Latent Dynamics Modulation Guidance for Semantic Obstacle Avoidance in Diffusion Policies

Sasidhar Alavala[1,*], Varuni Buereddy[1,*], and Ravi Prakash[1]

[1]Department of Cyber-Physical Systems, IISc Bangalore, India
[*]*Equal Contribution*

***Abstract*— Steering a pretrained diffusion policy at inference time can be thought of as a control problem in the latent space of a learned dynamics model, in which positive-condition embeddings act as attractors and negative-condition embeddings as obstacles. The standard approach across guided-diffusion methods composes the two by summing the gradients of their guidance scores with respect to the noisy action chunk. We show this composition has a gradient-cancellation failure mode, familiar from text-to-image negative prompting, when an obstacle lies on the line between the agent and a target. Treating attractors and obstacles geometrically suggests the multiplicative composition used in dynamical-system motion planning, where each obstacle contributes a full-rank modulation matrix that compresses the component of the positive flow normal to the obstacle and stretches the tangent. The resulting Latent Dynamics Modulation Guidance (LDMG) is cancellation-free, preserves at least the tangential component of the positive flow, and composes over multiple obstacles as a matrix product. On a 2D Pymunk benchmark with three obstacles, LDMG reaches the target on 92% of rollouts on a layout where additive composition reaches only 38%, and on 100% of rollouts on a multi-obstacle diagonal layout. On a CALVIN multi-objective task with door manipulation as the positive condition and switch-on plus button-on as negative conditions, LDMG attains the highest task-completion rate among the compared methods (26.0%).**

I. INTRODUCTION

Steering a pretrained diffusion policy [1] at inference time without retraining requires injecting a directional preference into the denoising loop. When that preference is expressed as semantic conditions, visual or perceptual goals to reach and outcomes to avoid, the conditions acquire a geometric role in the latent space: *positives* behave as *attractors* that the predicted future state should approach, *negatives* behave as *obstacles*. Phrased this way, test-time guidance becomes a control problem, analogous to classical obstacle avoidance, where a flow is shaped to move towards attractors while avoiding obstacles. The key difference is that geometry is defined in a learned latent space, where attractors and obstacles are defined by semantic representations rather than physical bodies.

The recent diffusion policy steering literature has converged on an *additive* composition of the directional preferences. One such method is DynaGuide [2], which scores positive and negative conditions through a log-sum-exp over latent distances and adds the resulting gradients into the denoiser inside a DDIM loop [3]. Inference-time policy steering [4], classifier-free guidance [5], and composable diffusion [6] all belong to the same family, they compose attractor and obstacle scores by summation and inject the gradient of the resulting scalar into the denoising step. The composition is convenient because the resulting objective is a single scalar that can be differentiated with one backward pass. Its weakness is that a sum of two gradients can vanish even when neither is zero.

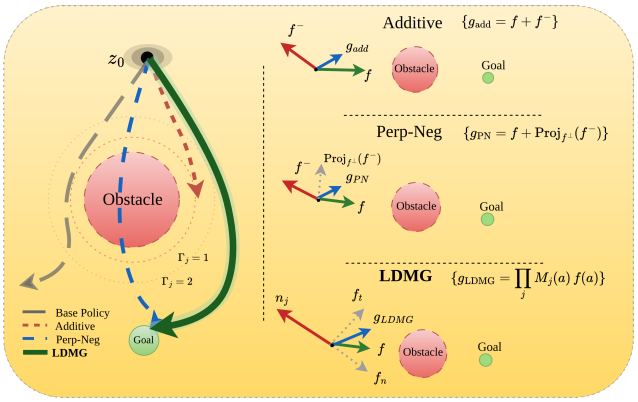

Fig. 1: Geometric view of test-time guidance. (Left) In the dynamics latent space, positive-condition embeddings act as attractors and negative-condition embeddings as obstacles. The base policy wanders, additive guidance stalls in the cancellation regime, Perp-Neg sidesteps cancellation but does not bend around the obstacle, and LDMG curves around it along the level-set tangent. (Right) Vector geometry of the three guidance composition rules on a shared scene. Additive collapses when f^- and f are nearly anti-parallel. Perp-Neg subtracts the projection of f^- onto $f^\perp$, using f as the decomposition axis. LDMG decomposes f along the obstacle normal n_j and applies eigenvalue scaling to compress the inflow and to stretch the tangent.

When the predicted state approaches an obstacle while moving toward the goal, the positive gradient pulls the action toward the target, while the negative gradient pushes it away from the obstacle along a similar direction. The two gradients become nearly anti-parallel, and their sum vanishes. This leads to degenerate guidance. The same phenomenon has been studied in text-to-image diffusion [7], [8], [9]. The closest fix in that literature is Perpendicular-Negative (Perp-Neg) [10], which projects the negative score onto the com-

ponent perpendicular to the positive flow. Perp-Neg sidesteps cancellation, but it is a single-axis correction defined by only the positive flow direction, and it does not use the geometry of the obstacles or their level sets.

The main problem is that the additive composition is not well-suited to combine attractors and obstacles in a vector field. In contrast, the dynamical-system motion-planning literature [11], [12], [13] formulates their interaction through *multiplicative* composition. This is achieved by reshaping the attractive flow $f(x)$ at each obstacle by a full-rank modulation matrix $M(x)$ using eigenvalues that suppress motion towards the obstacle while preserving tangential motion around it. We lift this construction into the latent space. Here, obstacles correspond to semantic embeddings rather than physical bodies, distances are defined by soft-min latent distances rather than Euclidean ones, and the flow being modulated is a gradient on action chunks rather than an acceleration in configuration space. We build on this formulation to obtain *Latent Dynamics Modulation Guidance* (LDMG) to steer diffusion policy. We show that LDMG avoids gradient cancellation and demonstrate its effectiveness on Pymunk and CALVIN tasks [14], where it improves robustness under competing objectives.

II. METHOD

A. Preliminaries

Diffusion guidance methods steer a pretrained action denoiser $\epsilon_\phi(a^k, o_t)$ that maps a noisy action chunk a^k and observation o_t to a noise estimate using a gradient of some guidance metric. In DynaGuide [2], a frozen observation encoder ϕ and a latent dynamics model h_θ together predict a future latent $\hat{z}_{t+H}(a) := h_\theta(z_t, a) = h_\theta(\phi(o_t), a)$ from a clean action chunk a and the current latent state z_t. Given positive conditions $\{o_i^+\}$ and negative conditions $\{o_j^-\}$, the guidance metric is

$$\begin{aligned} d(a) \;=\; & \log \sum_i e^{-\|\phi(o_i^+) - \hat{z}_{t+H}(a)\|^2/\sigma} \\ & - \log \sum_j e^{-\|\phi(o_j^-) - \hat{z}_{t+H}(a)\|^2/\sigma}, \end{aligned} \tag{1}$$

and the guidance injection into the DDIM loop is

$$\hat{\epsilon}(a^k, o_t) \;=\; \epsilon_\phi(a^k, o_t) \;-\; s\sqrt{1-\bar{\alpha}_k}\, \nabla_{a^k} d(a), \tag{2}$$

with guidance scale s and DDIM coefficient $\bar{\alpha}_k$. Writing $d = d^+ - d^-$, the two components d^+, d^- enter the sum with opposite signs: $\nabla_{a^k} d \;=\; \nabla_{a^k} d^+ \;-\; \nabla_{a^k} d^-$. When the two gradients become nearly parallel in action-chunk space, $\nabla_{a^k} d^+ \;\approx\; \nabla_{a^k} d^-$, and their difference collapses, $\|\nabla_{a^k} d\| \to 0$. Equivalently, in the force view of Fig. 1 with $f = \nabla_{a^k} d^+$ and $f^- = -\nabla_{a^k} d^-$, the attractive and repulsive forces become anti-parallel and their sum $f + f^-$ vanishes.

B. Latent Dynamics Modulation Guidance

We keep the latent dynamics model and condition embeddings of [2] unchanged and modify the guidance metric. Let

$$f(a) \;:=\; \nabla_{a^k} d^+ = \nabla_{a^k} \log \sum_i e^{-\|\phi(o_i^+) - \hat{z}_{t+H}(a)\|^2/\sigma}, \tag{3}$$

be the nominal attractive flow in action-chunk space, the positive half of (1), used on its own. We likewise write $f^-(a) := -\nabla_{a^k} d^-(a)$ for the repulsive flow, so the additive rule is $g_{\text{add}} = f + f^-$. Each obstacle j is represented by a set of negative conditions $\text{avoid}_\text{j} = \{o_k^-\}$ corresponding to undesired outcomes of same type. For each obstacle j define a soft-min latent distance

$$d_j(a) \;:=\; -\log \sum_{k \in \text{avoid}_j} e^{-\|\phi(o_k^-) - \hat{z}_{t+H}(a)\|^2/\sigma}, \tag{4}$$

which is small (≈ 0) when $\hat{z}_{t+H}(a)$ is close to the avoid embeddings and increases as it moves away. The level-set distance Γ_j analogous to that of [11] is

$$\Gamma_j(a) \;:=\; 1 + d_j(a)^{1/\rho}, \tag{5}$$

with $\rho > 0$ controlling how quickly the modulation changes with distance from the obstacle. Thus $\Gamma_j \in [1, \infty)$, with $\Gamma_j = 1$ on the obstacle's level set and $\Gamma_j \to \infty$ far from every avoid embedding.

Let $n_j(a)$ be the unit vector along $\nabla_{a^k} d_j(a)$, the direction along which the policy moves away from obstacle j. Let $T_j(a)$ be an orthonormal basis for the subspace orthogonal to $n_j(a)$, capturing directions tangent to the obstacle. For each obstacle, eigenbasis and eigenvalue matrix are

$$E_j(a) = [\, n_j(a)\; T_j(a)\,], \quad D_j(a) = \text{diag}(\lambda_j^n, \lambda_j^e, \ldots, \lambda_j^e). \tag{6}$$

$$\lambda_j^n \;=\; 1 - \tfrac{1}{\Gamma_j(a)}, \qquad \lambda_j^e \;=\; 1 + \tfrac{1}{\Gamma_j(a)}, \tag{7}$$

and the per-obstacle modulation matrix is $M_j(a) = E_j(a)\, D_j(a)\, E_j(a)^{-1}$. For multiple obstacles, the matrices compose as a product [11]:

$$g_{\text{LDMG}}(a) \;=\; \Big(\prod_{j \in \text{avoids}} M_j(a) \Big) f(a). \tag{8}$$

The guidance is injected into the DDIM denoising step as in (2), replacing the additive gradient $\nabla_{a^k} d(a)$ with the modulated guidance field $g_{\text{LDMG}}(a)$.

Comparison to Perp-Neg. Perp-Neg [10] replaces additive composition with $\nabla_{a^k} d^+ - \text{Proj}_{(\nabla_{a^k} d^+)^\perp}(\nabla_{a^k} d^-)$, avoiding cancellation in the anti-parallel case by removing the component of the negative gradient orthogonal to the positive flow. However, it remains a single-axis correction defined by the goal direction and does not explicitly incorporate obstacle geometry. In contrast, LDMG is defined with respect to each obstacle's normal n_j, modulates both normal and tangential components through Γ_j, and composes naturally across multiple obstacles via matrix products.

C. Properties of LDMG

Cancellation-freeness. For all a such that every $\Gamma_j(a) > 1$, each modulation matrix $M_j(a)$ is full rank, and hence $g_{\text{LDMG}}(a) \;=\; 0 \iff f(a) \;=\; 0$.

Norm bound. Writing $f = f_j^{\parallel} + f_j^{\perp}$, components parallel and orthogonal to n_j, the modulation gives $M_j f = \lambda_j^n f_j^{\parallel} + \lambda_j^e f_j^{\perp}$ with $0 < \lambda_j^n < 1$ and $\lambda_j^e \geq 1$. Hence $\|M_j f\| \geq \|f_j^{\perp}\|$, so the guidance magnitude is bounded below by the component tangent to the obstacle. The bound is stated for

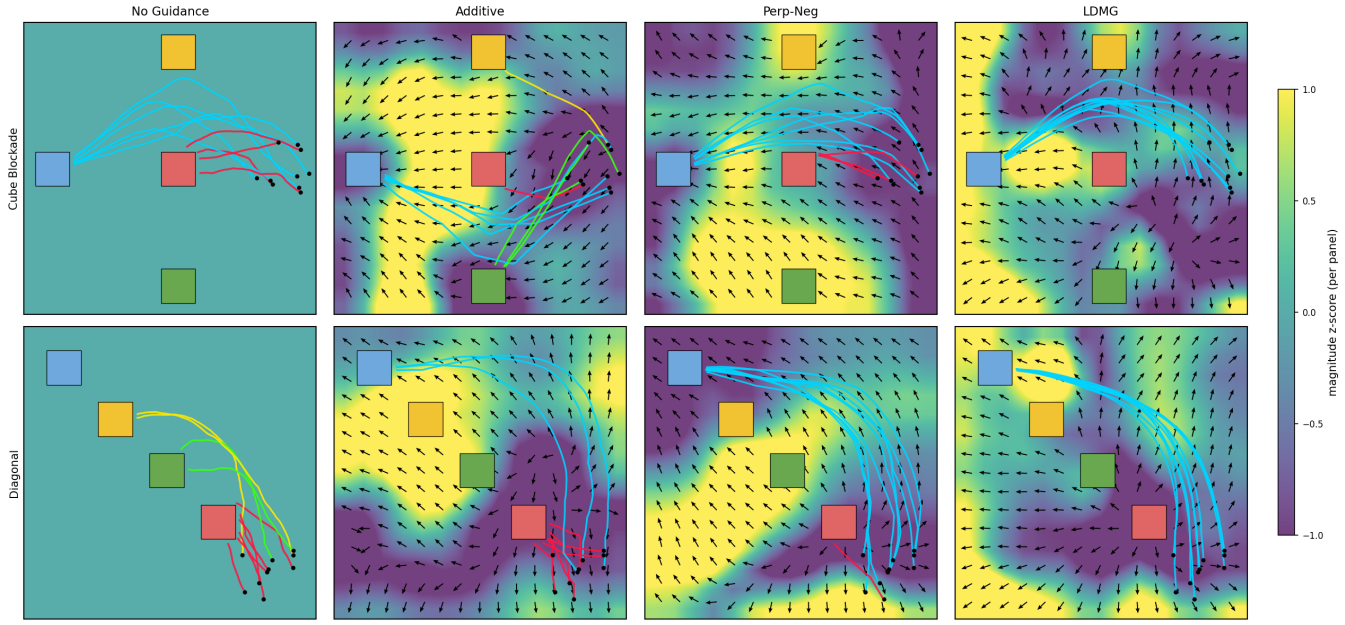

Fig. 2: Guidance vector fields and rollout trajectories. *Additive* fails to provide guidance during the initial rollout due to the cancellation of gradients. Under *Perp-Neg*, the projection along $f^{\perp}$ does not bend trajectories around red, so red collisions happen. Under *LDMG*, the modulation product produces tangent guidance flow around each obstacle and trajectories curve cleanly toward blue on both layouts.

a single M_j and for the product $\prod_j M_j$ it applies factor by factor.

Recovery of positive-only guidance. As $\Gamma_j(a) \to \infty$, $D_j(a) \to I$ and $M_j(a) \to I$, so $g_{\mathrm{LDMG}}(a) \to f(a)$.

Relation to prior formulations. Alternative modulation schemes exist in dynamical systems [13], but we adopt the eigenvalue form of [11] due to its simplicity and suitability for convex latent obstacles.

III. RESULTS

We compare three guidance compositions: *Additive*, *Perp-Neg* (rank-1 projection onto $f^{\perp}$), and *LDMG*. The base policy, dynamics model, condition embeddings, guidance scale $s = 1$, $\rho = 0.5$ in Γ_j and DDIM schedule are identical across methods.

A. Pymunk Experiment

We follow the toy setup of [2] in a Pymunk 2D scene at 128×128 pixels with four colored cubes and an agent moving under 2D velocity control. The base policy is a DDIM diffusion policy trained on 100k demonstrations that visit all four cubes uniformly. The classifier is a noise-augmented 4-way softmax head whose log probabilities are the latent distances entering (3) and (4). The DDIM loop runs for 10 steps with $M{=}4$ stochastic sub-samples per step. Blue cube is the goal and red, green, and yellow cubes are simultaneously obstacles, and LDMG composes three modulation matrices in (8). We experiment with Cube Blockade, and Diagonal layouts.

Table I reports per-cube arrival rates for both layouts. Under Additive composition, the guidance gradient cancels in the early portion of the rollout where the positive and negative flows run anti-parallel, and recovers strength only later, near the goal away from the obstacles. The Additive panel of Fig. 2 shows this directly. The two strong yellow lobes of the heatmap sit adjacent to the blue cube rather than along the path from the agent's start, so the guidance has no useful magnitude during the part of the rollout in which it would need to act. Perp-Neg projects the summed negative flow onto the component perpendicular to the positive flow, sidestepping the cancellation when the obstacles are spread along the path so that each one's contribution is partially perpendicular to the positive flow at the relevant moment. It sidesteps cancellation but does not use the obstacles' level-set geometry, so it has no mechanism to bend the trajectory around red when the positive flow points straight at it.

TABLE I: Per-cube arrival rate reported as Cube Blockade / Diagonal (50 runs per cell).

Method	Blue% ↑	Red% ↓	Green% ↓	Yellow% ↓	None%
No guidance	34 / 8	44 / 38	22 / 10	0 / 42	0 / 2
Additive	38 / 40	2 / 58	44 / 0	16 / 2	0 / 0
Perp-Neg	54 / 96	44 / 2	0 / 0	2 / 2	0 / 0
LDMG	**92 / 100**	4 / 0	2 / 0	2 / 0	0 / 0

LDMG reaches blue 92% on Cube Blockade and 100% on Diagonal. The modulation matrix product $M_{\mathrm{green}} M_{\mathrm{red}} M_{\mathrm{yellow}}$ shrinks the inflow component of the positive flow as the predicted future approaches each obstacle's level set, with the closest obstacle dominating the modulation along the trajectory. The geometric difference from the two baselines is visible in the arrow fields of Fig. 2.

TABLE II: Behavior distribution on the CALVIN multi-objective task over 50 runs per method. None % denotes roll-outs that exhausted the 80-step horizon without completing any recognized task.

Method	Move-Door% ↑	Button-On% ↓	Switch-On% ↓	None %↓
No guidance	11.3 ± 4.6	4.0 ± 2.0	4.7 ± 5.0	80.0 ± 8.7
Additive	13.3 ± 3.1	4.7 ± 1.2	3.3 ± 4.2	78.7 ± 2.3
Perp-Neg	17.3 ± 6.1	6.7 ± 2.3	2.7 ± 2.3	73.3 ± 4.6
LDMG	26.0 ± 2.0	5.0 ± 4.2	6.0 ± 0.0	63 ± 5.0

Additive and Perp-Neg both produce guidance arrows that lie along the positive-flow direction, which is also the direction the rollout trajectory tries to follow, so the arrows in those panels run roughly tangent to the trajectories everywhere in the scene. LDMG instead decomposes its guidance along each obstacle normal n_j. Far from every obstacle, where every Γ_j is large, the modulation matrices reduce to the identity and LDMG arrows align with the positive flow as seen near the goal. Near an obstacle, the inflow component along n_j is shrunk and the tangent component stretched, so the arrows curve around the obstacle along its level-set shell rather than driving through it. The trajectories inherit this structure, curving around each obstacle and straightening out as they approach blue. The remaining 8% of failures on Cube Blockade split across the three obstacles with no dominant mode, and we attribute these to dynamics-model error rather than to a structural limitation of the construction.

B. CALVIN Experiment

To test LDMG beyond the classifier-based Pymunk setting, we evaluated on the CALVIN benchmark [14], using the released DynaGuide policy and dynamics-model checkpoints. The environment is a 7-DoF Franka Research 3 manipulating a tabletop scene containing a drawer, a sliding door, a switch, and a button. Guidance gradients are computed through a DINOv2 latents [15] based dynamics model. We construct a multi-objective evaluation in which the positive condition is door manipulation, represented by 20 guidance images spanning both door-open and door-close (ten of each), and the negative condition is the joint avoidance of switch-on and button-on, represented by 20 guidance images split equally between the two. The guidance set of one negative condition forms a single avoid cluster in (4). Each method is evaluated over 50 rollouts from a fixed initial robot configuration, with a horizon of 80 environment steps and $\rho = 0.5$.

Table II reports the behavior distribution, each rate is the mean over 3 random seeds and we report $\pm$ standard deviation. LDMG gives the highest positive behavior rate (i.e., Move-Door). The button-on rate under LDMG is close to the unguided baseline, which shows that the avoidance term reduces motion toward the button. The switch-on rate, however, increases rather than decreases. This is not a failure of the avoidance term but a result of the tabletop layout. The switch is placed close to the door, and the end-effector must pass near the switch on its way to the door. Because LDMG drives the policy toward the door more often than the other methods, a small number of rollouts flip the switch in passing. LDMG also attains the lowest non-completion rate (None% $= 63$ vs. 78.7 for Additive). Fig. 3 visualizes representative end-effector trajectories projected onto the tabletop.

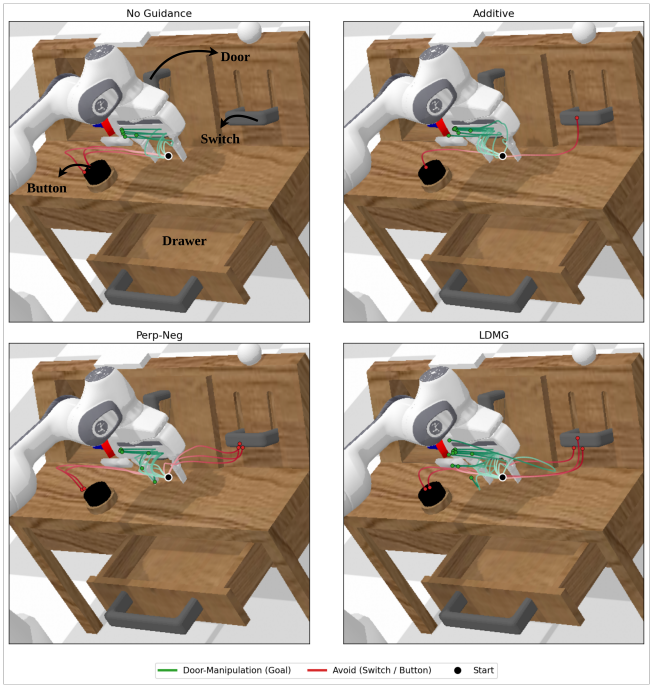

Fig. 3: End-effector trajectories on CALVIN over 50 runs per method, projected onto the tabletop.

IV. CONCLUSION

We framed test-time steering of a diffusion policy as a control problem in the latent space of a learned dynamics model, where positive condition embeddings act as attractors and negative embeddings as obstacles. The additive composition of attractor and negative gradients used by current methods admits a gradient-cancellation failure mode in this geometry, while the multiplicative composition used in dynamical-system motion planning does not, giving a guidance rule that is cancellation-free and composes across multiple obstacles. LDMG improves target reach in Pymunk and task completion in CALVIN under a fixed-base policy and dynamics model.

Several limitations follow from where the construction is applied rather than from the construction itself. Because the modulation acts on the predicted future latent rather than on a sensed current state, LDMG is a predictive guidance planner rather than a reactive one, and the guidance is only as reliable as the dynamics model and encoder it depends on. Where the training distribution is sparse or where two semantic categories are entangled in the latent space, the level set Γ_j and the obstacle normal n_j can be computed against a geometry that does not match the true future, and no guidance rule built on top of the latent can correct for this. The CALVIN switch-on task rate under LDMG is the clearest empirical sign of these limits. Our evaluation covers

a controlled 2D benchmark and one multi-objective CALVIN task, broader semantic obstacle avoidance settings, and a systematic study of how dynamics model and encoder error perturb the obstacle normals n_j and modulation matrices M_j, are left for future work.

REFERENCES

[1] C. Chi, Z. Xu, S. Feng, E. Cousineau, Y. Du, B. Burchfiel, R. Tedrake, and S. Song, "Diffusion Policy: Visuomotor Policy Learning via Action Diffusion," in *The International Journal of Robotics Research*, 2024.

[2] M. Du and S. Song, "Dynaguide: Steering Diffusion Policies with Active Dynamic Guidance," in *Proceedings of the 39th Conference on Neural Information Processing Systems (NeurIPS)*, 2025.

[3] J. Song, C. Meng, and S. Ermon, "Denoising Diffusion Implicit Models," in *International Conference on Learning Representations*, 2021.

[4] Y. Wang, L. Wang, Y. Du, B. Sundaralingam, X. Yang, Y.-W. Chao, C. Pérez-D'Arpino, D. Fox, and J. Shah, "Inference-time policy steering through human interactions," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2025, pp. 15 626–15 633.

[5] J. Ho and T. Salimans, "Classifier-Free Diffusion Guidance," in *arXiv preprint arXiv:2207.12598*, 2022.

[6] N. Liu, S. Li, Y. Du, A. Torralba, and J. B. Tenenbaum, "Compositional Visual Generation with Composable Diffusion Models," in *European Conference on Computer Vision*. Springer, 2022, pp. 423–439.

[7] Y. Ban, R. Wang, T. Zhou, M. Cheng, B. Gong, and C.-J. Hsieh, "Understanding the Impact of Negative Prompts: When and How do They Take Effect?" in *European Conference on Computer Vision*. Springer, 2024, pp. 190–206.

[8] F. Koulischer, J. Deleu, G. Raya, T. Demeester, and L. Ambrogioni, "Dynamic Negative Guidance of Diffusion Models," in *The Thirteenth International Conference on Learning Representations*, 2025.

[9] J. Chang, H. Chung, and J. C. Ye, "Contrastive CFG: Improving CFG in Diffusion Models by Contrasting Positive and Negative Concepts," in *arXiv preprint arXiv:2411.17077*, 2024.

[10] M. Armandpour, H. Zheng, A. Sadeghian, A. Sadeghian, and M. Zhou, "Re-imagine the Negative Prompt Algorithm: Transform 2D Diffusion into 3D, alleviate Janus problem and Beyond," in *arXiv preprint arXiv:2304.04968*, 2023.

[11] S. M. Khansari-Zadeh and A. Billard, "A dynamical system approach to realtime obstacle avoidance," in *Autonomous Robots*, vol. 32, no. 4. Springer, 2012, pp. 433–454.

[12] L. Huber, A. Billard, and J.-J. Slotine, "Avoidance of convex and concave obstacles with convergence ensured through contraction," in *IEEE Robotics and Automation Letters*, vol. 4, no. 2. IEEE, 2019, pp. 1462–1469.

[13] L. Huber, J.-J. Slotine, and A. Billard, "Avoidance of concave obstacles through rotation of nonlinear dynamics," in *IEEE Transactions on Robotics*, vol. 40. IEEE, 2023, pp. 1983–2002.

[14] O. Mees, L. Hermann, E. Rosete-Beas, and W. Burgard, "CALVIN: A Benchmark for Language-Conditioned Policy Learning for Long-Horizon Robot Manipulation Tasks," *IEEE Robotics and Automation Letters (RA-L)*, vol. 7, no. 3, pp. 7327–7334, 2022.

[15] M. Oquab, T. Darcet, T. Moutakanni, H. V. Vo, M. Szafraniec, V. Khalidov, P. Fernandez, D. HAZIZA, F. Massa, A. El-Nouby, M. Assran, N. Ballas, W. Galuba, R. Howes, P.-Y. Huang, S.-W. Li, I. Misra, M. Rabbat, V. Sharma, G. Synnaeve, H. Xu, H. Jegou, J. Mairal, P. Labatut, A. Joulin, and P. Bojanowski, "DINOv2: Learning Robust Visual Features without Supervision," in *Transactions on Machine Learning Research*, 2024.

V. RESPONSE TO REVIEWER COMMENTS

We thank both reviewers for their feedback and for recommending acceptance. We added the statistical robustness requested by Reviewer 1 by report CALVIN results over multiple seeds with mean and standard deviation in Table II.

The remaining concerns are extensions beyond the scope of this workshop paper and we answer them as future work. To the breadth concern raised by both reviewers, we will evaluate LDMG on a broader set of semantic obstacle avoidance tasks beyond the toy benchmark and the single CALVIN scenario. In response to the concern about sample-size raised by Reviewer 1, we will scale the number of rollouts by one to two orders of magnitude to further tighten the statistical claims. To the latent-geometry concern raised by Reviewer 2 we will run a systematic study of how dynamics model and encoder error perturb the obstacle normals n_j and the modulation matrices M_j.