PG-BIG: Personalized Guidance for Biomechanically Informed Generative Models in Exercise Science

Nicholas King NCKING@UCSD.EDU
Jared Maeyama JMAEYAMA@UCSD.EDU
Shubh Maheshwari SHMAHESHWARI@UCSD.EDU
Andrew McCulloch AMCCULLOCH@UCSD.EDU
Rose Yu ROSEYU@UCSD.EDU
University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093

Editors: G. Sukhatme, L. Lindemann, S. Tu, A. Wierman, N. Atanasov

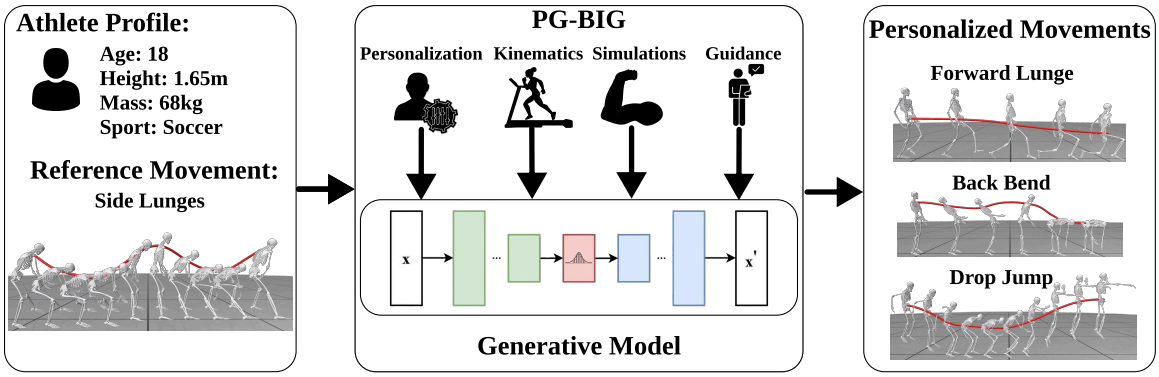

Figure 1: PG-BIG uses subject-specific motion references to learn individualized biomechanical motion priors, enabling personalized and physiologically plausible motion generation.

Abstract

Modeling human motion that is both biomechanically realistic and personalized to individual characteristics remains a key challenge in movement science. While biomechanically informed models such as BIGE Maheshwari (2025) incorporate physiological constraints to produce physically plausible motions, they operate at a population level and fail to capture individual variability in anatomy, strength, and motor strategy. Limiting their applicability to contexts like athletic performance analysis and rehabilitation, where personalization is critical. We introduce PG-BIG, a generative framework that integrates subject-specific personalization with biomechanical guidance in a unified generative pipeline. PG-BIG conditions on both an athlete profile and an action label to generate motion that aligns with individual style while maintaining physiological plausibility. Experiments on the Motus Global movement-screen dataset show that PG-BIG outperforms prior generative baselines in biomechanical realism and stylistic fidelity, enabling interpretable and personalized motion synthesis for applications in performance optimization and injury prevention. Code implementation is available atproject webpage.

Keywords: personalized human motion, generative AI, athlete-specific synthesis, musculoskeletal dynamics, motor control, movement analysis, digital biomechanics

1. Introduction

Accurate modeling of human motion allows for applications in athletic performance, virtual coaching, and rehabilitation. Recent works have been done to generate individualized motion patterns that reflect a person's physiology, performance level, and movement style. Physics-based simulations Keller et al. (2023); Rajagopal et al. (2016a) ensure physiological validity but rely on population-level models and require expert tuning, limiting real-time, individualized use. Generative AI offers a complementary approach, learning directly from large-scale motion data to produce visually realistic movements Min et al. (2010); Tashakori et al. (2025), but often neglects personalization and biomechanical validity.

Athletic performance is inherently personal. Differences in anthropometry, injury history, and training background create unique motion signatures that generic models miss. Simultaneously, biomechanical realism is essential. Motions that violate joint limits or muscle constraints cannot be clinically trusted or used for coaching. Previous work, such as BIGE Maheshwari (2025) for biomechanical guidance and PersonaBooth Kim et al. (2025) for subject-specific personalization, addresses one aspect, but not both.

We introduce PG-BIG, which integrates subject-specific personalization and biomechanical guidance in a unified pipeline. While building on prior work, we achieve this through a synergistic combination: neither personalization alone nor guidance alone achieves both properties. PG-BIG conditions generation on an athlete's profile encoding anthropometric and performance attributes, and an action label specifying the target movement. Generated motions are refined with biomechanical metrics to ensure physiological plausibility, producing motions that are both individualized and biomechanically consistent.

We evaluate PG-BIG on the Motus Global movement-screen dataset Zhao et al. (2023), containing over 5,000 motion-capture trials from 183 athletes across 30 mobility and stability tasks. PG-BIG reduces average foot penetration and floating errors by over 40% relative to state-of-the-art baselines and achieves the lowest joint-space reconstruction error (1.47 ± 0.79 cm). It also improves biomechanical smoothness and energy consistency, producing physiologically stable and stylistically faithful motions. This demonstrates a new approach to clinically interpretable, personalized movement synthesis.

2. Related Works

Biomechanical Simulation and Motion Synthesis. Classical biomechanical simulations are based on musculotendon dynamics models, such as the Hill-type formulation Hill (1938) and its extensions by Zajac Zajac (1989). These foundations underpin frameworks like OpenSim Seth et al. (2018); Delp et al. (2007b), Moco Dembia et al. (2021); De Groote et al. (2016), and other platforms (MuJoCo Todorov et al. (2012), MyoSuite Vittorio et al. (2022)). However, they are computationally expensive and impractical for real-time applications. Recent generative models attempt to address this: BIGE Maheshwari (2025) integrates clinical biomechanics into motion generation using differentiable surrogates, while FlexMotion Tashakori et al. (2025) employs physics-aware transformers with Euler–Lagrange and muscle coordination losses to enhance realism. These methods improve biomechanical validity but lack subject-specific personalization.

Personalized and Stylistic Human Motion Generation. Early motion synthesis methods reused pre-recorded motion sequences through motion graphs and interpolation techniques Kovar et al.

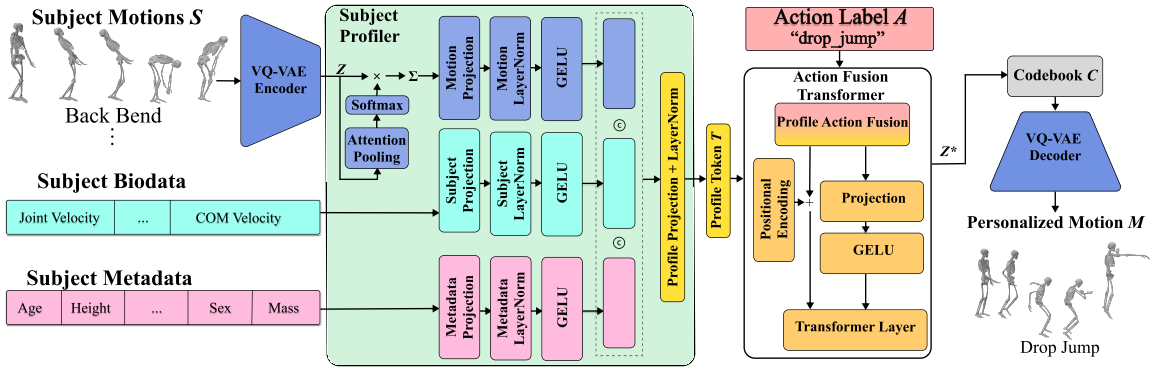

Figure 2: Overview of PG-BIG Inference Pipeline. Given subject-specific motion sequences S, we first encode them into a latent representation using a VQ-VAE. A *Subject Profiler* generates profile tokens from these latents, subject biodata, and metadata. More information on the *Subject Profiler* can be found A.6. We then pool the profile tokens and fuse them with an action label A and subsequently passed through a transformer layer. The outputs are quantized using the codebook C then the VQ-VAE decoder transforms the conditioned latent into the personalized motion M.

(2023). While they achieve visually coherent results, they fail to capture anatomical or stylistic variations across individuals. Subsequent multilinear and tensor-based models Min et al. (2010) factorized motion data into actor-specific and emotion-related components, enabling limited style transfer but with dense subject-specific capture data which struggles to generalize across diverse actions. Recent advances in deep generative modeling,particularly diffusion-based frameworks, have substantially improved realism and controllability in human motion synthesis Holden et al. (2016); Tevet et al. (2022a); Zhang et al. (2023); Tevet et al. (2022b). These models however, often average out subject-specific nuances, producing motions that are realistic yet impersonal. To address this limitation, recent approaches have introduced subject-conditioned generation, such as PersonaBooth Kim et al. (2025), which learns identity embeddings from limited samples to personalize diffusion-based text-to-motion models, and other frameworks that disentangle identity and content representations for motion style transfer Guo et al. (2024). While these methods capture stylistic movements, they lack biomechanical grounding and may produce anatomically implausible movements. PG-BIG builds on these advances by jointly modeling stylistic personalization and biomechanical guidance, generating motions that are both individualized and physiologically valid.

3. Methodology

We propose PG-BIG, a generative framework for personalized, biomechanically plausible motion synthesis. As shown in Figure 2, given subject-specific motion sequences S, athlete biodata and metadata s, and an action label A, the model learns the conditional distribution $p(M \mid s, A)$ over motion outputs $M = \{m_t\}_{t=1}^{T}$. PG-BIG combines three components: (i) a VQ-VAE-based discrete latent representation Z, (ii) profile–action conditioning via a learned subject profiler and cross-attention to produce conditioned latents Z^*, and (iii) a differentiable biomechanical guidance module that refines Z^* for physiological plausibility. The primary goal is the integration of personalization and biomechanical refinement. PG-BIG achieves a strong balances both, as supported by the ablation results and surrogate-validation analysis in Sections 4.3 and A.8.

3.1. Profile-Conditioned Latent Motion Modeling

PG-BIG represents motion compactly using a VQ-VAE Van Den Oord et al. (2017), which encodes trajectories X into discrete latent tokens via a learned codebook C. We map athlete metadata s to a profile embedding T, and action labels A to task embeddings. These conditioning vectors guide the resampling of the motion latent Z into a conditioned latent Z^*, which the decoder transforms into personalized motion M aligned with both individual style and task intent. We train the VQ-VAE with standard reconstruction and commitment losses to learn reusable motion primitives for efficient autoregressive sampling. We primarily choose VQ-VAE over diffusion-based generators because our setting requires fast, stable, and repeatedly guided inference in latent space: discrete codebook tokens provide compact motion primitives that are easy to condition and refine, while avoiding the multi-step sampling cost of diffusion that offers limited additional benefit for this profile-conditioned refinement task. Compared to standard continuous VAEs, the discrete codebook further improves latent structure and robustness through quantization regularization, helping prevent posterior collapse and supporting real-time generation.

During inference, the model autoregressively samples discrete motion tokens conditioned on (z_s, z_c), then decodes them into continuous kinematic trajectories. This factorized modeling strategy isolates motion structure from subject identity, supporting personalized synthesis without overfitting to individual data.

3.2. Cross-Attention Fusion

To integrate subject-specific style and task intent, PG-BIG employs a transformer-based cross-attention mechanism. The athlete embedding $\mathbf{z}_s \in \mathbb{R}^{d_s}$ encodes individual traits such as anthropometry and skill. In contrast, the action embedding $\mathbf{z}_a \in \mathbb{R}^{d_a}$ represents the target movement type. We concatenate these into a joint conditioning vector $\mathbf{z}_{\text{cond}} = [\mathbf{z}_s; \mathbf{z}_a]$, which a lightweight MLP expands into $N_m = 8$ memory tokens $\mathbf{M} = \text{MLP}_{\text{expand}}(\mathbf{z}_{\text{cond}}) \in \mathbb{R}^{N_m \times d}$. For each motion sequence of length T, we form temporal query tokens as $\mathbf{Q} = \mathbf{E}_{\text{pos}} + \mathbf{1}_T \otimes \mathbf{z}_{\text{cond}}$, where $\mathbf{E}_{\text{pos}} \in \mathbb{R}^{T \times d}$ denotes positional encodings and $\mathbf{1}_T$ broadcasts the conditioning to each time step. A transformer decoder performs cross-attention between $\mathbf{Q}$ and $\mathbf{M}$, producing contextualized features $\mathbf{Z} \in \mathbb{R}^{T \times d}$ that we discretize via VQ-VAE codebook lookup. This enables PG-BIG to reason over athlete identity and biomechanical task constraints jointly during motion generation.

3.3. Biomechanical Guidance

Pose-driven models may yield visually plausible yet biomechanically implausible motion. To address this, PG-BIG introduces a post-generation refinement stage using a differentiable biomechanical surrogate trained on ground-truth OpenSim outputs. This transformer-based module estimates joint kinematics, muscle activations, and center-of-mass (COM) dynamics, achieving high correlation with full OpenSim simulations on validation data. The surrogate enables lightweight gradient-based corrections that enforce physiologically meaningful motion while maintaining computational efficiency for real-time applications.

We define five biomechanical guidance components using statistical normalization:

Angular Velocity. Joint angular velocities $\dot{\boldsymbol{q}}_t$ are constrained within 95th percentile ranges $\dot{q}_j^{\text{high}}$ derived from training data. Using the upper bound prevents unnaturally fast joint rotations while

allowing movement-appropriate speeds, ensuring smooth transitions between poses without restricting athletic motions like jumps or lunges.

Foot Contact Consistency. Foot sliding is penalized using foot velocity $\mathbf{v}_t^{(\text{foot})}$ and binary contact mask $\mathbf{m}_t$. This physical constraint enforces zero velocity during stance phases, preventing visual artifacts and ensuring biomechanically plausible weight transfer during ground contact across all movement types.

COM Velocity. Center-of-mass velocity $\dot{\mathbf{c}}_t$ is bounded by 5-95th percentile limits $\dot{c}^{\text{low}}$ and $\dot{c}^{\text{high}}$. The bilateral bounds maintain natural locomotor momentum for dynamic tasks while preventing unnaturally slow or fast global motion, accommodating both rapid direction changes and controlled balance poses.

COM Acceleration. COM acceleration $\ddot{\mathbf{c}}_t$ is constrained to percentile bounds $\ddot{c}^{\text{low}}$ and $\ddot{c}^{\text{high}}$. These limits ensure physically plausible force generation patterns, preventing jerky motion and sudden momentum changes while allowing rapid accelerations needed in athletic movements like jumping and cutting.

Muscle Activation. Muscle activations $\mathbf{a}_t$ are bounded by 5-95th percentile ranges a_m^{low} and a_m^{high} from training data. This statistical approach prevents physiologically implausible co-contraction or under-activation while accommodating the varied recruitment patterns needed across different exercises and skill levels.

Table 1: Biomechanical guidance components with statistical normalization.

Metric	Definition	Guidance Loss
Angular Velocity	$\dot{\boldsymbol{q}}_t$: joint angular velocity	$L_\omega = \frac{1}{T}\sum_t \sum_j \max(0, \|\dot{\boldsymbol{q}}_{t,j}\|_2 - \dot{q}_j^{\text{high}})$
Foot Contact Consistency	$\mathbf{v}_t^{(\text{foot})}$: foot velocity; $\mathbf{m}_t$: contact mask	$L_{\text{foot}} = \frac{1}{T}\sum_t \|\mathbf{v}_t^{(\text{foot})} \odot \mathbf{m}_t\|_2^2$
COM Velocity	$\dot{\mathbf{c}}_t$: COM velocity	$L_v = \frac{1}{T}\sum_t \max(0, \|\dot{\mathbf{c}}_t\|_2 - \dot{c}^{\text{high}}, \dot{c}^{\text{low}} - \|\dot{\mathbf{c}}_t\|_2)$
COM Acceleration	$\ddot{\mathbf{c}}_t$: COM acceleration	$L_a = \frac{1}{T}\sum_t \max(0, \|\ddot{\mathbf{c}}_t\|_2 - \ddot{c}^{\text{high}}, \ddot{c}^{\text{low}} - \|\ddot{\mathbf{c}}_t\|_2)$
Muscle Activation	$\mathbf{a}_t$: activations bounded by 5–95th percentiles	$L_{\text{act}} = \frac{1}{T}\sum_t \sum_m \max(0, a_{t,m} - a_m^{\text{high}}, a_m^{\text{low}} - a_{t,m})$

The overall biomechanical guidance loss combines all components as:

$$\mathcal{L}_{\text{guide}} = \lambda_\omega L_\omega + \lambda_{\text{foot}} L_{\text{foot}} + \lambda_v L_v + \lambda_a L_a + \lambda_{\text{act}} L_{\text{act}}$$

4. Experiments

We evaluate PG-BIG on subject-specific human motion synthesis across functional athletic tasks. Our experiments aim to answer three key questions: (i) Does PG-BIG generate biomechanically valid motion? (ii) Does conditioning on athlete profiles enhance personalization? and (iii) How does PG-BIG perform relative to state-of-the-art generative motion models?

4.1. Datasets and Experimental Setup

We evaluate PG-BIG on the Motus Global Movement-Screen dataset (Zhao et al., 2023), comprising over 5000 motion-capture trials from 183 athletes performing 30 tasks. For biomechanical guidance, each sequence is fitted to a scaled OpenSim model (Seth et al., 2018) to estimate joint kinematics and muscle activations used at inference.

For biomechanical guidance during inference, we fit each motion sequence to a scaled OpenSim musculoskeletal model. Specifically, we employ the full-body Rajagopal model (Rajagopal et al., 2016b), scaled per subject using static calibration poses following OpenSim procedures (Delp et al., 2007a). Joint trajectories are found using inverse kinematics with marker-error minimization, and muscle activation envelopes are estimated using simplified static optimization (Thelen et al., 2003). These biomechanical quantities are used for inference-time guidance, avoiding the need for differentiable physics during training.

Our generative backbone is a VQ-VAE (Van Den Oord et al., 2017) with an autoregressive Transformer prior conditioned on athlete profiles and action labels. Athlete profiles encode anthropometric and performance-related attributes, enabling personalized motion synthesis. More implementation details are deferred to A.

4.2. Evaluation Metrics

We evaluate both biomechanical validity and subject-specific fidelity. All metrics are computed per motion sequence and averaged across the test set.

- **Fidelity.** Measures subject-specific accuracy via joint-space error $\|\boldsymbol{q}_t^{\text{pred}} - \boldsymbol{q}_t^{\text{ref}}\|_2$.
- **Floor Penetration.** Measures average negative foot height relative to the ground ($\max(0, -z_t^{(\text{foot})})$), indicating physical violations.
- **Motion Entropy.** Quantifies motion diversity using state-space entropy $H = -\sum_k p_k \log p_k$, where higher values indicate richer, non-collapsed motion.
- **Floor Floating.** Penalizes foot hovering during stance ($\max(0, z_t - \varepsilon)$), for stable ground contact.
- **Foot Contact Velocity.** Evaluates the horizontal velocity of the foot $\|\dot{\mathbf{p}}_{t,xy}^{(\text{foot})}\|_2$ during contact phases, discouraging sliding.
- **Jerk Index (J).** Quantifies neuromotor smoothness via squared joint jerk $\|\dddot{\boldsymbol{q}}_t\|_2^2$, with lower values corresponding to smoother, more human-like motion.
- **Energy Consistency (E_{var}).** Measures temporal fluctuations in total mechanical energy $E = E_{\text{kin}} + E_{\text{pot}}$; lower variance indicates biomechanically stable motion.

4.3. Baseline Comparison

We compare PG-BIG against representative state-of-the-art baselines spanning text-conditioned generation and personalization:

- **Reference** (Zhao et al., 2023): Motion capture data from the Motus Global movement-screen dataset and associated muscle activation simulations.
- **MDM** (Tevet et al., 2022b): A diffusion-based model for human motion synthesis, serving as a state-of-the-art baseline.
- **T2M-GPT** (Zhang et al., 2023): A transformer-based text-conditioned generative model capturing task semantics and stylistic realism.

- **PersonaBooth** (Kim et al., 2025): A personalized motion synthesis framework that adapts stylistic traits to individual subjects.

Baseline Comparison and Ablation Insight. Table 2 presents quantitative results across realism and biomechanical metrics. PG-BIG achieves the lowest penetration, floating, and foot contact errors, demonstrating superior physical grounding compared to all baselines.

Ablation Insight. A key observation is that *w/o Guidance* already outperforms population-level baselines on fidelity. This indicates that subject-conditioned generation contributes strongly to identity/style preservation. Biomechanical guidance then provides complementary gains by reducing physical violations: penetration, floating, and sliding, while maintaining competitive fidelity. Together, these results support that personalization and guidance play distinct, complementary roles.

Table 2: Baseline comparison across realism and biomechanical metrics. ↓ indicates lower is better, → indicates values closer to reference are better.

Metric	Reference	MDM	PersonaBooth	T2M-GPT	PG-BIG (Ours)
Fidelity (cm) (↓)	–	5.12 ± 3.50	3.60 ± 2.30	9.45 ± 4.70	**1.47 ± 0.79**
Penetration (cm) (↓)	2.29 ± 1.38	1.46 ± 0.65	1.39 ± 0.44	1.70 ± 0.54	**1.28 ± 0.54**
Floating (cm) (↓)	2.07 ± 1.04	1.84 ± 0.92	2.14 ± 0.91	2.11 ± 0.87	**1.28 ± 0.64**
Sliding (cm) (↓)	0.52 ± 0.26	3.20 ± 1.03	3.12 ± 0.88	5.11 ± 1.52	**0.75 ± 0.38**
Entropy (→)	3.47 ± 0.14	3.34 ± 0.11	3.41 ± 0.08	**3.47 ± 0.10**	3.39 ± 0.12
Jerk (m/s^3) (↓)	0.58 ± 0.29	38.96 ± 19.50	48.18 ± 14.17	64.35 ± 24.22	**5.74 ± 2.03**
Energy (J) (↓)	1.90 ± 0.95	2.68 ± 1.34	5.40 ± 1.18	2.95 ± 0.95	**2.48 ± 1.18**

Figure 3 visualizes generated motions across representative tasks selected to probe complementary aspects of motor control. These tasks are illustrative examples from the 30+ trials in our evaluation, chosen to highlight both movement generation and postural maintenance under varying biomechanical demands.

Baseline models exhibit several characteristic failures. MDM and T2M-GPT frequently exhibit foot sliding during contact phases and inconsistent foot placement, leading to unstable weight transitions. PersonaBooth shows improved stylistic consistency but struggles with contact handling, producing exaggerated postural sway and unnatural compensatory movements. Critically, baseline methods tend to generate shallower, less committed movements.

In contrast, PG-BIG produces deeper, more athletic lunges with stable knee-hip alignment and proper heel-to-toe placement. During the T-balance, PG-BIG maintains consistent ground contact with the stance foot, generates appropriate postural adjustments, and achieves fuller extension of the free limb. These results demonstrate that biomechanical guidance enables PG-BIG to generate motion that is simultaneously physically grounded and athletically realistic.

4.4. Ablation Study

To assess the contributions of profile conditioning and biomechanical guidance, we evaluate two ablated variants of PG-BIG:

- **w/o Personalization**: Retains biomechanical guidance but removes subject-specific profile conditioning, producing motion that adheres to physical constraints but lacks stylistic individuality.

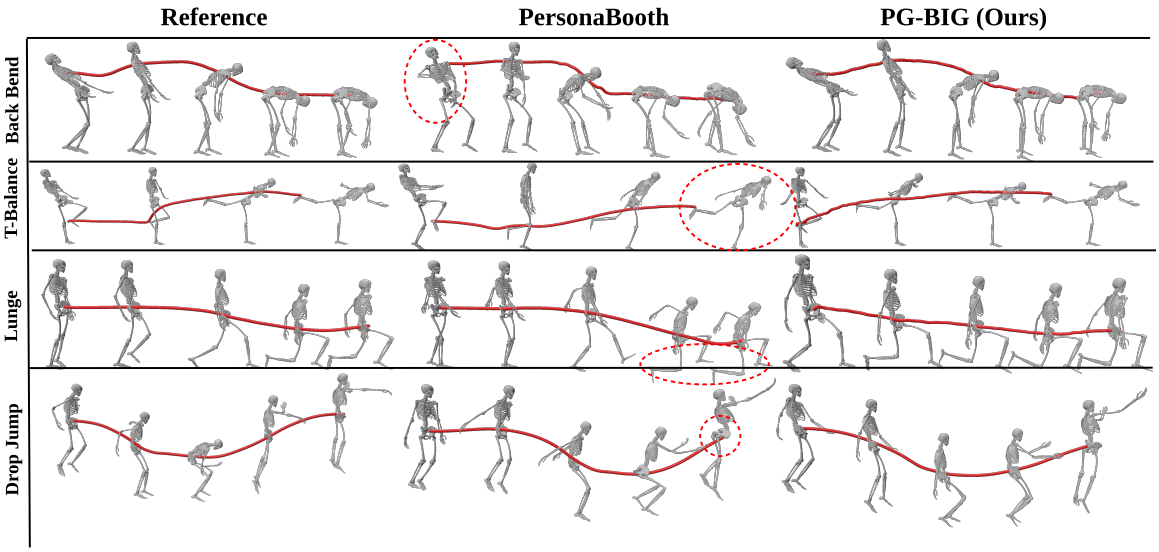

Figure 3: Qualitative comparison across four representative tasks. We generate motions using PG-BIG, state-of-the-art baselines (MDM, PersonaBooth, T2M-GPT), and show ground-truth reference motion. Baseline models exhibit foot drift, postural instability, and exaggerated sway. In contrast, PG-BIG maintains consistent ground contact, stable center of mass control, and subject-specific movement style. Red circles highlight biomechanical errors in baseline outputs.

- **w/o Guidance**: Includes athlete profile conditioning but omits biomechanical guidance, generating personalized motion that may exhibit physical inaccuracies.

Table 3 summarizes the quantitative results. The *w/o Guidance* variant shows degraded physical realism, with higher penetration, floating, and sliding errors. Conversely, the *w/o Personalization* model produces the lower penetration and floating values, reflecting improved biomechanical consistency, but exhibits weaker fidelity and less stable kinematics.

The full *PG-BIG* model achieves the best overall trade-off, maintaining the top fidelity while closely matching ablated variants across physical realism metrics, including jerk, energy, and sliding, which suggests that the model balances personalization with dynamic stability.

Table 3: Ablation analysis. ↓ means lower is better, and → is better when closer to reference values.

Metric	Reference	w/o Guidance	w/o Personalization	PG-BIG
Fidelity (cm) (↓)	–	3.19 ± 1.38	1.72 ± 2.58	**1.47 ± 0.79**
Penetration (cm) (↓)	2.29 ± 1.38	1.42 ± 1.30	**0.91 ± 0.63**	1.28 ± 0.54
Floating (cm) (↓)	2.07 ± 1.04	1.33 ± 1.02	**0.85 ± 0.34**	1.28 ± 0.64
Sliding (cm) (↓)	0.52 ± 0.26	0.94 ± 0.65	0.78 ± 0.34	**0.75 ± 0.38**
Entropy (→)	3.47 ± 0.14	**3.40 ± 0.28**	3.31 ± 0.23	3.39 ± 0.12
Jerk (J) (m/s^3) (↓)	0.58 ± 0.29	7.41 ± 2.86	5.78 ± 2.04	**5.74 ± 2.03**
Energy (↓)	1.90 ± 0.95	**3.02 ± 1.48**	2.22 ± 1.27	**2.08 ± 1.18**

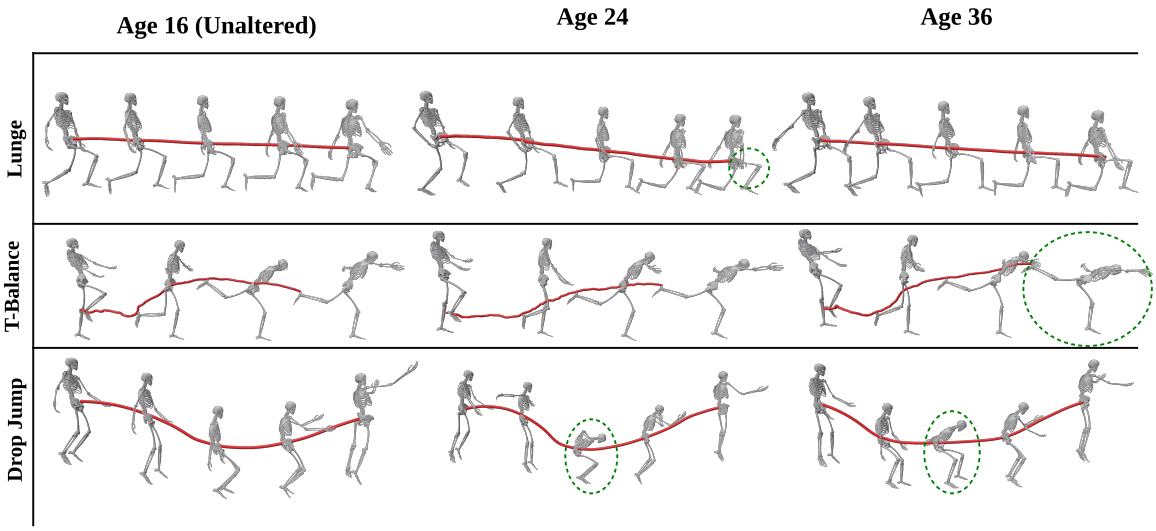

Figure 4: Synthetic age manipulation across mobility and stability tasks. We manipulate athlete age profiles and generate motions for mobility and stability tasks. PG-BIG produces age-appropriate kinematic adaptations: older profiles exhibit reduced range of motion, conservative postures, and decreased movement amplitude, while maintaining biomechanical validity. Green dashed circles highlight kinematic features altered by the age manipulation.

4.5. Synthetic Profile Manipulation

To assess the controllability and interpretability of profile conditioning, we synthetically tune athlete attributes and generate corresponding motions while holding the task constant. Modifications include altering age group, injury status, and training experience. This demonstrates that PG-BIG can produce appropriate responses to profile changes, highlighting its potential for personalized adaptation studies and hypothesis-driven biomechanical manipulation.

Quantitative Validation. To verify that profile manipulations produce meaningful kinematic changes, we train lightweight classifiers on the motion embeddings produced by the VQ-VAE encoder. Specifically, we use simple MLP models to predict athlete age via regression and athlete identity via classification. The age regression model achieves a mean squared error of 2.4 years, while the identity classifier attains 94% accuracy. These results confirm that the learned motion representations capture subject-specific and age-related movement signatures, validating that synthetic profile perturbations induce detectable and interpretable kinematic changes. Additional classification experiments for injury status, limb dominance, and training level are provided in the appendix.

Qualitative Analysis. Figure 4 shows generated motions under synthetic profile manipulations for mobility and stability tasks. We synthetically manipulate athlete profiles. Increasing age produces reduced joint range of motion, narrower stance, and smaller movement amplitude, consistent with age-related mobility decline, whereas reducing training experience produces shallower lunges, less committed weight transfers, and altered trunk compensation, reflecting skill-level differences. These adaptations occur smoothly without mode collapse or biomechanical degradation: PG-BIG maintains stable foot contact, generates physiologically valid joint trajectories, and controls center-of-mass dynamics. PG-BIG's profile conditioning provides both interpretable and controllable mo-

tion generation, enabling exploration of individualized movement adaptations across demographic and training variables.

5. Discussion and Conclusion

We introduced PG-BIG, a generative motion framework that combines personalization and biomechanical guidance to synthesize athlete-specific movements. By conditioning the generative prior on learned athlete-specific profile embeddings and applying differentiable biomechanical guidance losses to the latent motion sequence during inference, PG-BIG integrates subject-level stylistic cues with constraints on joint kinematics, foot–ground interaction, and COM dynamics. Experimental results demonstrate that personalized conditioning improves fidelity to individual movement patterns, while biomechanical guidance enhances balance, contact stability, and kinematic smoothness, yielding natural and individualized motion sequences.

A central finding is the complementarity between personalization and biomechanical guidance. Removing profile conditioning produces physically stable yet generic motion, whereas omitting biomechanical guidance results in stylistically expressive but dynamically inconsistent trajectories. This highlights the importance of jointly modeling subject identity and physical validity in generative motion systems. Our findings also suggest that interpretable biomechanical metrics—such as center-of-mass stability, joint smoothness, and muscle activation can effectively regularize motion synthesis without the need for full physics-based simulation.

Limitations and Future Work. PG-BIG currently uses static athlete profiles that do not capture longitudinal changes such as fatigue and adaptation. The Rajagopal musculoskeletal model employs population-level muscle parameters. Subject-specific parameters would improve guidance precision but was unavailable in our dataset. Additionally, the model includes only lower-body muscles, limiting precision for upper-body-focused movements (e.g., shoulder rotations). Evaluation is also limited to controlled laboratory movement screens; open-environment and text-conditioned generation remain untested.

We include sport-specific personalization analysis in the appendix. A sex-specific analysis is less informative in our current setup because all trajectories are retargeted and scaled to a common musculoskeletal representation, which reduces sex-dependent anatomical differences for guidance. Future work will explore subject-specific musculoskeletal parameterization, full-body biomechanical guidance, and longitudinal profile adaptation.

Acknowledgment

This work was supported in part by the Wu Tsai Human Performance Alliance, U.S. Army Research Office under Army-ECASE award W911NF-07-R-0003-03, the U.S. Department Of Energy, Office of Science, IARPA HAYSTAC Program, and NSF Grants #2205093, #2146343, #2134274, #2441832, CDC-RFA-FT-23-0069, DARPA AIE FoundSci and DARPA YFA.

References

Friedl De Groote, Allison Kinney, Anil Rao, and Benjamin Fregly. Evaluation of direct collocation optimal control problem formulations for solving the muscle redundancy problem. *Annals of Biomedical Engineering*, 44, 03 2016. doi: 10.1007/s10439-016-1591-9.

Scott L Delp, Frank C Anderson, Allison S Arnold, Peter Loan, Ayman Habib, Chand T John, Eran Guendelman, and Darryl G Thelen. Opensim: open-source software to create and analyze dynamic simulations of movement. *IEEE transactions on biomedical engineering*, 54(11):1940–1950, 2007a.

Scott L. Delp, Frank C. Anderson, Allison S. Arnold, Peter Loan, Ayman Habib, Chand T. John, Eran Guendelman, and Darryl G. Thelen. Opensim: Open-source software to create and analyze dynamic simulations of movement. *IEEE Transactions on Biomedical Engineering*, 54(11): 1940–1950, 2007b. doi: 10.1109/TBME.2007.901024.

Christopher L. Dembia, Nicholas A. Bianco, Antoine Falisse, Jennifer L. Hicks, and Scott L. Delp. Opensim moco: Musculoskeletal optimal control. *PLOS Computational Biology*, 16(12):1–21, 12 2021. doi: 10.1371/journal.pcbi.1008493. URL https://doi.org/10.1371/journal.pcbi.1008493.

Antoine Falisse, Gil Serrancolí, Christopher L Dembia, Joris Gillis, Ilse Jonkers, and Friedl De Groote. Rapid predictive simulations with complex musculoskeletal models suggest that diverse healthy and pathological human gaits can emerge from similar control strategies. *Journal of The Royal Society Interface*, 16(157):20190402, 2019.

Chuan Guo, Shihao Zou, Xinxin Zuo, Sen Wang, Wei Ji, Xingyu Li, and Li Cheng. Generating diverse and natural 3d human motions from text. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5152–5161, 2022.

Chuan Guo, Yuxuan Mu, Xinxin Zuo, Peng Dai, Youliang Yan, Juwei Lu, and Li Cheng. Generative human motion stylization in latent space. *arXiv preprint arXiv:2401.13505*, 2024.

Archibald Vivian Hill. The heat of shortening and the dynamic constants of muscle. *Proceedings of the Royal Society of London. Series B - Biological Sciences*, 126(843):136–195, 1938. doi: 10.1098/rspb.1938.0050. URL https://royalsocietypublishing.org/doi/abs/10.1098/rspb.1938.0050.

Daniel Holden, Jun Saito, and Taku Komura. A deep learning framework for character motion synthesis and editing. *ACM Transactions on Graphics (ToG)*, 35(4):1–11, 2016.

Marilyn Keller, Keenon Werling, Soyong Shin, Scott Delp, Sergi Pujades, C. Karen Liu, and Michael J. Black. From skin to skeleton: Towards biomechanically accurate 3d digital humans. In *ACM ToG, Proc. SIGGRAPH Asia*, volume 42, December 2023.

Boeun Kim, Hea In Jeong, JungHoon Sung, Yihua Cheng, Jeongmin Lee, Ju Yong Chang, Sang-Il Choi, Younggeun Choi, Saim Shin, Jungho Kim, et al. Personabooth: Personalized text-to-motion generation. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 22756–22765, 2025.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Lucas Kovar, Michael Gleicher, and Frédéric Pighin. Motion graphs. In *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*, pages 723–732. 2023.

Shubh Maheshwari. Bige: Biomechanics-informed genai for exercise science. Master's thesis, University of California, San Diego, 2025.

Jianyuan Min, Huajun Liu, and Jinxiang Chai. Synthesis and editing of personalized stylistic human motion. In *Proceedings of the 2010 ACM SIGGRAPH symposium on Interactive 3D Graphics and Games*, pages 39–46, 2010.

Apoorva Rajagopal, Christopher L. Dembia, Matthew S. DeMers, Denny D. Delp, Jennifer L. Hicks, and Scott L. Delp. Full-body musculoskeletal model for muscle-driven simulation of human gait. *IEEE Transactions on Biomedical Engineering*, 63:2068–2079, 2016a. URL https://api.semanticscholar.org/CorpusID:3798525.

Apoorva Rajagopal, Christopher L Dembia, Matthew S DeMers, Denny D Delp, Jennifer L Hicks, and Scott L Delp. Full-body musculoskeletal model for muscle-driven simulation of human gait. *IEEE transactions on biomedical engineering*, 63(10):2068–2079, 2016b.

Ajay Seth, Jennifer L. Hicks, Thomas K. Uchida, Ayman Habib, Christopher L. Dembia, James J. Dunne, Carmichael F. Ong, Matthew S. DeMers, Apoorva Rajagopal, Matthew Millard, Samuel R. Hamner, Edith M. Arnold, Jennifer R. Yong, Shrinidhi K. Lakshmikanth, Michael A. Sherman, Joy P. Ku, and Scott L. Delp. Opensim: Simulating musculoskeletal dynamics and neuromuscular control to study human and animal movement. *PLOS Computational Biology*, 14(7):1–20, 07 2018. doi: 10.1371/journal.pcbi.1006223. URL https://doi.org/10.1371/journal.pcbi.1006223.

Arvin Tashakori, Arash Tashakori, Gongbo Yang, Z Jane Wang, and Peyman Servati. Flexmotion: Lightweight, physics-aware, and controllable human motion generation. *arXiv preprint arXiv:2501.16778*, 2025.

Guy Tevet, Brian Gordon, Amir Hertz, Amit H Bermano, and Daniel Cohen-Or. Motionclip: Exposing human motion generation to clip space. In *European Conference on Computer Vision*, pages 358–374. Springer, 2022a.

Guy Tevet, Sigal Raab, Brian Gordon, Yonatan Shafir, Daniel Cohen-Or, and Amit H Bermano. Human motion diffusion model. *arXiv preprint arXiv:2209.14916*, 2022b.

Darryl G Thelen, Frank C Anderson, and Scott L Delp. Generating dynamic simulations of movement using computed muscle control. *Journal of biomechanics*, 36(3):321–328, 2003.

Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5026–5033. IEEE, 2012. doi: 10.1109/IROS.2012.6386109.

Aaron Van Den Oord, Oriol Vinyals, et al. Neural discrete representation learning. *Advances in neural information processing systems*, 30, 2017.

Caggiano Vittorio, Wang Huawei, Durandau Guillaume, Sartori Massimo, and Kumar Vikash. Myosuite – a contact-rich simulation suite for musculoskeletal motor control, 2022. URL https://arxiv.org/abs/2205.13600.

Andreas Wächter and Lorenz T Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming*, 106(1): 25–57, 2006.

Keenon Werling, Dalton Omens, Jeongseok Lee, Ioannis Exarchos, and Karen Liu. Fast and feature-complete differentiable physics engine for articulated rigid bodies with contact constraints. 07 2021. doi: 10.15607/RSS.2021.XVII.034.

Felix E. Zajac. Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control. *Critical reviews in biomedical engineering*, 17 4:359–411, 1989. URL https://api.semanticscholar.org/CorpusID:25888803.

Jianrong Zhang, Yangsong Zhang, Xiaodong Cun, Yong Zhang, Hongwei Zhao, Hongtao Lu, Xi Shen, and Ying Shan. T2m-gpt: Generating human motion from textual descriptions with discrete representations. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14730–14740, 2023.

Xiong Zhao, Gwyneth Ross, Brittany Dowling, and Ryan B Graham. Three-dimensional motion capture data of a movement screen from 183 athletes. *Scientific Data*, 10(1):235, 2023.

Appendix A. Implementation Details

A.1. Hyperparameters

Table 4 and Table 5 list the hyperparameters used for PG-BIG's VQ-VAE training and biomechanical guidance, respectively. During guidance, we employ a staged schedule: an initial warm-up phase (the first 500 iterations) that applies only the proximity loss, followed by the progressive activation of biomechanical constraints α (see Table 5). In the final phase, we reweight proximity loss to refine motion smoothness and contact stability.

Table 4: Hyperparameters for VQ-VAE training

Argument	Value	Description
total-iterations	300000	Total training iterations
batch-size	128	Training batch size
learning-rate	0.0002	Initial learning rate
lr-scheduler	50000	LR decay frequency
gamma	0.05	LR decay factor
loss-commitment	0.02	Codebook commitment loss weight
loss-penetration	0.05	Penetration loss weight
embedding-dimension	512	Latent embedding dimension
nb-codes	512	Codebook size

Table 5: Hyperparameters for Action Fusion Encoder

Argument	Value	Description
total-iterations	250	Total iterations
physics-loss-iteration	100	Iteration where physics loss was added
quantization-iteration	200	Iteration where Z^* was quantized
batch-size	32	Number of sampled trajectories
λ_ω	10	Angular velocity
λ_v	0.1	Center of mass velocity
λ_a	1	Center of mass acceleration
λ_{act}	10	Muscle activation

A.2. Motion Retargeting

To enable biomechanical analysis and ensure anatomical consistency, we retarget raw motion-capture marker data from the Motus Global dataset to the Rajagopal full-body musculoskeletal model (Rajagopal et al., 2016b) using inverse kinematics (IK).
The retargeting pipeline consists of three stages:

Calibration and Marker Offset Computation. For each subject, we first process static calibration trials to compute subject-specific marker-to-body offsets. Given calibration frame markers

$\mathbf{m}_j \in \mathbb{R}^3$ and their corresponding body nodes b_j on the skeleton, we solve:

$$\min_{\mathbf{q},\{\boldsymbol{\delta}_j\}} \sum_j w_j \|\mathbf{m}_j - (\mathbf{T}_j(\mathbf{q}) + \mathbf{R}_j(\mathbf{q})\boldsymbol{\delta}_j)\|_2^2, \tag{1}$$

where $\mathbf{q}$ represents generalized joint coordinates, $\mathbf{T}_j(\mathbf{q})$ and $\mathbf{R}_j(\mathbf{q})$ are the world position and rotation of body b_j, and $\boldsymbol{\delta}_j$ is the local marker offset. Offsets are averaged across calibration frames using statistics (5th–95th percentile filtering) to reduce noise.

Per-Frame Inverse Kinematics. For each trial frame, we fit the skeleton pose to observed marker positions via weighted least-squares IK with body scaling enabled:

$$\mathbf{q}_t = \arg\min_{\mathbf{q}} \sum_j w_j \|\mathbf{m}_{j,t} - (\mathbf{T}_j(\mathbf{q}) + \mathbf{R}_j(\mathbf{q})\boldsymbol{\delta}_j)\|_2^2, \tag{2}$$

where $\mathbf{m}_{j,t}$ are observed marker positions at time t. We warm-start each frame with the solution from the previous frame ($\mathbf{q}_{t-1}$) to maintain temporal continuity. Invalid markers (NaN, zero coordinates, or physically implausible positions) are excluded from the optimization.

Temporal Filtering. To remove high-frequency noise introduced by marker dropout and IK optimization artifacts, we apply a second-order Butterworth low-pass filter with a cutoff frequency of 10 Hz to the resulting joint angle trajectories. The retargeting process yields kinematically consistent pose sequences in the Rajagopal model's representation, enabling downstream biomechanical analysis, including joint kinematics, center-of-mass dynamics, and muscle activation estimation. All retargeted motion data are stored in the Biomechanics 3D (B3D) format (Werling et al., 2021) for efficient access during training and inference.

A.3. Ground Truth Muscle Activation Generation

To obtain ground-truth muscle activation profiles for training the biomechanical surrogate model, we employ muscle-driven optimal control in OpenSimAD (Falisse et al., 2019). This process transforms retargeted motion sequences into physiologically realistic muscle recruitment patterns. Details on the surrogate model can be found in A.8

Processing Pipeline. The pipeline consists of three main stages. First, we extract marker trajectories from B3D motion capture files (Werling et al., 2021) and convert them to OpenSim's MOT format. Second, we solve inverse kinematics (IK) to obtain joint angle trajectories $\{\mathbf{q}_t\}_{t=1}^T$ that minimize marker reprojection error. Third, we estimate muscle activations via trajectory tracking optimization:

$$\min_{\{\mathbf{a}_t\}} \sum_{t=1}^{T} \left(\|\mathbf{q}_t - \mathbf{q}_t^{\text{ref}}\|_2^2 + \lambda_a \|\mathbf{a}_t\|_2^2 \right), \tag{3}$$

to match musculoskeletal dynamics constraints including Hill-type muscle models (Hill, 1938; Zajac, 1989), joint torque limits, and ground contact forces. We solve this using direct collocation with IPOPT (Wächter and Biegler, 2006), applying mesh scaling ($\alpha = 0.25$) and relaxed tolerances ($\epsilon = 10^{-4}$) for computational efficiency.

The optimized activations $\{\mathbf{a}_t^*\}_{t=1}^T$ represent normalized recruitment levels for $N_m = 80$ muscles in the Rajagopal model. We apply low-pass filtering (6 Hz Butterworth) to remove numerical

artifacts while preserving physiological dynamics. These activation sequences, along with center-of-mass trajectories, serve as ground truth for training the Transformer-based surrogate model used during inference-time biomechanical guidance.

A.4. Text Prompts for Baseline Evaluations

Table 6 contains the prompts used for motion generation from baselines. Because these models operate in a text-to-motion framework, we construct and tune natural-language prompts to best reproduce the target movements evaluated using PG-BIG. These prompts were derived from and refined based on the exercise descriptions in the supplementary material of the Motus Global Movement-Screen dataset.

Table 6: Exercise prompts adapted and rewritten from the original *Motus Global Movement-Screen* dataset descriptions, with left and right variants shown separately where applicable.

#	Description
1	Ankle (Left): Start on the left knee. With hands at the side or resting on the thigh, lean forward, bending at the ankle until the left heel rises. Return to the start.
2	Ankle (Right): Start on the right knee. With hands at the side or resting on the thigh, lean forward, bending at the ankle until the right heel rises. Return to the start.
3	Back Bend: Stand with feet hip-width apart and toes pointed forward. Keep knees straight. Arch the back to the end range, then bend forward at the waist. If reaching the toes, pull arms slightly out and back to allow full pelvic and trunk motion.
4	Left Crossover Adduction: Stand with toes and pelvis facing forward. Cross the right leg over the left for maximum distance while keeping legs fairly straight. After reaching maximum stretch, continue walking through, returning to standing.
5	Right Crossover Adduction: Stand with toes and pelvis facing forward. Cross the left leg over the right for maximum distance while keeping legs fairly straight. After reaching maximum stretch, continue walking through, returning to standing.
6	Left Crossover Rotation: Stand with the left foot planted and toes facing forward. Rotate the pelvis inward, crossing the right leg over. After achieving maximum rotation, walk through to return to neutral.
7	Right Crossover Rotation: Stand with the right foot planted and toes facing forward. Rotate the pelvis inward, crossing the left leg over. After achieving maximum rotation, walk through to return to neutral.
8	Elbows (Left): Stand with the left elbow at shoulder height, bent 90°, and shoulder rotated so the hand is near head height with a closed fist. Straighten to full extension and then bend to full flexion. Rotate the forearm fully in one direction, then the other.
9	Elbows (Right): Stand with the right elbow at shoulder height, bent 90°, and shoulder rotated so the hand is near head height with a closed fist. Straighten to full extension and then bend to full flexion. Rotate the forearm fully in one direction, then the other.

Continued on next page

Table 6 – continued from previous page

#	Description
10	Head: Stand with head facing forward. Drop the chin to the chest, tilt back, then rotate toward each shoulder. Finally, tilt side to side, bringing each ear toward the shoulder.
11	Left Hip Turn: Stand with feet hip-width apart. Keep the right foot stationary, lift the left foot slightly, turn the pelvis open, and maximally rotate on the hip of the right (planted) foot. Step through to finish.
12	Right Hip Turn: Stand with feet hip-width apart. Keep the left foot stationary, lift the right foot slightly, turn the pelvis open, and maximally rotate on the hip of the left (planted) foot. Step through to finish.
13	Scorpion (Left): Begin in quadruped position with wrists below shoulders and knees below hips. Keep elbows locked, back flat, and abdominals tight. Extend the left leg back with a 90° knee bend, then bring the knee toward the chest. Return to the start.
14	Scorpion (Right): Begin in quadruped position with wrists below shoulders and knees below hips. Keep elbows locked, back flat, and abdominals tight. Extend the right leg back with a 90° knee bend, then bring the knee toward the chest. Return to the start.
15	Shoulder Abduction (Left): Stand with elbows bent 90° at sides. Raise the left upper arm, bringing the elbow as high as possible while keeping it between the chest and back planes. Return to the start.
16	Shoulder Abduction (Right): Stand with elbows bent 90° at sides. Raise the right upper arm, bringing the elbow as high as possible while keeping it between the chest and back planes. Return to the start.
17	Shoulder Azimuth (Left): With the left elbow out at shoulder height, bent 90°, and palm down, move at the shoulder to bring the elbow toward the opposite shoulder, then back toward the spine.
18	Shoulder Azimuth (Right): With the right elbow out at shoulder height, bent 90°, and palm down, move at the shoulder to bring the elbow toward the opposite shoulder, then back toward the spine.
19	Shoulder Rotation (Left): With the left elbow out at shoulder height, bent 90°, and palm down, rotate the arm back at the shoulder to end range, then down to the opposite end range.
20	Shoulder Rotation (Right): With the right elbow out at shoulder height, bent 90°, and palm down, rotate the arm back at the shoulder to end range, then down to the opposite end range.
21	Side Bends (Left): Stand upright with arms at sides. Slide the left hand down the left leg, bending laterally at the waist. Return to the center.
22	Side Bends (Right): Stand upright with arms at sides. Slide the right hand down the right leg, bending laterally at the waist. Return to the center.
23	Side Lunge (Left): Stand with feet spread 80% of maximum. Keep chest upright, toes straight, and feet flat. Lunge to the left side, reaching maximum depth. Return to the center.

Continued on next page

Table 6 – continued from previous page

#	Description
24	Side Lunge (Right): Stand with feet spread 80% of maximum. Keep chest upright, toes straight, and feet flat. Lunge to the right side, reaching maximum depth. Return to the center.
25	Trunk Rotation (Left): Kneel with arms crossed in front of chest. Keep the pelvis still and rotate the upper trunk left to end range, then return to start.
26	Trunk Rotation (Right): Kneel with arms crossed in front of chest. Keep the pelvis still and rotate the upper trunk right to end range, then return to start.
27	Drop Jump: Stand on a 30 cm platform. Drop to the floor and immediately perform a maximal vertical jump, using arm swing for height.
28	Hop Down (Left): Stand on the left foot on a 15–30 cm box. Hop off, land on the same foot, and immediately perform a maximal vertical jump. Use arm swing for height.
29	Hop Down (Right): Stand on the right foot on a 15–30 cm box. Hop off, land on the same foot, and immediately perform a maximal vertical jump. Use arm swing for height.
30	L-Cut (Left): Stand with feet hip-width apart. Jump forward, land on the left foot, and immediately cut 90° to the right side, landing on both feet.
31	L-Cut (Right): Stand with feet hip-width apart. Jump forward, land on the right foot, and immediately cut 90° to the left side, landing on both feet.
32	Lunge (Left): Stand with feet hip-width apart and hands on hips. Lunge forward with the left leg until the knee nearly touches the ground, then step back into a backward lunge without touching the floor. Return to the start.
33	Lunge (Right): Stand with feet hip-width apart and hands on hips. Lunge forward with the right leg until the knee nearly touches the ground, then step back into a backward lunge without touching the floor. Return to the start.
34	Rotary Stability (Left Arm/Right Leg): Begin in quadruped position with hands below shoulders and knees below hips. Lift and straighten the left arm and right leg to the back level. Hold for a three-count, bring elbow and knee together, then extend again.
35	Rotary Stability (Right Arm/Left Leg): Begin in quadruped position with hands below shoulders and knees below hips. Lift and straighten the right arm and left leg to the back level. Hold for a three-count, bring elbow and knee together, then extend again.
36	Step Down (Left): Stand with the left foot on a mid-shank–height box. Let the right foot hang off with toes up. With hands in a "prayer" position, squat until the right heel taps the ground, then return to standing.
37	Step Down (Right): Stand with the right foot on a mid-shank–height box. Let the left foot hang off with toes up. With hands in a "prayer" position, squat until the left heel taps the ground, then return to standing.
38	T-Balance (Left): Stand with hands in "prayer" position. Lift the right leg so the thigh is parallel to the ground. Hold, then lean forward, straightening arms and right leg to form a "T," balancing on the left leg. Return to the start.

Continued on next page

Table 6 – continued from previous page

#	Description
39	T-Balance (Right): Stand with hands in "prayer" position. Lift the left leg so the thigh is parallel to the ground. Hold, then lean forward, straightening arms and left leg to form a "T," balancing on the right leg. Return to the start.

A.5. Baselines and Inverse-Kinematics Limitations

Baseline models (MDM, PersonaBooth, T2M-GPT) output motion in the HumanML3D format (Guo et al., 2022), representing 22-joint 3D positions in a subject-agnostic skeleton. To enable biomechanical comparison with PG-BIG, we retarget these motions to the Rajagopal musculoskeletal model via joint-position-based inverse kinematics (IK).

We map 14 HumanML3D joints (hips, knees, ankles, metatarsals, shoulders, elbows, wrists) to Rajagopal nodes and solve, for each frame t:

$$\mathbf{q}_t = \arg\min_{\mathbf{q}} \sum_{j \in \mathcal{J}} \|\mathbf{p}_{j,t}^{\text{HML3D}} - \mathbf{f}_j(\mathbf{q})\|_2^2, \tag{4}$$

where $\mathbf{p}_{j,t}^{\text{HML3D}}$ are target joint positions and $\mathbf{f}_j(\mathbf{q})$ computes joint locations given configuration $\mathbf{q}$. We apply body scaling, line search, and temporal warm-starting for smoother solutions.

Cross-format retargeting introduces key limitations, however. HumanML3D joints lack anatomical precision, leading to misaligned hip, shoulder, and spine configurations. The sparse 14-joint mapping underconstrains many Rajagopal degrees of freedom, forcing the IK solver to infer missing details. Baseline motions also contain non-physical contacts—such as sliding, penetration, and hovering—that propagate as discontinuities in IK trajectories.

These issues manifest in higher foot contact errors, jerk, and energy variance (Table 2). In contrast, PG-BIG generates motion directly in anatomically grounded joint space, eliminating the lossy retargeting step and yielding biomechanically consistent results by design.

A.6. Subject Profiler Architecture

To effectively create profile tokens that embed motion, bio, and metadata about a profile, we use our subject profiler architecture. As shown in Figure 5, the VQ-VAE encoder first encodes a subject-specific motion sequence, and attention pooling aggregates the resulting latent features to produce a weighted representation of motion information. We then individually project, normalize, and pass this motion vector, along with subject biodata (age, sex, mass, height, sport, level, and injury history) and metadata (joint velocities, center-of-mass velocity, and accelerations), through GELU activations. We concatenate the processed embeddings, then project and normalize them to yield the subject-specific profile token.

A.7. Loss Surrogates for Center of Mass and Muscle Activation

Estimating muscle activations requires solving inverse dynamics, which is computationally expensive and non-differentiable. We address this by training a sequence-to-sequence surrogate model Maheshwari (2025) that efficiently estimates muscle activations and COM dynamics directly from joint kinematics. This surrogate eliminates the computational burden of inverse dynamics solvers,

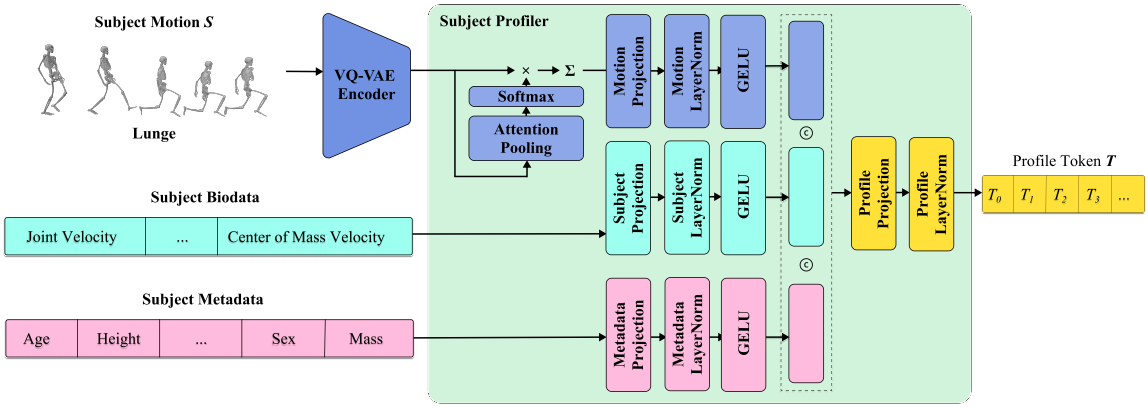

Figure 5: Subject Profiler Architecture: A subject-specific motion sequence is encoded by the VQ-VAE encoder, then, through attention pooling, a motion vector is produced. This motion vector, with subject biodata and metadata, is individually projected, normalized, and passed through GELU activations. The processed embeddings are concatenated, projected, and normalized, yielding the subject-specific Profile Token.

ensures anatomical plausibility and biomechanical consistency, and enables gradient-based optimization during inference.

Specifically, we train a transformer-based network to map the predicted joint angles $\{q_t\}_{t=1}^{T}$ to their corresponding COM coordinates $\{\mathbf{c}_t\}_{t=1}^{T}$ and muscle activations $\{\mathbf{a}_t\}_{t=1}^{T}$.

A.8. Surrogate Model Validation

Because the surrogate replaces OpenSim's non-differentiable static-optimization solve during inference-time guidance, we summarize its training recipe, report validation against OpenSim-derived targets, and relate accuracy to gradient-based refinement.

The surrogate is a transformer encoder that maps joint kinematics to muscle activations and COM quantities (input dim. 33 per frame, output dim. 80 activations plus COM terms; 3 layers, 3 attention heads, feedforward dim. 128). It was trained with dropout rate 0.1 to limit overfitting, initial learning rate 2×10^{-4}, and 1000 epochs using the Adam optimizer Kingma and Ba (2014) with batch size 128. The learning rate was reduced by a factor of 0.5 whenever the validation loss plateaued for 10 consecutive epochs. The training objective was smooth L1 (Huber) loss between predicted and ground-truth muscle activations from OpenSim. Gradient clipping with maximum norm 1.0 and weight decay 10^{-5} stabilized optimization. Model checkpoints and inference scripts are provided in the project repository under `PG-BIG/`.

Evaluation. Targets are obtained by static optimization (OpenSimAD) on retargeted Rajagopal trajectories Rajagopal et al. (2016a). We use a 70%/10%/20% subject-level split; the test set has 37 subjects and 1004 sequences spanning all movements. For each muscle we compute coefficient of determination (R^2) and RMSE between surrogate and OpenSim activations (unitless in $[0, 1]$) and aggregate over the test set. Table 7 lists the mean over all 80 muscles and representative lower-limb muscles heavily loaded in squat-like and propulsive tasks. On the same split, aggregate COM velocity and acceleration predictions yield mean R^2 of 0.94 and 0.91 with RMSE $0.045\,\text{m/s}$ and $0.079\,\text{m/s}^2$, respectively. Mean peak-to-peak activation phase lag vs. OpenSim is $7.2 \pm 3.1\,\text{ms}$ at

Table 7: Surrogate vs. OpenSim on held-out test data: overall mean over all muscles and key muscles of interest (R^2 and RMSE of activations).

Muscle / scope	R^2 (↑)	**RMSE** (↓)
All muscles (mean, N=80)	0.878	0.067
Soleus (left)	0.839	0.117
Soleus (right)	0.814	0.119
Vastus intermedius (left)	0.865	0.068
Vastus intermedius (right)	0.876	0.063
Vastus lateralis (left)	0.873	0.112
Vastus lateralis (right)	0.899	0.134
Vastus medialis (left)	0.827	0.089
Vastus medialis (right)	0.839	0.092

60 Hz. On one NVIDIA A40 GPU, OpenSim static optimization requires 42–58 s per 3 s clip versus $<$8 ms for the surrogate, which makes iterative guidance practical.

Implication for guidance. Biomechanical guidance applies percentile-based bounds to surrogate-predicted activations and COM terms. The R^2 and RMSE in Table 7 show that the surrogate approximates OpenSim well on average and on biomechanically critical muscles, so refinement gradients track the physics-based reference rather than a systematically biased proxy; combined with bounded phase lag and fast evaluation, this supports using the surrogate inside the inference-time loop instead of repeated OpenSim solves.

Appendix B. Additional Results

B.1. Profile Classifiers

To verify that profile conditioning effectively encodes subject-specific and demographic information, we analyze VQ-VAE latent embeddings extracted from generated motion sequences. Each motion is encoded and temporally pooled via attention pooling to produce a fixed-dimensional embedding $\mathbf{z} \in \mathbb{R}^{512}$. We then evaluate two supervised tasks on these embeddings: (i) subject identity classification among 183 athletes, and (ii) continuous age regression in years.

We assess multiple supervised models, including logistic regression, support vector machines, random forests, gradient boosting, and multi-layer perceptrons, to ensure robust results across architectures. Since the evaluation is performed on generated motions, the goal is to determine whether the embeddings retain information consistent with the intended profile attributes, rather than to generalize to held-out training data.

For identity classification, we report accuracy, macro F1 score, and top-3 accuracy. For age regression, we report mean absolute error (MAE), root mean squared error (RMSE), and R^2. The results show that generated motions preserve individual kinematic signatures: the identity classifier achieves 94% accuracy, while age regression yields a mean squared error of 2.4 years. These findings confirm that (i) PG-BIG embeddings reflect subject-specific characteristics, (ii) generated motions remain individually attributable rather than collapsing to a generic style, and (iii) synthetic

profile manipulations (e.g., age or training variations) produce consistent, detectable changes in the latent space aligned with metadata.

Table 8: Evaluation results for Subject Identification. XGBoost achieved the highest classification accuracy among all models.

Model	Error	MSE	RMSE	Accuracy	Expl. Var.
Logistic Regression	0.162	6.412	2.533	0.812	0.810
SVM (Linear)	0.143	5.872	2.422	0.835	0.833
SVC (Kernel)	0.129	5.013	2.239	0.861	0.858
Random Forest	0.094	3.014	1.736	0.904	0.902
k-NN	0.118	4.284	2.070	0.876	0.873
MLP	0.083	2.751	1.658	0.917	0.915
XGBoost	**0.062**	**2.001**	**1.414**	**0.940**	**0.939**

Table 9: Evaluation results for Age Prediction. XGBoost demonstrated the lowest error and highest R^2 among all regression models.

Model	MAE	MSE	RMSE	R^2	Expl. Var.
Gradient Boosting	2.774	14.223	3.771	0.090	0.091
Random Forest	2.795	14.374	3.791	0.081	0.081
Linear Regression	2.851	14.802	3.847	0.053	0.054
Ridge Regression	2.851	14.801	3.847	0.053	0.054
MLP	3.876	25.826	5.082	-0.652	-0.651
XGBoost	**1.129**	**2.034**	**1.427**	**0.938**	**0.937**

B.2. Quantitative and Qualitative Results

Extended Quantitative Metrics. These results supplement Figure 3 and Figure 4 in the main text and provide a finer-grained view of biomechanical stability and stylistic fidelity across models.

Additional Visualizations. Figure 6 presents ablation results for mobility (lunge) and stability (T-balance) exercises, while Figure 3 compares PG-BIG to data-driven baselines (MDM, PersonaBooth, and T2M-GPT).

Figure 6 visualizes these trade-offs qualitatively. The profiling variant successfully reproduces subject-specific stylistic traits but displays physical instabilities, including foot penetration, floating during contact phases, and incorrect joint placement (particularly at the ankles and knees). These artifacts arise from the absence of biomechanical constraints during generation, allowing the model to prioritize stylistic fidelity over physical plausibility.

The Guidance variant produces biomechanically sound motion with stable ground contact, smooth joint trajectories, and physiologically realistic center-of-mass control. However, the generated movements are generic and fail to capture individual expression: all subjects perform lunges and balance poses with similar depth, timing, and postural strategies, losing the personalized movement signatures present in the reference data.

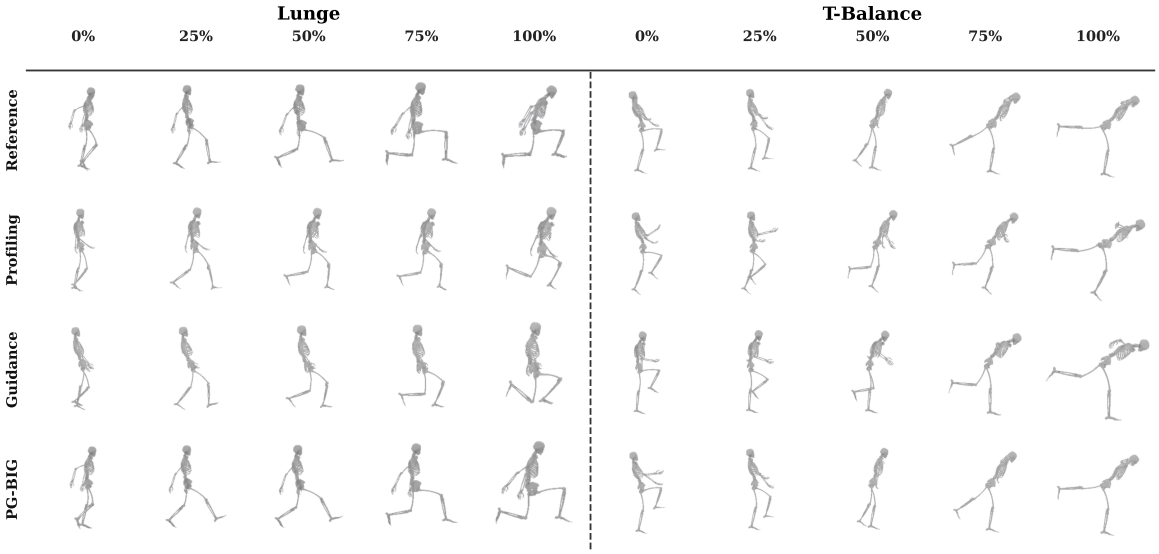

Figure 6: Ablation analysis of PG-BIG components across two representative tasks: forward lunge and T-balance. We compare reference motion with outputs from the profile-only, guidance-only, and full PG-BIG models. The profile-only variant reproduces stylistic traits but lacks physical stability, while the guidance-only model maintains balance yet loses individual expression. The full PG-BIG framework shows both, producing biomechanically grounded movements personalized to each subject.

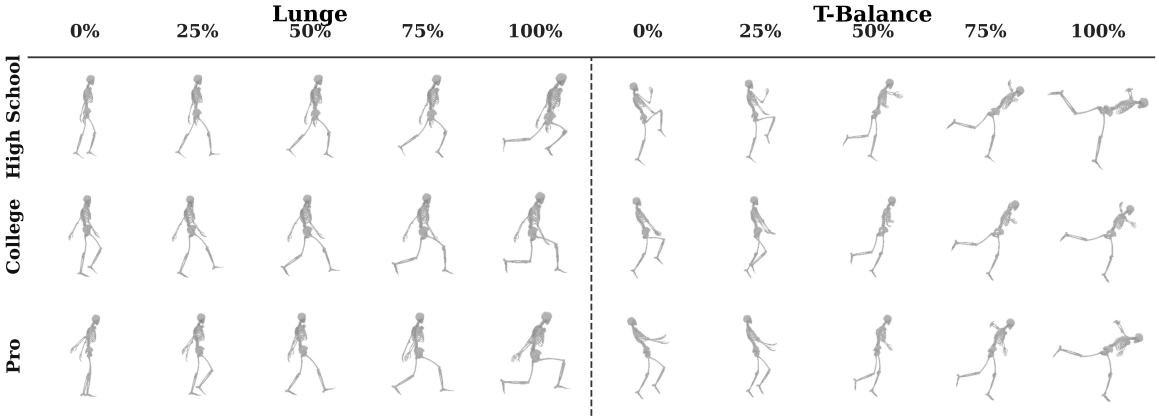

Figure 7: Synthetic training level manipulation. Decreasing training experience from professional to high school yields shallower movement execution, altered trunk strategies, and less committed weight transfers, consistent with skill-level differences in motor control.

The complete PG-BIG model integrates both strengths. It maintains the physical stability while preserving the individualized movement style. The result is motion that is both biomechanically grounded and personalized, demonstrating that profile conditioning and biomechanical guidance can work in concert to achieve high-fidelity, physically valid human motions.