APPENDIX

I. IMPLEMENTATION DETAILS

A. Rewards

Table I lists both the contact reward terms and regularization terms together with the weights used in all experiments.

TABLE I
REWARD COMPONENTS

Term	Expression	Weight
Contact		
Correct strikes	$\sum_{e \in \mathcal{E}_t} \mathbb{I}\{\text{drum}(e) \in D_i\}$	$+1.0$
Wrong strikes	$\sum_{e \in \mathcal{E}_t} \mathbb{I}\{\text{drum}(e) \notin D_i\}$	-0.5
Missed strikes	$\sum_{d \in D_i} \mathbb{I}\{\text{drum missed} = d\}$	-2.0
Proximity	$\mathbb{I}\{|D_i| > 0\} \sum_{s \in \{s_L, s_R\}} min_{d \in D_i} \|p_{s,t} - p_d\|$	-1.0
Regularization		
Action-rate	$\|a_t - a_{t-1}\|^2$	$-1e^{-3}$
DoF acceleration	$\|\ddot{q}\|^2$	$-2.5e^{-7}$

Note: $\mathcal{E}_t$ is the set of drums the robot actually hits at time t; $p_{s,t}$ is the position of stick s at time t; p_d is the position of drum d.

B. PD Controller

Table II lists the stiffness (K_p) and damping (K_d) values used in the PD controller for all joints.

TABLE II
PD CONTROLLER PARAMETERS

Joint	K_p (N·m/rad)	K_d (N·m·s/rad)
Waist	100	2
Shoulder pitch	90	2
Shoulder roll	60	1
Shoulder yaw	20	0.4
Elbow	60	1
Wrist roll	20	0.1
Wrist pitch	4	0.2
Wrist yaw	4	0.2

C. Policy Training and Evaluation Protocol

We use a 3-layer MLP architecture for both the actor and critic networks, with hidden sizes of (128, 64, 32). For each song or training condition, policies are learned across 5 random seeds over 1500 PPO updates. Experiments are conducted on a server with Nvidia RTX 4090 GPUs, running 8192 parallel environments in Isaac Gym. After training, each policy is evaluated over 20 independent rollouts, and performance is measured using the mean F1 score between actual and target drum strikes across all timesteps. The final performance metric is the average F1 score and standard deviation across the 5 seeds.

II. ADDITIONAL RESULTS

A. Song-level Performance Analysis

Figure 1 presents song-level performance analysis across all specialist policies. (a) shows absolute Spearman rank correlations between individual song features and per-song F_1 score, while (b–d) show the relationship between F_1 and selected features using scatter plots with linear trendlines. Together, these plots summarize how performance varies across songs as a function of rhythmic and spatial attributes of the target drum patterns.

B. Temporal Decomposition Ablation

Table III reports per-song F_1 scores (mean $\pm$ standard deviation over five seeds) obtained with and without temporal decomposition. Temporal decomposition trains fixed-length segments concurrently rather than end-to-end on full tracks. Although the policies achieve similar F_1 scores in both cases, wall-clock time is substantially reduced for each run with temporal decomposition (from 8–9 hours to 2–3 hours).

C. Listener Study

The listener study evaluates perceptual aspects of the drumming pattens performed by Robot Drummer using subjective ratings collected from human participants. Fifteen participants each watched and listened to three performance videos of specialist policies. After watching the videos, participants answered a series of questions, rating the robot's drumming on a 1–5 Likert scale (1 = not at all; 5 = very much) for the following aspects:

- Timing Consistency - How consistent and well-timed was the rhythm of the drumming?
- Expressiveness - To what extent did the robot convey a sense of musical feel?
- Naturalness - How natural or human-like did the robot's drumming sounded?
- Enjoyment - Overall enjoyment of the performances.

Across all responses, we observed the following results:

- Timing Consistency: 3.69 ± 0.63
- Expressiveness: 3.38 ± 0.96
- Naturalness: 2.46 ± 0.78
- Enjoyment: 3.08 ± 1.60

These results show that listeners generally perceived the robot's timing and musical feel in the upper mid-range, with lower ratings for human-like naturalness and mixed enjoyment responses reflected by greater variance.

TABLE III
F1 SCORES WITH AND WITHOUT TEMPORAL DECOMPOSITION

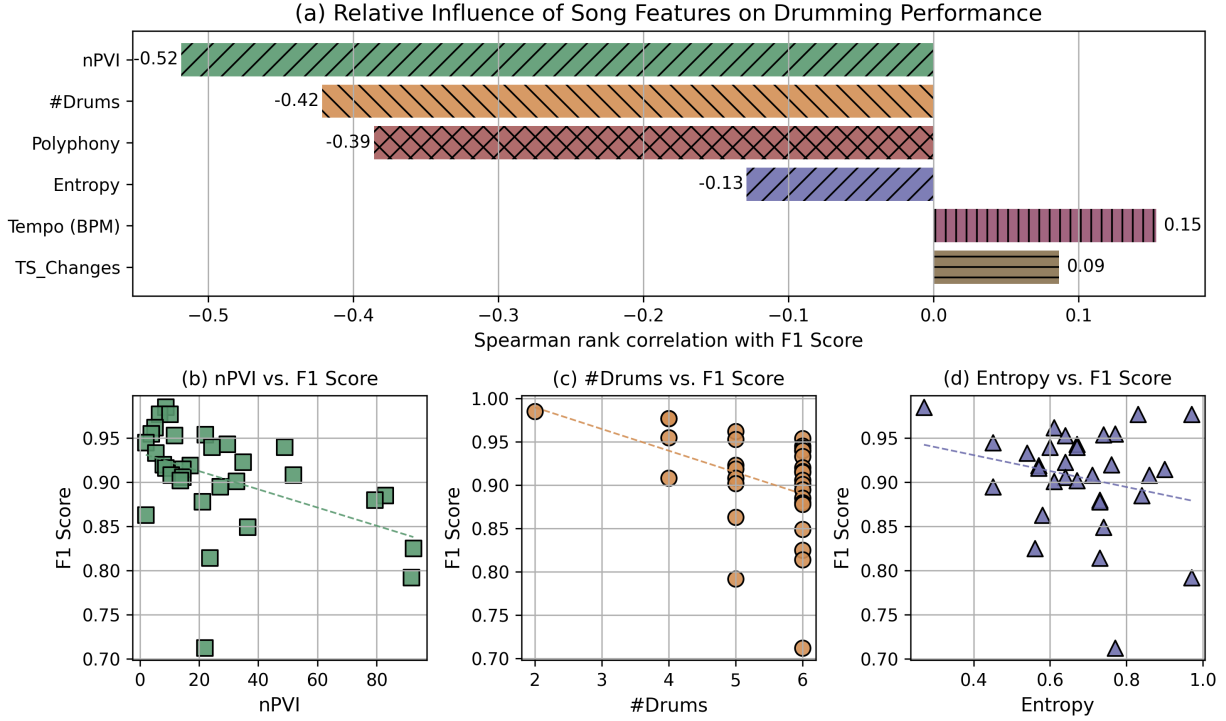| Song	w/o Decomposition	w/ Decomposition
Rebel Rebel	0.98 ± 0.01	0.99 ± 0.00
Lithium	0.96 ± 0.02	0.95 ± 0.02
Fire	0.93 ± 0.02	0.94 ± 0.02
In The End	0.92 ± 0.02	0.90 ± 0.02
Livin' on a Prayer	0.86 ± 0.02	0.89 ± 0.03
Roxanne	0.89 ± 0.02	0.88 ± 0.02

Fig. 1. Song-level performance analysis across specialist policies. (a) Absolute Spearman correlations between song features and per-song F$_1$. (b–d) Scatter plots showing F$_1$ versus nPVI, number of drums, and entropy.

TABLE IV
QUANTITATIVE EVALUATION OF POLICIES ON SONGS WITH NON-ZERO POLYPHONY

Song	#Drums	Entropy	nPVI	BPM	Polyphony (%)	F$_1$ ↑
Limp Bizkit - Break Stuff	4	0.83	10.22	107	0.27	0.977 ± 0.005
Metallica - Nothing Else Matters	5	0.61	5.08	69	0.23	0.962 ± 0.003
Foo Fighters - Everlong	4	0.77	3.79	156	0.86	0.955 ± 0.017
Metallica - One	5	0.64	11.74	142	1.21	0.953 ± 0.006
AC/DC - You Shook Me All Night Long	6	0.45	2.05	129	1.71	0.945 ± 0.017
Europe - The Final Countdown	6	0.60	48.96	118	1.56	0.940 ± 0.011
Rage Against the Machine - Killing in the Name	6	0.67	24.38	102	0.14	0.940 ± 0.006
Eagles - Hotel California	6	0.54	5.38	149	0.98	0.933 ± 0.020
Iron Maiden - Where Eagles Dare	5	0.64	34.82	120	1.16	0.923 ± 0.016
Green Day - Boulevard of Broken Dreams	6	0.76	7.96	167	0.20	0.920 ± 0.009
Nirvana - The Man Who Sold the World	5	0.57	17.08	118	0.40	0.919 ± 0.017
Nirvana - Come as You Are	6	0.57	8.73	120	0.71	0.916 ± 0.013
Nirvana - Dumb	6	0.90	14.59	120	1.60	0.915 ± 0.050
Metallica - Enter Sandman	5	0.86	10.63	123	0.55	0.908 ± 0.025
The Police - Message in a Bottle	6	0.64	14.55	150	0.21	0.906 ± 0.021
Guns N' Roses - Sweet Child o' Mine	5	0.67	13.66	128	0.21	0.902 ± 0.040
AC/DC - Thunderstruck	6	0.45	27.17	107	3.36	0.895 ± 0.013
Guns N' Roses - Knockin' on Heaven's Door	6	0.73	79.42	128	0.31	0.880 ± 0.023
Iron Maiden - Wasted Years	5	0.58	2.02	152	4.58	0.863 ± 0.013
Bon Jovi - Bed of Roses	6	0.74	36.49	60	0.64	0.849 ± 0.011
Bon Jovi - It's My Life	6	0.56	92.40	120	1.69	0.825 ± 0.022
Bon Jovi - Always	6	0.73	23.61	67	3.28	0.814 ± 0.020
Rolling Stones - Paint It Black	5	0.97	91.76	110	3.20	0.792 ± 0.055
Nirvana - Smells Like Teen Spirit	6	0.77	22.02	119	14.16	0.712 ± 0.030

All songs have zero time signature changes, except for "Metallica - Nothing Else Matters" (17), and "Nirvana - The Man Who Sold the World" (4).