .204 .214 .219	$t \cap$.127 .152 .197 .191 .198 .353 .162 .134 .173 .169 .174 .178 .791 .655 .847 .824 .851 .869	$ \begin{array}{c} \cup t \\ .145 \\ .202 \\ .257 \\ .256 \\ .260 \\ .220 \\ .125 \\ .147 \\ .187 \\ .186 \\ .189 \\ .190 \\ 0.00 \\ 0.$	Avg 197 228 262 258 263 303 223 236 268 263 268 272 430 444 502 490 500 452
Dataset FB15k-237 NELL995	VariantsPERM-1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1t	It .410 .406 .396 .405 .445 .428 .425 .415 .424 .432 .706 .700 .694 .678 .693 .706 .416	2t .180 .232 .283 .283 .286 .268 .191 .197 .241 .241 .243 .244 .005 .005 .006 .006 .006 .006 .006 .006	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210 .213 .217 N.A. N.A. N.A. N.A. N.A. N.A. N.A. N.A	Mean 2∩ .217 .245 .277 .280 .306 .234 .261 .294 .295 .298 .296 .665 .741 .838 .841 .254	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431 .436 .438 .541 .727 .862 .848 .855 .862 .341	Procal R 2∪ .209 .210 .204 .194 .202 .253 .332 .331 .322 .307 .319 .332 .564 .562 .547 .521 .542 .564 .269	ank ∩t .085 .115 .151 .143 .150 .269 .094 .092 .120 .113 .119 .122 .169 .164 .215 .204 .214 .219 .100	$t \cap$.127 .152 .197 .191 .198 .353 .162 .134 .173 .169 .174 .178 .791 .655 .847 .824 .851 .869 .157	$ \begin{array}{c} \cup t \\ .145 \\ .202 \\ .257 \\ .256 \\ .260 \\ .220 \\ .125 \\ .147 \\ .187 \\ .186 \\ .189 \\ .190 \\ 0.00 \\ 0.$	Avg 197 228 262 258 263 303 223 236 268 263 268 272 430 444 502 490 500 452 220
Dataset FB15k-237 NELL995 DBPedia	VariantsPERM-1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslations1ttranslations	It .410 .406 .396 .405 .445 .428 .425 .415 .424 .432 .706 .700 .694 .678 .693 .706 .416	2t .180 .232 .283 .283 .286 .268 .191 .197 .241 .241 .243 .244 .005 .005 .006 .006 .006 .006 .006 .006	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210 .213 .210 .213 .217 N.A. N.A. N.A. N.A. N.A. N.A. N.A. N.A	Mean 2∩ .217 .245 .277 .280 .306 .234 .261 .294 .295 .298 .296 .665 .741 .838 .841 .254 .286	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431 .436 .438 .541 .727 .862 .848 .855 .862 .341 .392	Procal R 2∪ .209 .210 .204 .194 .202 .253 .332 .331 .322 .307 .319 .332 .564 .562 .547 .521 .542 .564 .269 .270	ank ∩t .085 .115 .151 .151 .150 .269 .094 .092 .120 .113 .119 .122 .169 .164 .215 .204 .214 .219 .100 .135	$t \cap$.127 .152 .197 .191 .198 .353 .162 .134 .173 .169 .174 .178 .791 .655 .847 .824 .851 .869 .157 .188	$ \begin{array}{c} \cup t \\ .145 \\ .202 \\ .257 \\ .256 \\ .260 \\ .220 \\ .125 \\ .147 \\ .187 \\ .186 \\ .189 \\ .190 \\ 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \\ 0.00 \\ 0.157 \\ .218 \end{array} $	Avg 197 228 262 258 263 203 223 236 268 263 268 272 268 272 430 444 502 490 500 452 220
Dataset FB15k-237 NELL995 DBPedia	VariantsPERM-1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingle	It .410 .406 .396 .405 .445 .428 .425 .415 .424 .432 .706 .700 .694 .678 .693 .706 .416 .413	2t .180 .232 .283 .283 .286 .268 .191 .197 .241 .241 .243 .244 .005 .005 .006 .006 .006 .006 .006 .006	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210 .213 .210 .213 .217 N.A. N.A. N.A. N.A. N.A. N.A. N.A. N.A	Mean 2∩ .217 .245 .277 .280 .306 .234 .261 .294 .295 .298 .296 .665 .741 .838 .841 .254 .286 .323	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431 .436 .438 .541 .727 .862 .848 .855 .862 .341 .392 .462	Procal R 2∪ .209 .210 .204 .194 .202 .253 .332 .331 .322 .307 .319 .332 .564 .562 .547 .521 .542 .564 .269 .270 .263	ank $\cap t$.085 .115 .151 .153 .150 .269 .094 .092 .120 .113 .119 .122 .169 .164 .215 .204 .214 .219 .100 .135 .176	$t \cap$.127 .152 .197 .191 .198 .353 .162 .134 .173 .169 .174 .178 .791 .655 .847 .824 .851 .869 .157 .188 .243	Ut .145 .202 .257 .256 .260 .220 .125 .147 .187 .186 .189 .190 0.00 0.00 0.00 0.00 0.00 0.00 0.0	Avg .197 .228 .262 .258 .263 .233 .236 .268 .263 .268 .272 .430 .444 .502 .440 .500 .452 .220 .255 .293
Dataset FB15k-237 NELL995 DBPedia	VariantsPERM-1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslations1ttranslations	It .410 .406 .396 .405 .445 .428 .425 .415 .424 .432 .706 .700 .694 .678 .693 .706 .416 .413 .404	2t .180 .232 .283 .286 .268 .191 .197 .241 .241 .243 .244 .005 .005 .006 .006 .006 .006 .006 .006	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210 .213 .217 N.A. N.A. N.A. N.A. N.A. N.A. N.A. N.A	Mean 2∩ .217 .245 .277 .280 .306 .234 .294 .295 .298 .296 .665 .741 .838 .841 .254 .223 .323 .324	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431 .436 .438 .541 .727 .862 .848 .855 .862 .341 .392	Procal R 2∪ .209 .210 .204 .194 .202 .253 .332 .331 .322 .307 .319 .332 .564 .562 .547 .521 .564 .269 .270 .263 .250	ank ∩t .085 .115 .151 .151 .150 .269 .094 .092 .120 .113 .119 .122 .169 .164 .215 .204 .214 .219 .100 .135	$t \cap$.127 .152 .197 .191 .198 .353 .162 .134 .173 .169 .174 .178 .791 .655 .847 .824 .851 .869 .157 .188 .243 .236	$ \begin{array}{c} \cup t \\ .145 \\ .202 \\ .257 \\ .256 \\ .260 \\ .220 \\ .125 \\ .147 \\ .187 \\ .186 \\ .189 \\ .190 \\ 0.00 \\ 0.$	Avg .197 .228 .262 .258 .263 .233 .236 .268 .263 .268 .263 .268 .272 .430 .444 .502 .440 .500 .452 .220 .255 .293 .288
Dataset FB15k-237 NELL995 DBPedia	VariantsPERM-1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingleaverageMLP(final)1ttranslationssingle	It .410 .406 .396 .405 .445 .428 .425 .415 .424 .432 .706 .700 .694 .678 .693 .706 .416 .413	2t .180 .232 .283 .283 .286 .268 .191 .197 .241 .241 .243 .244 .005 .005 .006 .006 .006 .006 .006 .006	3t .122 .171 .217 .214 .217 .201 .160 .168 .213 .210 .213 .210 .213 .217 N.A. N.A. N.A. N.A. N.A. N.A. N.A. N.A	Mean 2∩ .217 .245 .277 .280 .306 .234 .261 .294 .295 .298 .296 .665 .741 .838 .841 .254 .286 .323	n Recip 3∩ .274 .314 .370 .367 .372 .409 .275 .369 .435 .431 .436 .438 .541 .727 .862 .848 .855 .862 .341 .392 .462	Procal R 2∪ .209 .210 .204 .194 .202 .253 .332 .331 .322 .307 .319 .332 .564 .562 .547 .521 .542 .564 .269 .270 .263	ank $\cap t$.085 .115 .151 .153 .150 .269 .094 .092 .120 .113 .119 .122 .169 .164 .215 .204 .214 .219 .100 .135 .176	$t \cap$.127 .152 .197 .191 .198 .353 .162 .134 .173 .169 .174 .178 .791 .655 .847 .824 .851 .869 .157 .188 .243	Ut .145 .202 .257 .256 .260 .220 .125 .147 .187 .186 .189 .190 0.00 0.00 0.00 0.00 0.00 0.00 0.0	Avg .197 .228 .262 .258 .263 .233 .236 .268 .263 .268 .272 .430 .444 .502 .440 .500 .452 .220 .255 .293

Table 7: Performance comparison of (final) PERM model against its variants to study the contributions of its components. The columns present the query structures and the overall average performance.