

SCONE: A Novel Stochastic Sampling to Generate Contrastive Views and Hard Negative Samples for Recommendation



Chaejeong Lee^{1*}, Jeongwhan Choi^{2*}, Hyowon Wi³, Sung-Bae Cho², Noseong Park³

¹Korea Telecom ²Yonsei University ³Korea Advanced Institute of Science & Technology

Killing Two Recommendation Challenges with One Stochastic Stone!

A unified framework, **SCONE**, improves data sparsity and negative sampling.



Contrastive Learning and Negative Sampling for Recommendation

Table: Existing Methods vs. SCONE

Model	Contrastive Views	Negative Samples
LightGCN	$\times$	Random sampling
SGL	Perturbing graph structure	Random sampling
SimGCL	Injecting noises into embeddings	Random sampling
DNS	$\times$	Promising candidate
MixGCF	$\times$	Linear interpolation
SCONE	Stochastic sampling with SGMs	

Main Idea: SCONE Architecture Overview

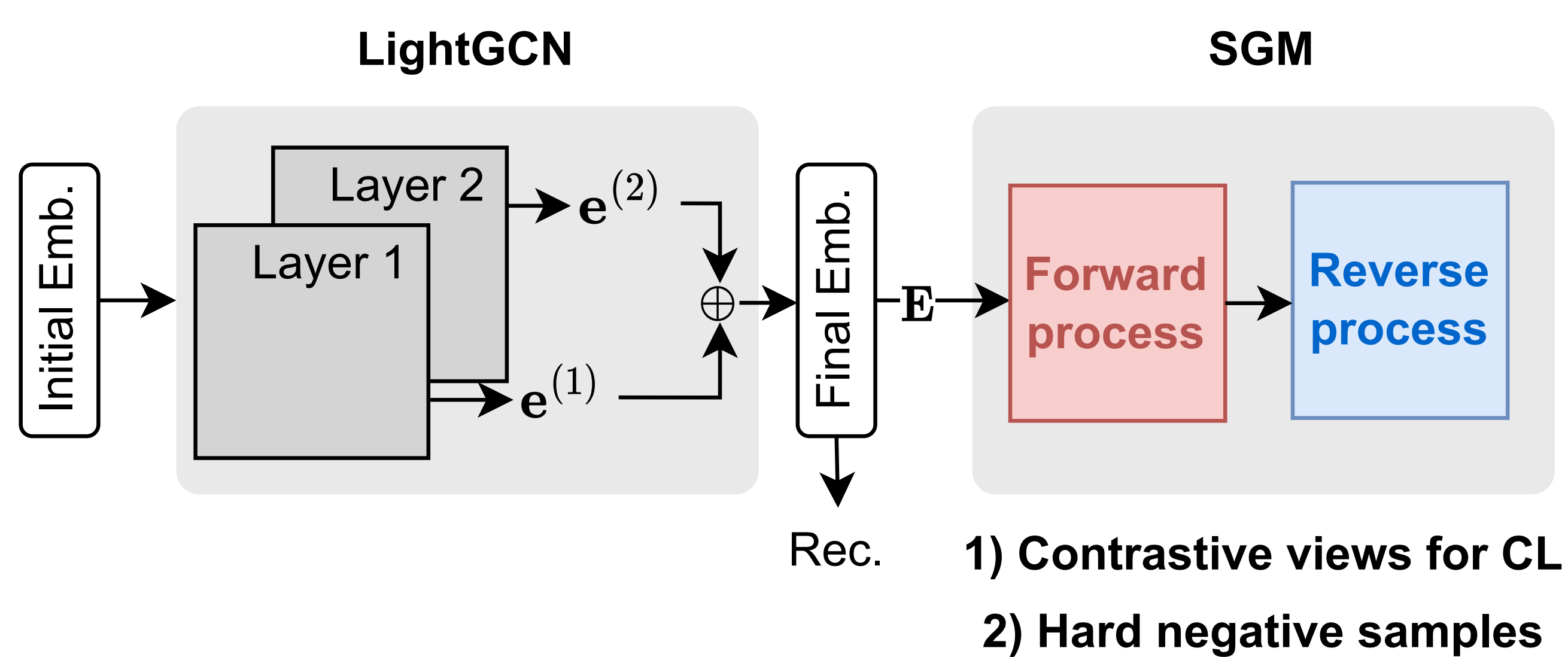


Figure: The overall model architecture consisting of LightGCN and SGM.

Score-based Generative Models (SGM)

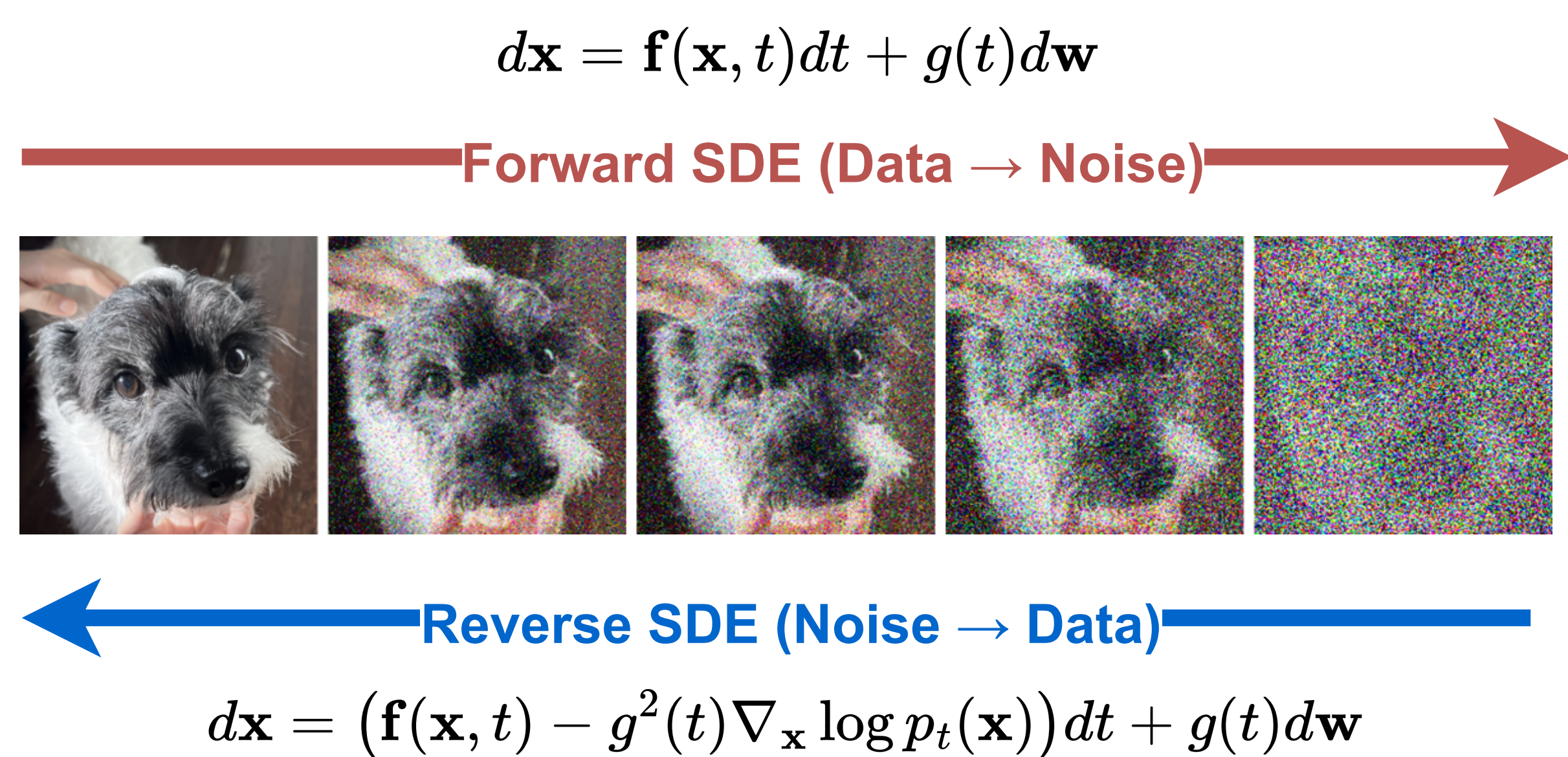
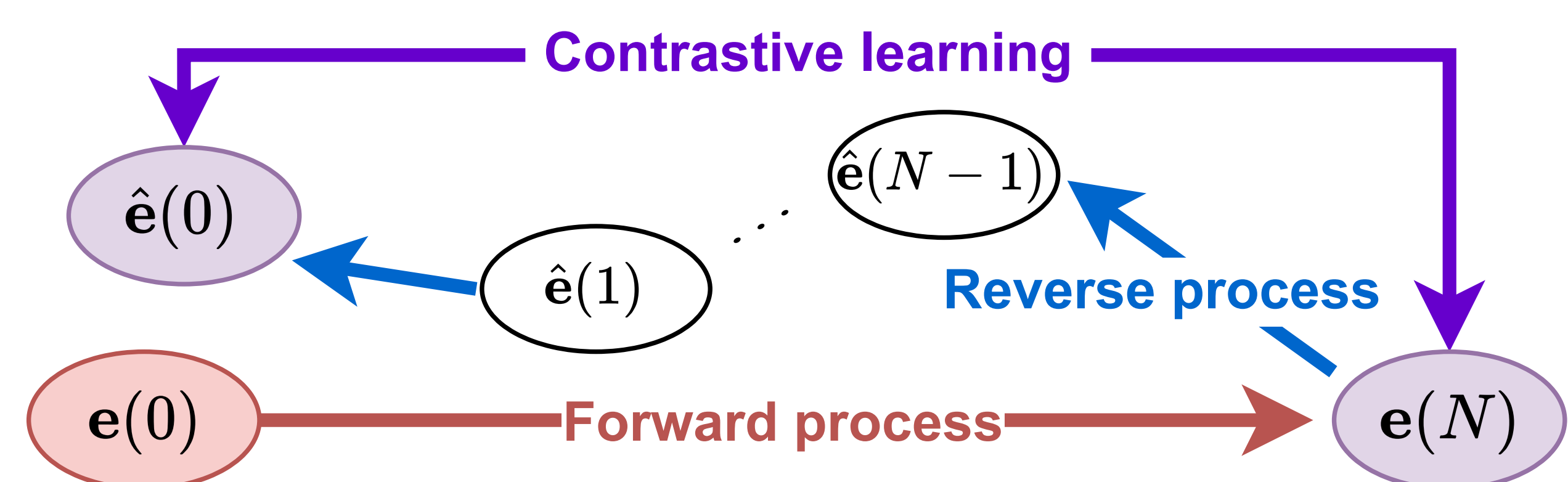


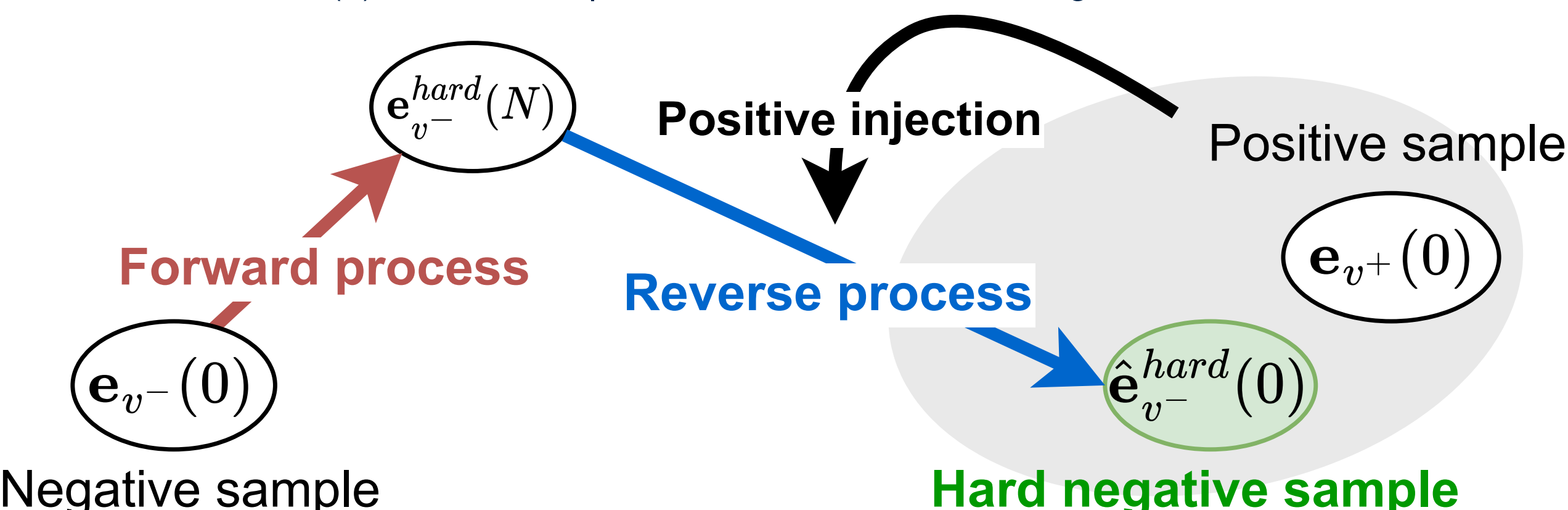
Figure: The overall framework of score-based generative models. SGMs corrupt the input data with increasing noise and generate a new sample from the noisy data.

SCONE: Our Unified Approach

Stochastic sampling for **C**ontrastive views and hard **N**egative samples (**SCONE**)



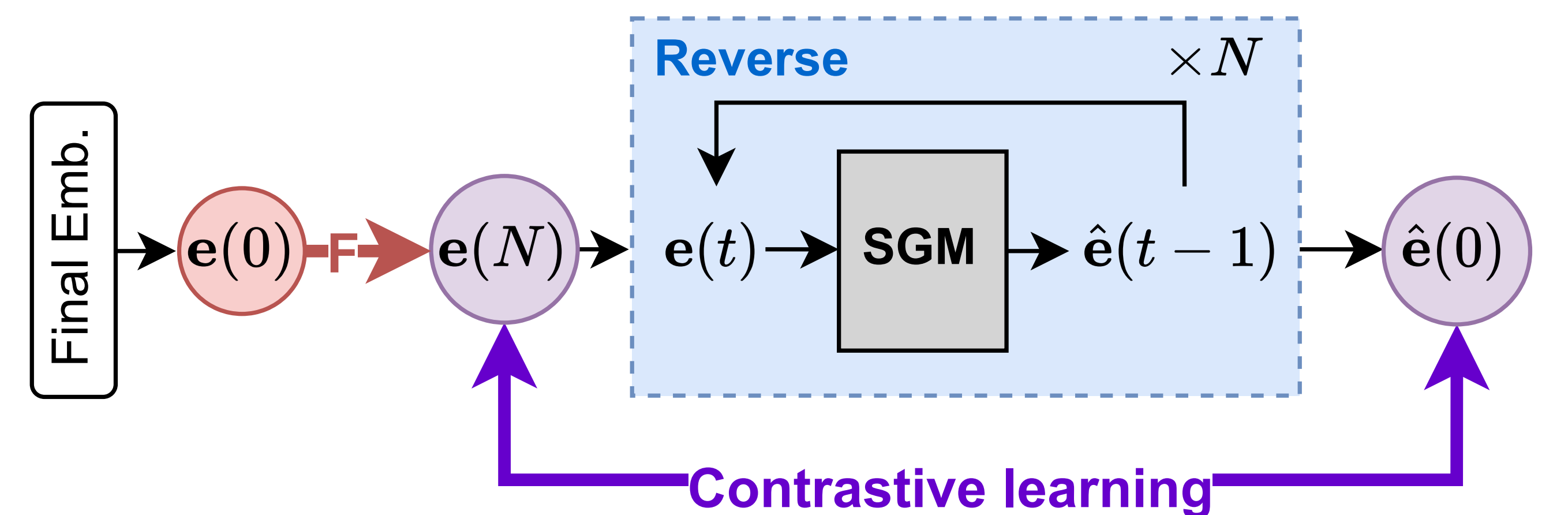
(a) A stochastic process for contrastive views generation



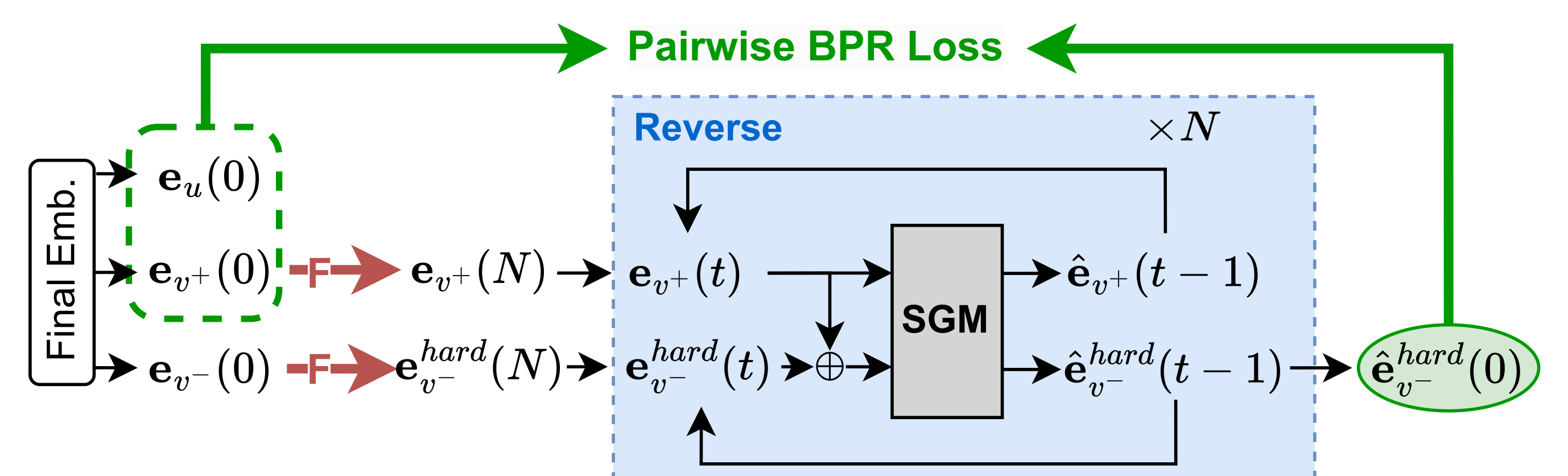
(b) A stochastic positive injection for hard negative samples generation

Figure: Our proposed user-item specific stochastic sampling method based on score-based generative models for contrastive learning and negative sampling.

Contrastive Views Generation & Hard Negative Sample Generation



(a) Contrastive views generation for a user/item embedding $e(0)$ via a stochastic process.



(b) Hard negative sample generation via a stochastic positive injection given a user-item embedding triplet $(e_u(0), e_{v+}(0), e_{v-}(0))$. The synthesized sample $\hat{e}_{v-}^{hard}(0)$ is a hard negative sample.

Figure: The overall workflow of SCONE

Performance Comparison

Table: Performance comparison with state-of-the-art models.

Dataset	Metric	GCN		CL		NS		SCONE
		LightGCN	SGL	NCL	SimGCL	DNS	MixGCF	
Douban	Recall@20	0.1474	0.1728	0.1638	0.1780	0.1645	0.1784	0.1815
	NDCG@20	0.1240	0.1510	0.1403	0.1567	0.1393	0.1576	0.1611
Gowalla	Recall@20	0.2086	0.2223	0.2235	0.2269	0.2233	0.2147	0.2295
	NDCG@20	0.1264	0.1349	0.1353	0.1386	0.1358	0.1295	0.1409
Tmall	Recall@20	0.0686	0.0754	0.0700	0.0898	0.0860	0.0819	0.0908
	NDCG@20	0.0477	0.0532	0.0492	0.0643	0.0600	0.0570	0.0654
Yelp2018	Recall@20	0.0588	0.0669	0.0654	0.0718	0.0671	0.0703	0.0721
	NDCG@20	0.0485	0.0552	0.0540	0.0593	0.0551	0.0576	0.0594
Amazon-CDs	Recall@20	0.1283	0.1550	0.1483	0.1584	0.1545	0.1542	0.1586
	NDCG@20	0.0779	0.0986	0.0940	0.1006	0.0972	0.0960	0.1009
ML-1M	Recall@20	0.2693	0.2657	0.2777	0.2868	0.2791	0.2897	0.2910
	NDCG@20	0.3015	0.3027	0.3150	0.3230	0.3171	0.3287	0.3244

Robustness of SCONE

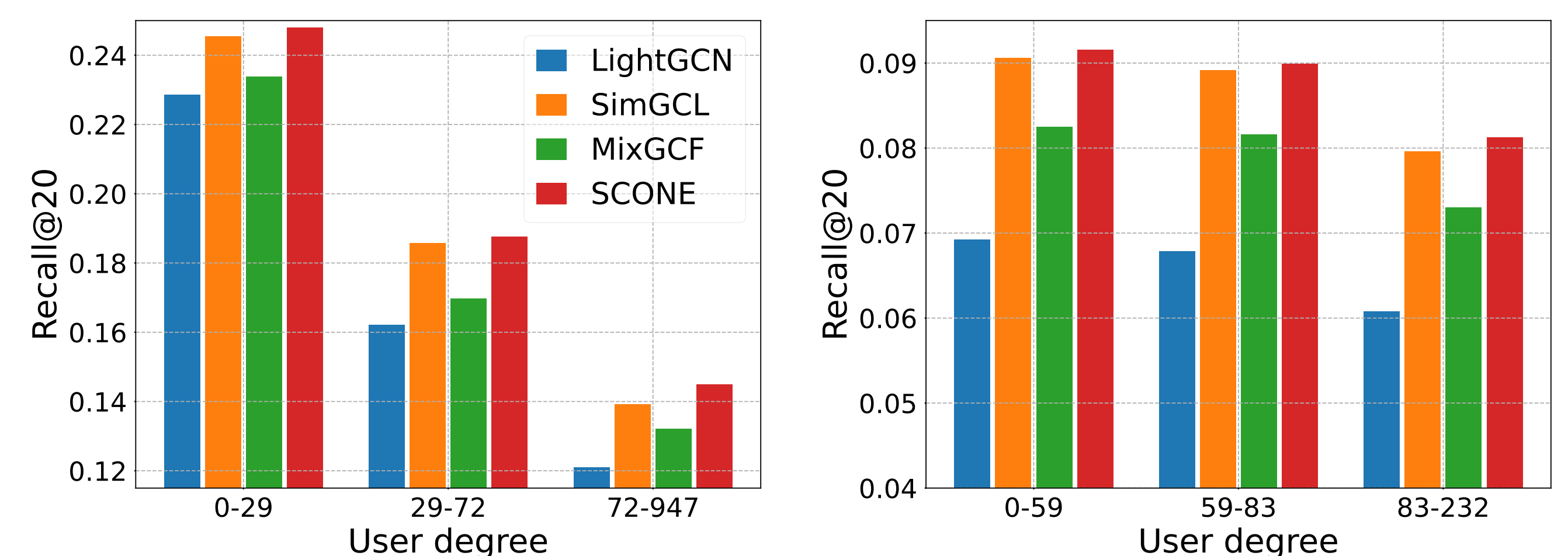


Figure: Performance comparison over different user groups

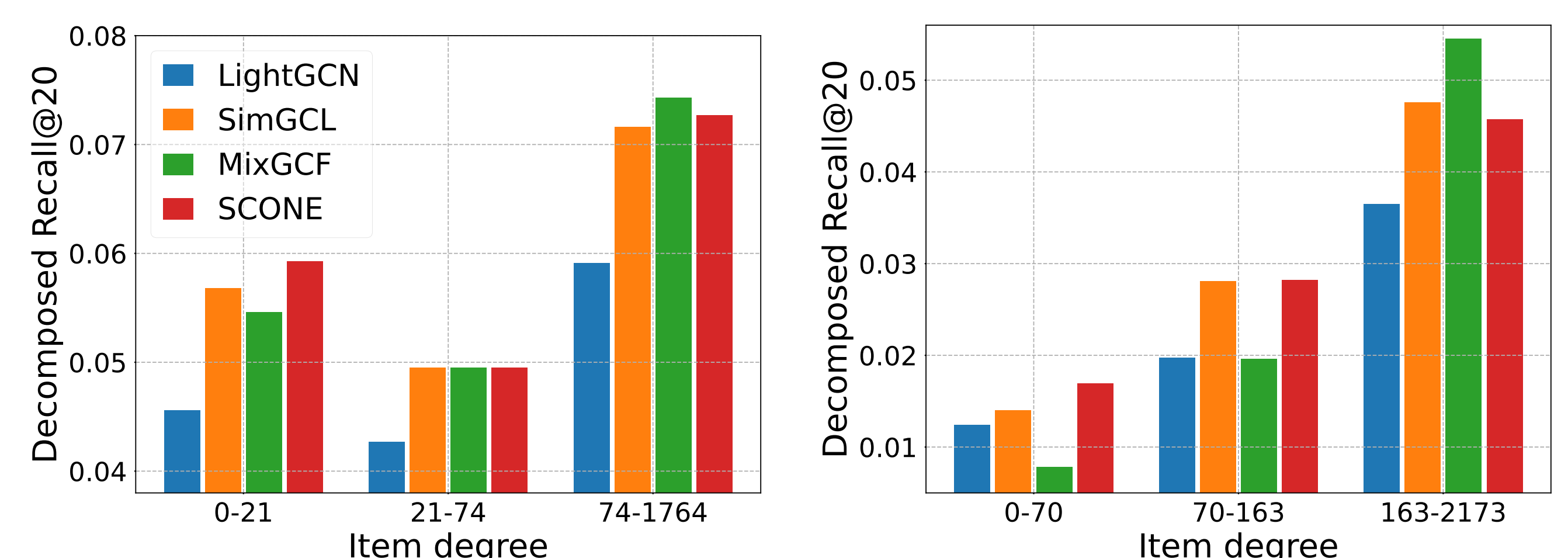


Figure: Performance comparison over different item groups.