Stabilizing GANs' Training with Brownian Motion Controller

Tianjiao Luo [1] **Ziyu Zhu** [1] **Jianfei Chen** [1] **Jun Zhu** [1 2]

Abstract

The training process of generative adversarial networks (GANs) is unstable and does not converge globally. In this paper, we examine the stability of GANs from the perspective of control theory and propose a universal higher-order noise-based controller called Brownian Motion Controller (BMC). Starting with the prototypical case of Dirac-GANs, we design a BMC to retrieve precisely the same but reachable optimal equilibrium. We theoretically prove that the training process of DiracGANs-BMC is *globally exponential stable* and derive bounds on the rate of convergence. Then we extend our BMC to normal GANs and provide implementation instructions on GANs-BMC. Our experiments show that our GANs-BMC effectively stabilizes GANs' training under StyleGANv2-ada frameworks with a faster rate of convergence, a smaller range of oscillation, and better performance in terms of FID score.

1. Introduction

Generative Adversarial Networks (GANs) (Goodfellow et al., 2014) are popular deep generative models. Given a multi-dimensional input dataset with an unknown distribution $p(x)$, GANs can obtain an estimated $p_G(x)$ and produce new samples that are almost indistinguishable from the training samples. GANs have many application on image generation (Brock et al., 2018), representation learning (Radford et al., 2015b; Salimans et al., 2016; Mathieu et al., 2016), and image to image translation (Zhu et al., 2017).

GANs consist of two neural networks: a *generator* and a *discriminator*. The generator creates new samples that resemble those from the input dataset as closely as possible. The discriminator, on the other hand, aims to distinguish

Table 1: Summarization of the converging behaviors under Dirac-GANs' setting on various GANs from existing theoretical stability analysis (Mescheder et al., 2018; Fedus et al., 2018; Farnia & Ozdaglar, 2020). Notice that $\checkmark^*$ indicates observations from our experiments.

	Unstable	Local Stability	Global Stability
Original (Goodfellow et al., 2014)	$\checkmark$		
WGAN (Arjovsky & Bottou, 2017)	$\checkmark$		
WGAN-GP (Gulrajani et al., 2017)	$\checkmark$		
DRAGAN (Kodali et al., 2017)	$\checkmark$		
Instance Noise (Sønderby et al., 2016)		$\checkmark$	
CLC (Xu et al., 2020)		$\checkmark$	
TTUR (Heusel et al., 2017a)		$\checkmark$	
StyleGAN (Karras et al., 2019; 2020a; 2021)	$\checkmark^*$		
Our Method		$\checkmark$	$\checkmark$

the (counterfeit) samples produced by the generator from the members of the input dataset. The two networks can be modeled as a minimax problem; they compete against one another while striving to reach a *Nash-equilibrium*, an optimal solution where the generator can produce fake samples that are, from the point of view of the discriminator, in all respects indistinguishable from real ones.

Unfortunately, the process of training GANs often suffers from instabilities, preventing them from reaching the Nash equilibrium. There are many attempts for stabilizing GANs, including alternative architectures (Miyato et al., 2018; Brock et al., 2018; Karras et al., 2019), regularizations (Kodali et al., 2017; Karras et al., 2020b), alternative training algorithms (Heusel et al., 2017a; Karras et al., 2020b), and alternative objective functions (Gulrajani et al., 2017; Nowozin et al., 2016; Li et al., 2017). However, as shown by Mescheder et al. (2018); Farnia & Ozdaglar (2020), though being empirically effective, many of these methods still suffer from instability even for very simple GANs' architectures.

---

[1]Dept. of Comp. Sci. & Tech., Institute for AI, BNRist Center, Tsinghua-Bosch Joint ML Center, THBI Lab, Tsinghua University [2]Pazhou Lab (Huangpu), Guangzhou, China. Correspondence to: Jianfei Chen <jianfeic@tsinghua.edu.cn>.

One promising approach for analyzing and stabilizing GANs' training is through the perspective of dynamical systems (Arjovsky & Bottou, 2017; Kodali et al., 2017; Xu et al., 2020). Rather than the original minimax problem, they directly analyze the continuous-time training dynamics, which is a system of ordinary differential equations (ODEs) known as gradient flows. In this formulation, the behavior of GAN training closely relates to the concept of *asymptotic stability*, which depicts whether the training trajectory can eventually converge in the vicinity of an equilibrium of the gradient flow. Moreover, the stability of GANs' dynamical system can be improved by designing custom controllers from the control theory. With some modifications to the dynamics, several works (Kodali et al., 2017; Xu et al., 2020) can achieve asymptotic stability. However, these works have limitations. The stability they achieve is *local*, which assumes that the network is initialized sufficiently close to an equilibrium. Such an assumption cannot be guaranteed in practice. Furthermore, they only obtain *asymptotic* stability, without a guarantee to the convergence rate.

In this work, we leverage potent tools of noise-based controllers from control theory to propose a *Brownian motion controller* (BMC) to stabilize GANs' training. BMC stabilizes the training dynamics by adding noise, and the new noise-induced dynamics can be formulated as a system of stochastic differential equations (SDEs).

We extensively analyze BMC for Dirac-GANs (Mescheder et al., 2018), a popular class of GANs for analysis, to derive bounds on the rate of convergence. For Dirac-GANs, BMC are able to converge *globally* with *exponential* stability. Such theoretical result is much stronger than previous *local*, *asymptotic* ones (Heusel et al., 2017a; Kodali et al., 2017; Xu et al., 2020). Our global converge result implies that the trajectory can converge with *any* initialization rather than those near equilibria. Our globally exponential stability result further guarantees a fast convergence rate.

Empirically, BMC improves the stability of the training of several popular GAN architectures, including classical DC-GAN (Radford et al., 2015a) and the advanced StyleGANv2-ada (Karras et al., 2020a), which already has a carefully designed architecture and path-length regularization for stabilization. We evaluate BMC on a wide variety of datasets including CIFAR-10, LSUN-Bedroom 256x256, and FFHQ 1024x1024. Experiment results indicate that BMC can accelerate the training with a smaller range of oscillation, significantly reduced the training time, and achieved better image quality in terms of FID score.

2. Related Work

In this section, we discuss previous attempts for stabilizing GANs and present some background of noise-based
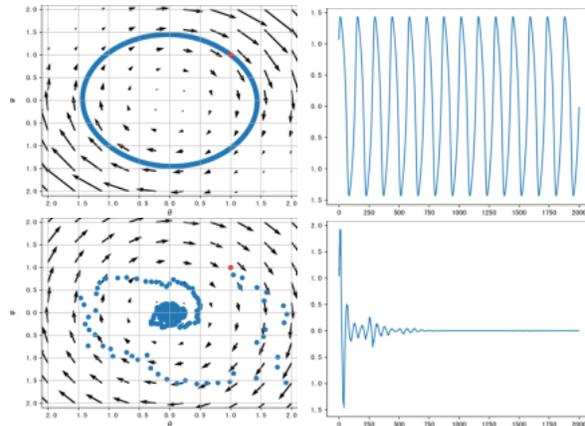

Figure 1: The gradient map (left) and convergence behavior of generator parameter θ (right) of Dirac-WGAN (first row) and DiracWGAN-BMC (second row), where the Nash equilibria of both models are at $(0,0)^\top$.

controllers in control theory. We compare the theoretical guarantee of stability of some representative works in Tab. 6.

Alternative Training Methods or Model Architecture To stabilize GANs' training process, a lot of work has been done on modifying its architecture. Wang et al. (2021) observe that during training, the discriminator converges faster and dominates the dynamics. They produce an attention map from the discriminator and use it to improve the spatial awareness of the generator. In this way, they claim to push GANs' solution closer to equilibrium. Heusel et al. (2017a) use a two-time scale update rule (TTUR) to push GANs' training process converging to a local Nash equilibrium. Karras et al. (2018) train the generator and the discriminator progressively to stabilize the training process. Later on, StyleGANs models (Karras et al., 2019; 2020a;b; 2021) use style transfer to alternate the architecture for the generator and accomplish the state-of-the-art performance on high-resolution image synthesis. However, their methods do not have guarantees of global stability.

Regularization on Objective Functions Many works stabilizes GANs' training process with modified objective functions. Kodali et al. (2017) add gradient penalty to their objective function to avoid local equilibrium with their model called DRAGAN. This method has fewer mode collapses and can be applied to a lot of GANs frameworks. Other work, such as Generative Multi-Adversarial Network (GMAN) (Durugkar et al., 2017), packing GANs (PacGAN) (Lin et al., 2017), GANs with spectrum control (Jiang et al., 2019), and energy-based GANs (Zhao et al., 2016), modifies the discriminator to achieve better stability. However, in their works, neither local nor global stability is promised and thus is orthogonal to our approach.

Other Methods with Control Theory Xu et al. (2020) formulate GANs as a system of differential equations and add closed-loop control (CLC) on a few variations of the GANs to enforce stability. However, the design of their controller depends on the objective function of the GANs models and does not work for all variations of the GANs models. Additionally, their method is neither globally nor exponentially stable, which we find unable to stabilize later proposed StyleGANv2-ada (Karras et al., 2020a).

Dynamic Systems with Noise-based Controller In control theory, theoretical works have proved that white noise is able to stabilize linear (Arnold et al., 1983; Scheutzow, 1993) and non-linear (Mao, 1994) dynamic systems. Later on, numerical experiments by Toral et al. (2001); He et al. (2003); Lin & Chen (2006); Sun & Yang (2011) observe a phenomenon called *Noise-induced synchronization*, which indicates adding designed independent Gaussian white noise successfully stabilizes dynamic systems. Our BMC incorporates the idea of noise-induced synchronization, in which we analyze GANs' training process from control theory's perspective and design an invariant Brownian Motion Controller (BMC) to stabilize GANs' training process.

3. Formulating GANs' Training as a Dynamical System

In this section, we briefly review the formulation of GAN training as a stability problem of dynamical systems. Generative adversarial networks (GANs) fit a data distribution $p(x)$ with a generator G and a discriminator D. The objective functions of GANs can be written as the following optimization problem:

$$\begin{cases} \max_D L_D(D; G) = \mathbb{E}_{p(x)}[h_1(D(x))] + \mathbb{E}_{p_G(x)}[h_2(D(x))] \\ \max_G L_G(G; D) = \mathbb{E}_{p_z(z)}[h_3(D(G(z)))], \end{cases}$$
(1)

where $h_1(\cdot)$, $h_3(\cdot)$ are increasing functions and $h_2(\cdot)$ is a decreasing function around zero. $p_G(x)$ is the distribution of the generator and p_z is a Gaussian distribution.

GANs seek for a *Nash equilibrium* (D_*, G_*), where

$$\forall D, L_D(D; G_*) \le L_D(D_*; G_*),$$
$$\forall G, L_G(G; D_*) \le L_G(G_*; D_*).$$

However, finding the global Nash equilibrium is challenging. Gradient-based methods are usually adopted for Eq. (1). A convenient way to analyze GANs is thus to directly consider the dynamics defined by the optimization algorithm.

Following Xu et al. (2020)'s notation, the training dynamic (under continuous time limit) for the generator and discrim-

inator over the time domain t can be written as:

$$\begin{cases} \dfrac{\mathrm{d}D(x,t)}{\mathrm{d}t} = p(x)\dfrac{\mathrm{d}h_1(D(x,t))}{\mathrm{d}D(x,t)} + p_G(x)\dfrac{\mathrm{d}h_2(D(x,t))}{\mathrm{d}D(x,t)}, \forall x \\ \dfrac{\mathrm{d}G(z,t)}{\mathrm{d}t} = p_z(z)\dfrac{\mathrm{d}h_3(D(G(z,t)),t)}{\mathrm{d}D(G(z,t),t)}\dfrac{\mathrm{d}D(G(z,t),t)}{\mathrm{d}G(z,t)}, \forall z \end{cases}$$
(2)

where $D(x,t)$ and $G(z,t)$ are respectively the generator and the discriminator over a time domain t. Define the state $X(t) = (D(x,t), G(z,t))^\top$ and the transition function

$$f(X(t)) = \begin{pmatrix} p(x)\frac{\mathrm{d}h_1(D(x,t))}{\mathrm{d}D(x,t)} + p_G(x)\frac{\mathrm{d}h_2(D(x,t))}{\mathrm{d}D(x,t)} \\ p_z(z)\frac{\mathrm{d}h_3(D(G(z,t)),t)}{\mathrm{d}D(G(z,t),t)}\frac{\mathrm{d}D(G(z,t),t)}{\mathrm{d}G(z,t)} \end{pmatrix},$$

we can rewrite Eq. (2) as

$$\mathrm{d}X(t) = f(X(t))\mathrm{d}t.$$
(3)

Rather than Nash equilibria, we analyze whether the algorithm can reach an *equilibrium of the dynamics* Eq. (3), i.e., a point X_e where $f(X_e) = 0$, which we refer by "equilibrium" in the rest of the paper. Such an equilibrium cannot be improved by local gradient updates, and is closely related to the concept of *local Nash equilibrium* (Heusel et al., 2017a).

If there exists an equilibrium, the question of interest is

Can the training algorithm find such an equilibrium?

This closely relates to the concept of *stability* in dynamical system theory, defined as follows:

Definition 3.1. (asymptotic stability) Given a dynamic system of with equilibirum X_e, this system is said to be *asymptotically stable*, if there exists a $\delta > 0$, such that whenever $\|X(0) - X_e\| \le \delta$, we have $\lim_{t \to \infty} \|X(t) - X_e\| = 0$.

Intuitively, an equilibrium X_e of a system is stable if the system can converge to X_e given any initial point close enough to it. Here, the convergence radius δ is of practical interest. While several works (Heusel et al., 2017a; Xu et al., 2020; Sønderby et al., 2016) show that (possibly modified) GAN training dynamics are stable, the concept of stability only guarantees convergence in a small neighborhood. However, the practical implication of such theoretical results is questionable since we cannot guarantee to initialize the network in the vicinity of an equilibrium. Additionally, in practice, we are concerned about the training time required for GANs to converge. However, current theoretical results only guarantee convergence as time goes to infinity.

4. Brownian Motion Controller for GANs

Viewing GANs' training as a dynamic system, we can leverage advances in control theory to design novel controllers to

stabilize the dynamics. The controllers define a new dynamical system with identical equilibria to the original system Eq. (3), but with better stabilization properties.

Before presenting formal theoretical results, we notice that while existing work (Mescheder et al., 2018; Xu et al., 2020) extracts GANs' training dynamic as a system of ordinary differential equations (ODEs), our new dynamics Eq. (8) is a system of stochastic differential equations (SDEs) due to the noise terms. Here, we present some preliminary concepts on SDEs and generalize the concept of stability to SDEs.

4.1. Preliminary

4.1.1. A SYSTEM OF SDEs

In statistical physics and engineering, the motion of many systems can be expressed as a SDE called *Langevin Equation* (Kloeden & Platen, 1992) such that

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = f(X(t)) + g(X(t))B(t), \tag{4}$$

where $X(t)$ is a m-dimensional vector representing the state of the system and $B(t) = (B_1(t), \ldots, B_m(t))^T$ is a m-dimensional *Brownian motion*, where each one-dimensional Brownian motion $\{B_i(t)\}_{t \geq 0}$ is a real-valued process with the following properties (Mao, 2007):

1. $B_i(0) = 0$ a.s.;
2. for $0 \leq s < t < \infty$, the increment $B_i(t) - B_i(s)$ is normally distributed with mean 0 and variance $t - s$;
3. The increments of the above property are independent to each other.

Given an initial state X_0, the training trajectory of GAN can be viewed as a solution of an initial value problem.

Definition 4.1. (Itô, 1951) Given a system described with the SDEs Eq. (4) with the initial value $X(0) = X_0$, a *solution* $\{X(t)\}_{t \geq 0}$ to this system is a stochastic process such that for $0 \leq t \leq \infty$, we have

$$X(t) = X_0 + \int_0^t f(X(s))ds + \int_0^t g(X(s))dB(s). \tag{5}$$

For any given system, we are concerned about whether this system exists a solution for a given initial point, and if so, whether the solution is unique (Mao, 1991). While for ODE, the existence and uniqueness of an initial value problem is guaranteed by the Picard–Lindelöf theorem, extra care needs to be taken for SDEs. We give the definitions below:

Definition 4.2. (Veretennikov, 1981) A system of SDEs *exists a global solution* if for any initial value X_0, its solution $\{X(t)\}_{t \geq 0}$ does not diverge and approach vertical asymptotic (line $x = a$ for some a) on $t \in [0, \infty)$.

Definition 4.3. (Mao & Yuan, 2006) A system of SDEs has *a unique solution* if given *any* initial value $X(0) = X_0$, there almost surly exists one and only one solution $\{X(t)\}_{t \geq 0}$ starting with this initial value.

4.1.2. STABILITY ANALYSIS

Modern stability theory of dynamical systems is developed under Lyapunov's second method (Goldhirsch et al., 1987; Shevitz & Paden, 1994; Sastry, 1999). Throughout our work, we evaluate the stability of GANs' training process under the framework of Lyapunov stability analysis. First, we generalize the definition of equilibria to SDEs:

Definition 4.4. X_e is said to be an *equilibrium* of a system of SDE, if there exists a t_0 such that $X(t_0) = X_e$ and for all $t \geq t_0$ satisfies $\mathbb{P}\big(f(X(t)) + g(X(t))B(t) = 0\big) = 1$.

As that for ODEs, equilibrium implies for a state which cannot be improved with the stochastic updates Eq. (4) and relates to a local Nash equilibrium. Then, whether the solution can converge to an equilibrium relates to the following definitions of stability:

Definition 4.5. Given a system of SDEs with an initial condition $X(0) = X_0$ and equilibrium X_e, if we assume $\{X_t\}_{t \geq 0}$ is its unique solution. This system is said to be

1. **Lyapunov stable**, if given $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $\|X(0) - X_e\| \leq \delta$, we have $\mathbb{P}\big(\|X(t) - X_e)\| < \epsilon\big) = 1$ for $0 \leq t \leq \infty$.
2. **asymptotically stable**, if this system is Lyapunov stable and there exists a $\delta > 0$, such that whenever $\|X(0) - X_e)\| \leq \delta$, we have $\mathbb{P}\big(\lim_{t \to \infty} \|X(t) - X_e\| = 0\big) = 1$.
3. **exponentially stable**, if it is asymptotically stable and there exists $\alpha > 0, \beta > 0, \delta > 0$ such that whenever $\|X(0) - X_e\| \leq \delta$, for $0 \leq t \leq \infty$, we have $\mathbb{P}\big(\|X(t) - X_e\| \leq \alpha \|X(0) - X_e\| e^{-\beta t}\big) = 1$.
4. **unstable**, if neither of the three conditions above is satisfied.

In what we follow, we say a dynamic system is ***globally Lyaponouv/ asymptotically/ exponentially stable***, if for *any* initial value, this system has a unique solution converging to equilibrium and exhibits Lyaponouv/ asymptotically/ exponential stability. Under this circumstance, the δ in definition 4.5 can be arbitrary, leading to stronger stability than what we defined in definition 4.5.

4.1.3. NOISE-BASED CONTROL ON ODEs

Given a system of ODEs $\frac{\mathrm{d}X(t)}{\mathrm{d}t} = f(X(t))$ with an equilibruim X_e which exihibts unstable nature, we say the controller $g(X(t))B(t)$ globally Lyapunov / asymptotically / exponentially *stabilizes* the system $\frac{\mathrm{d}X(t)}{\mathrm{d}t} = f(x(t))$ if and only if X_e is still an equilibruim and the system of SDEs $\frac{\mathrm{d}X(t)}{\mathrm{d}t} = f(X(t)) + g(X(t))B(t)$ is globally Lyapunov / asymptotically / exponentially stable.

Theorem 4.6. *(Mao, 1994) A system of ODEs* $\frac{dX(t)}{dt} = f(X(t))$ *with* $X(0) = X_0$ *can be almost surely exponentially stabilized by some noise term*

4

$g(X(t))B(t)$ *if their exists* $K > 0$, *such that* $f(X(t))$ *satisfies*

$$|f(X(t))| \leq K \|X\|, \forall X \in \mathbb{R}^m, t \geq 0 \qquad (6)$$

4.2. Brownian Motion Controller

In control theory, noise-based controllers are useful tools for stabilizing dynamical systems and pushing the solution towards optimal value over time domain t (Mao et al., 2002).

In our work, we propose *Brownian motion controller* (BMC), a higher order noise-based controller as a universal control function invariant to objective functions of GANs. Eq. (3) is a second-order dynamical system. We follow the framework of controllers for high order system proposed by (Wu & Hu, 2009) to propose our BMC as below:

$$u(t) = \varrho_1 X(t)\dot{B}_1(t) + \varrho_2 |X(t)|^{\beta} X(t)\dot{B}_2(t), \quad (7)$$

where $B_1(t)$ and $B_2(t)$ are independent one-dimensional Brownian motions, ϱ_1 and ϱ_2 are non-negative constants. Incorporating BMC Eq. (7), the controlled system becomes

$$\mathrm{d}X(t) = f(X(t))\mathrm{d}t + u(t). \qquad (8)$$

Intuitively, BMC adds multiplicative noise, i.e., where the strength of stochasticity depends on the current state of the system, to the original system. We will show in Sec. 5-6 (theoretically) and Sec. 7 (empirically) that training GANs with BMC can have better stabilization properties than previous GAN training dynamics and controllers (Kodali et al., 2017; Xu et al., 2020). Importantly, on Dirac-GAN, a popular class of GANs for analysis, BMC is *globally exponentially stable*. This means the training process can converge *exponentially fast* to the equilibrium starting from *any* initialization point. This theoretical result is much stronger than previous ones, which are only *locally asymptotically stable*.

5. Stability Analysis for Dirac-GANs

In this section, we focus on Dirac-GANs, a simplified GANs' setting with linear generator and discriminator networks proposed by Mescheder et al. (2018) to illustrate the unstable nature of GANs. We prove that Dirac-GAN with BMC is globally exponentially stable, and we derive bounds on its converge rate.

5.1. Dynamic System of Dirac-GANs

In Dirac-GANs' settings, the generator is a point mass following $p_G(x) = \delta(x - \theta)$ and the discriminator is a linear $D_{\phi}(x) = \phi x$. The true data distribution is given by $p_D(x) = \delta(x - c)$ with a constant c. Here, $\delta(\cdot)$ is a Dirac distribution with $\delta(x) = 0$ for $x \neq 0$ and $\int_{-\infty}^{\infty} \delta(x) = 1$.

The objective functions of Dirac-GANs can be written as:

$$\begin{cases} \max_{\phi} L_{D_{\phi}}(\phi; \theta) = h_1(D_{\phi}(c)) + h_2(D_{\phi}(\theta)) \\ \max_{\theta} L_{G_{\theta}}(\theta; \phi) = h_3(D_{\phi}(\theta)), \end{cases} \qquad (9)$$

where $h_1(\cdot)$ and $h_3(\cdot)$ are increasing functions and $h_2(\cdot)$ is a decreasing function around zero (Xu et al., 2020).

Taking the reparameterization $\tilde{\theta}(t) = \theta(t) - c$, we have the following dynamic system with a unique equilibrium $(0,0)^T$:

$$\begin{cases} \frac{\mathrm{d}\phi(t)}{\mathrm{d}t} = h_1'(\phi(t)c)c + h_2'(\phi(t)(\tilde{\theta}(t) + c))(\tilde{\theta}(t) + c) \\ \frac{\mathrm{d}\tilde{\theta}(t)}{\mathrm{d}t} = h_3'(\phi(t)(\tilde{\theta}(t) + c))\phi(t). \end{cases} \qquad (10)$$

Depending on the specific choice of h_1, h_2, and h_3, the training dynamics of many GANs can be represented by Eq. (10) with the minimalistic generator and discriminator architecture. These GANs are referred as Dirac-GAN, Dirac-WGAN, and so on. Unfortunately, Mescheder et al. (2018) find that many Dirac-GANs are unstable. Our experiments also confirm the unstable nature of Eq. (10). As in Fig. 1, Dirac-WGAN is unstable and circles around the origin.

To stabilize Eq. (10), we apply our designed BMC and transform Dirac-GANs' training dynamics to Eq. (8), where $X(t) = (\phi(t), \theta(t))^{\top}$, $f(X(t)) = (h_1'(\phi(t)c)c + h_2'(\phi(t)(\tilde{\theta}(t)+c))(\tilde{\theta}(t)+c), h_3'(\phi(t)(\tilde{\theta}(t)+c))\phi(t))^{\top}$, and $u(t)$ is defined as in Eq. (7).

5.2. Dirac-GANs are Exponentially Stable with BMC

In this section, we derive the existence of unique global solution and stability of system (8). For the stability analysis, we impose the following assumption on the smoothness of functions h_1, h_2, h_3 in system (8).

Assumption 5.1. There exist positive constants α_1, α_2, α_3 such that for any $x, y \in \mathbb{R}^n$,

$$|h_1'(x) - h_1'(y)| \leq \alpha_1 \|x - y\|, \quad |h_2'(x) - h_2'(y)| \leq \alpha_2 \|x - y\|,$$

$$|h_3'(x) - h_3'(y)| \leq \alpha_3 \|x - y\|.$$

In what follows, we first prove that the BMC from equation (7) yields a unique global solution in Theorem 5.2. Then, in Theorem 5.3, we show that this unique global solution exponentially converges to the equilibrium point *a.s.* with bounds on the hyper-parameters ϱ_1, ϱ_2 and β which in turn affect the rate of convergence. Combining Theorem 5.2 and Theorem 5.3, we claim that with our BMC, Dirac-GANs representing by system (8) is globally exponentially stable.

Theorem 5.2. *(Proof in Appendix A) Under Assumption 5.1, for any initial value $X(0) = \xi \in \mathbb{R}^2$, if $\varrho_2 \neq 0$ and $\beta > 1$, then there a.s. exists a unique global solution $X(t)$ to system (8) on $t \in [0, \infty)$.*

Theorem 5.3. *(Proof in Appendix B) Let Assumption 5.1 hold. Assume that $\varrho_2 \neq 0$ and $\beta > 1$. If $\frac{\varrho_1^2}{2} - \varphi > 0$, where φ takes the value of*

$$\max_{x \geq 0} \left\{ -\frac{\varrho_2^2}{2} x^{2\beta} + (\alpha_2^2 + \frac{1}{2}\alpha_3^2) x^2 + [(1 + \frac{1}{2}\alpha_1^2)c^2 + 2c + \frac{1}{2}] \right\}, \tag{11}$$

then for any $X(0) = \xi$ with sufficiently small constant $\epsilon \in (0, \varrho_1^2/2 - \varphi)$, the global solution $X(t)$ of system (8) has the property that

$$\limsup_{t \to \infty} \frac{\log |X(t)|}{t} \leq -\left(\frac{\varrho_1^2}{2} - \varphi\right) + \epsilon, \quad a.s.$$

that is, the solution of system (8) is a.s. exponentially stable.

Here since $\frac{\varrho_1^2}{2} - \varphi > 0$ and ϵ is a sufficiently small constant, then when Eq. (11) is satisfied, we have $\limsup_{t \to \infty} \frac{\log |X(t)|}{t} \leq -\lambda$, a.s. for some positive constant λ. Rearranging we get $\limsup_{t \to \infty} |X(t)| \leq e^{-\lambda t}$, which implies $\limsup_{t \to \infty} X(t) = (0,0)^\top$ as required. Notice that the rate of convergence depends only on constant λ, which in turn depends on ϱ_1 and φ. Thus the convergence rate is decided by the choice of hyper-parameters ϱ_1, ϱ_2, and β. In practice, we can tune these three variables as desired, as long as they satisfy the constraint from Eq. (11).

Notice that our BMC work for any h_1, h_2 and h_3 as long as they satisfy the smoothness condition under assumption 5.1. In other words, we have proven that with BMC, Dirac-GANs are *globally exponentially stable* regardless of h_1, h_2 and h_3, and we have given theoretical bounds on the rate of convergence. In Fig. 1, we present visual proof that the Dirac-GAN with BMC is stable and converges to the optimal equilibrium as required.

6. Stabilize General GANs with BMC

In this section, we present BMC for general GANs beyond Dirac-GAN. Under GANs' setting, the generator and discriminator learn to compete with each other, and ideally, they would converge to their global optimal solution simultaneously. However, in reality, the generator and discriminator typically learn at different paces (Wang et al., 2021). For example, in DCGAN (Radford et al., 2015a) the discriminator learns faster, while in StyleGANv2 (Karras et al., 2020a) the generator learns faster. This inconsistency would result in the faster network being stuck in its local equilibrium while the slower one stops learning useful information and lead to mode-collapses or unstable behaviors.

BMC considers both the generator and discriminator as a *coupled dynamic system* (a system of two differential equations with two dependent variables D and G, and one independent variable t) (Peng & Wu, 1999) and the goal of

our BMC is to make sure the generator and discriminator learn at the same pace so that they are able to converge to optimal equilibrium at the same time.

For simplicity, we take $\beta = 2$ (the order of noise often take an even number) for general GANs, then BMC in this setting becomes

$$u(t) = \varrho_1 X(t)\dot{B}_1(t) + \varrho_2 |X(t)|^2 X(t)\dot{B}_2(t) \tag{12}$$

Our BMC can be reflected on objective functions as a regularization term on both generator and discriminators, with the derivative being Eq. (12). We thus take integration of Eq. (12) and modify the objective functions in (1) to:

$$\begin{cases} \max_D L_D'(D;G) = L_D(D;G) + \frac{1}{2}\varrho_1 D^2(x)\dot{B}_1(t) \\ \quad + [\frac{1}{4}\varrho_2 D^4(x) + \frac{1}{2}\varrho_2 D^2(G(z))D^2(x)]\dot{B}_2(t). \\ \max_G L_G'(G;D) = L_G(G;D) + \frac{1}{2}\varrho_1 D^2(G(z))\dot{B}_1(t) \\ \quad + [\frac{1}{4}\varrho_2 D^4(G(z)) + \frac{1}{2}\varrho_2 D^2(G(z))D^2(x)]\dot{B}_2(t). \end{cases} \tag{13}$$

Although in general GANs' setting, we are not aware of the equilibrium point of the generator, according to Theorem 5.3, this equilibrium would affect the constraints between $\varrho 1$ and ϱ_2. As a result, the stability is only guaranteed for certain pairs of $\varrho 1$ and ϱ_2. We implement our designed objective functions in Sec. 7 and analyze results for various pairs of $\varrho 1$ and ϱ_2. Our numerical experiments show that BMC successfully stabilizes GANs models and are able to generate images with promising quality.

7. Evaluation

In this section, we show the effectiveness of BMC by providing both quantitative and qualitative results. We first evaluate the performance of BMC for Dirac-GANs ("Dirac-GAN-BMC") in Sec. 7.1 and present Fig. 1 to illustrate that our BMC successfully stabilizes Dirac-GANs. Additionally, we compare the number of iterations needed for Dirac-GAN-BMC to converge under various parameters of BMC. Then in Sec. 7.2, we compare our GAN-BMC with the StyleGANv2-ada baseline (Karras et al., 2020a) on multiple datasets with various parameter settings of BMC. Experiments show that our BMC effectively stabilizes the training dynamics by reducing the range of oscillation, speeding up the convergence rate, and improving the FID score.

7.1. Convergence of Dirac-GANs with BMC

In Dirac-GAN's setting, we know the optimal equilibrium is $(0,0)^\top$, so we can measure the convergence speed and draw the gradient map for comparison, which are presented in Fig. 1. These results show that our Dirac-GAN-BMC has better convergence patterns and speed than Dirac-GANs. Without adding BMC to the training objective, Dirac-GANs
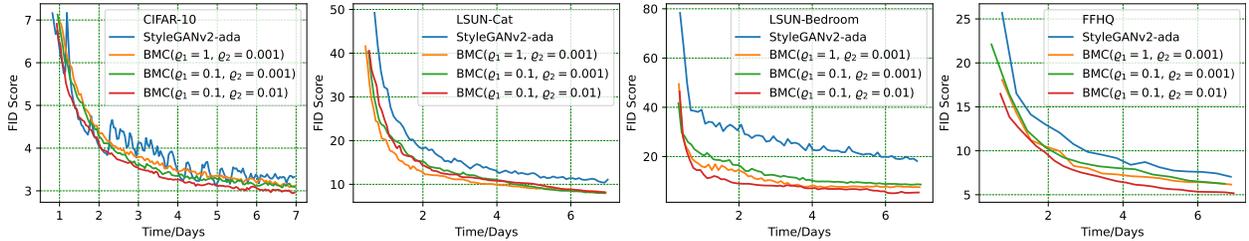
Figure 2: The FID score curves with respect to time. The blue curve represents StyleGANv2-ada baseline using the same parameter settings as their original paper (Karras et al., 2020a) , while the green, orange and red curves add BMC and are evaluated under the same setting.

cannot reach equilibrium. The parameters of the generator and the discriminator oscillate in a circle, as shown in Fig. 1. In comparison, these parameters of the Dirac-GAN-BMC only oscillate in the first 500 iterations and soon converge in 800 iterations.

We also study different combinations of the hyperparameters ϱ_1 and ϱ_2 under $\beta = 1$ and $\beta = 2$. As shown in Tab. 2, Dirac-GANs-BMC converge better when we set $\varrho_1 = 0.1$ and $\varrho_2 = 0.01$. Generally, a larger ϱ will lead to a faster convergence rate, but when ϱ is large enough, the effect of increasing ϱ will become saturated. On the other hand, when ϱ is too small, Dirac-GAN-BMC will take more than 100000 iterations to converge.

Table 2: Converge iterations required for $\beta = 1$ and $\beta = 2$ on DiracGANs-BMC. Symbol N below indicates not converging after 100k iterations.

$\beta = 1/2$	$\varrho_2 = 0.0001$	$\varrho_2 = 0.001$	$\varrho_2 = 0.01$
$\varrho_1 = 0.1$	0.6k / 1.5k	0.4k / 0.75k	0.4k / 0.7k
$\varrho_1 = 0.01$	25k / N	15k / 9k	10k / 8.5k
$\varrho_1 = 0.001$	N / N	N / N	40k / N

7.2. GANs-BMC: Converge better and lower FID

Dataset: We evaluate our proposed GANs-BMC on well-established CIFAR10 (Krizhevsky et al., 2009), LSUN-Bedroom with resolution 256x256 (Yu et al., 2015), LSUN-Cat with resolution 256x256 (Yu et al., 2015), and FFHQ with resolution 1024x1024 (Karras et al., 2019).

Implementation details: We compare the FID score (Heusel et al., 2017b) of StyleGANv2-ada (Karras et al., 2020a) and its stabilized version with our BMC ("StyleGANv2-ada-BMC"). We reproduce the identical configuration settings as reported in the StyleGANv2-ada paper within the period of 7 days on 4 cards of NVIDIA GeForce GTX TITAN X. The detailed experimental setups can be found in Appendix C. We find that adding BMC results in better performance without changing any hyperparameters

of the network, as shown in Fig. 2 . We conduct multiple trials under different coefficients of BMC and report the FID score with its range of oscillation in Tab. 3. Additionally, we calculate the shifting difference of the generator throughout the training process, shown in Fig. 3.

7.3. Quantitative Results

As in Fig. 2 and Tab. 3, StyleGAN2-ada has a large range of oscillation on FID scores. Especially on the low-resolution dataset CIFAR-10, we observe that the FID curve oscillates even after training for a long time. On the other hand, StyleGANv2-ada-BMC is able to reduce this range of oscillation by a multiple of 10 in CIFAR-10 32*32, LSUN-Bedrrom 256*256, LSUN-Cat 256*256 and a multiple of 4 on FFHQ 1024*1024. As a result, BMC largely reduces the range of oscillation, providing better convergence behavior.

Table 3: FID scores for each dataset after training for 7 days. ($a \pm b$ in this table should be interpreted as a is the FID score after 7 days, b is the range of oscillation on the 24 hours span of the seventh day).

		$\varrho_1 = 1$	$\varrho_1 = 0.1$	$\varrho_1 = 0.01$
Cifar10 (32x32)	$\varrho_2 = 0.01$	3.12 ± 0.03	$\mathbf{2.94 \pm 0.01}$	3.56 ± 0.06
	$\varrho_2 = 0.001$	3.05 ± 0.02	2.95 ± 0.02	3.31 ± 0.04
	$\varrho_2 = 0.0001$	2.98 ± 0.02	2.96 ± 0.02	3.15 ± 0.05
	Baseline		3.32 ± 0.16	
LSUN Bedroom (256x256)	$\varrho_2 = 0.01$	$\mathbf{5.34 \pm 0.09}$	5.40 ± 0.08	7.06 ± 0.09
	$\varrho_2 = 0.001$	7.52 ± 0.07	8.62 ± 0.12	8.82 ± 0.15
	$\varrho_2 = 0.0001$	7.48 ± 0.06	9.01 ± 0.14	9.25 ± 0.17
	Baseline		18.16 ± 1.00	
LSUN Cat (256x256)	$\varrho_2 = 0.01$	8.42 ± 0.06	8.07 ± 0.04	9.09 ± 0.16
	$\varrho_2 = 0.001$	8.03 ± 0.04	$\mathbf{8.01 \pm 0.03}$	9.23 ± 0.18
	$\varrho_2 = 0.0001$	9.56 ± 0.23	9.30 ± 0.21	9.48 ± 0.19
	Baseline		10.57 ± 0.31	
FFHQ (1024x1024)	$\varrho_2 = 0.01$	5.74 ± 0.29	$\mathbf{5.17 \pm 0.10}$	7.96 ± 0.14
	$\varrho_2 = 0.001$	6.16 ± 0.14	6.24 ± 0.17	7.32 ± 0.18
	$\varrho_2 = 0.0001$	6.02 ± 0.09	6.03 ± 0.11	8.01 ± 0.23
	Baseline		7.75 ± 0.42	

While StyleGANv2-ada's takes a long training time, especially on the dataset with high resolution, our BMC speeds up the training time for producing images with the identical FID scores by more than 2 days as observed in Fig. 2.

Additionally, we measure the training stability by defining
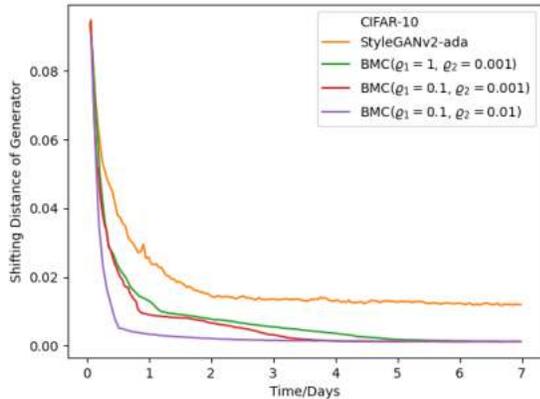
Figure 3: The shifting difference against time on CIFAR-10 baseline and its BMC trails. Clearly the shifting difference with BMC approaches 0.

the shifting difference of the generator as

$$D(G_1, G_2) = \mathbb{E}_{p_z(z)} \|G_1(z) - G_2(z)\|.$$

We calculate the shifting difference of the generator in between iterations and, as in Fig. 3, while StyleGANv2-ada has a non-vanishing shifting difference at around 0.02, the shifting difference of our BMC approaches 0, indicating our BMC successfully stabilizes the generator.

7.4. Qualitative results

We provide qualitative results on FFHQ 1024x1024 in Fig. 4. Noticing that our GANs-BMC is able to produce images with higher quality at a faster speed.

8. Conclusion and Discussion

In this paper, we revisit GANs' instability problem from the perspective of control theory. Our work novelly incorporates noise-based controller of on the training dynamics of GANs and modifies its objective function accordingly to stabilize GANs. We innovatively design a Brownian Motion Control (BMC) to achieve globally exponential stability. Notably, our BMC is compatible with all GANs variations. In our paper, theoretical analysis has been done under the Dirac-GANs setting, and we are able to stabilize both the generator and discriminator simultaneously. Experimental results demonstrate that our BMC accelerates the convergence of GANs and performs better in terms of FID than StyleGANv2-ada on CIFAR-10, LSUN-Bedroom, LSUN-Cat, and FFHQ.

Possible Future Directions While in practice, the training process of GANs is discrete, we model it as a continuous dynamic system in our work. As a result, our method performs better on training with smaller time intervals per iteration.
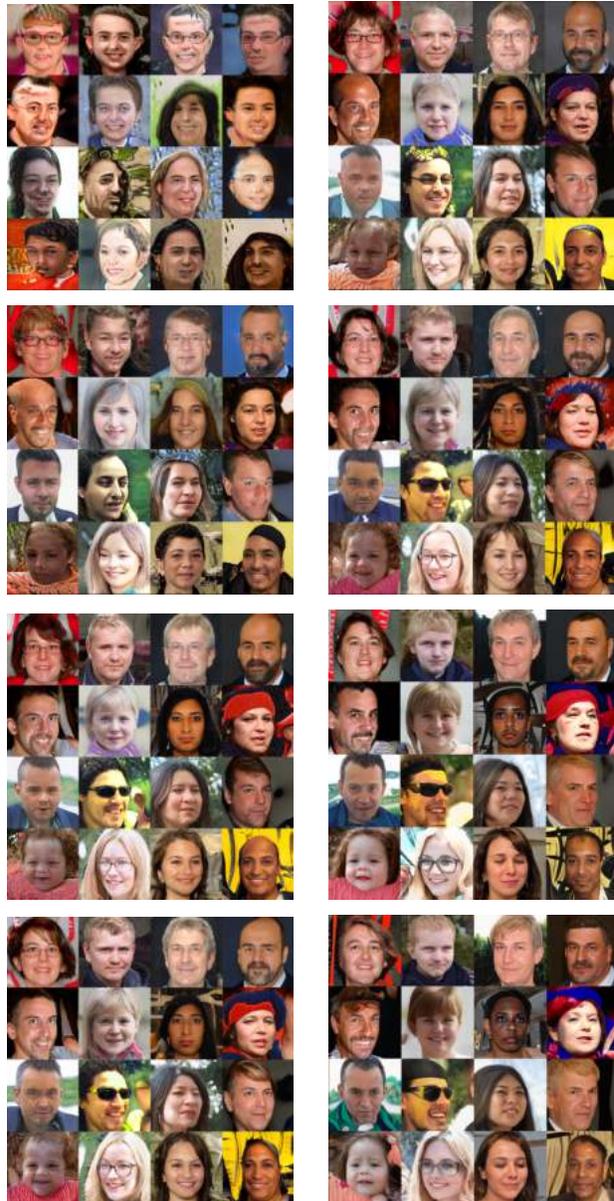
Figure 4: Each row represents images generated from the FFHQ 1024*1024 dataset within the first 20 hours (5, 10, 15, and 20 hours). The first column of images is generated by StyleGANv2-ada, and the second column of images is generated after adding BMC.

Additionally, although many advanced control methods, including our BMC, are available to stabilize the dynamic system, few of them have been applied in deep generative models. As a result, future work can be done on modeling training as a discrete process and exploring more control methods to improve its stability.

Social Impacts and Potential Harmful Consequences Like other generative models, our GAN-BMC has ethical

problems related to deep fake. The generated photos and videos are indistinguishable from the real ones and may be used illegally to spread misleading information for harmful uses. Secondly, our GAN-BMC may be biased toward concern identity categories, such as race and gender. Product farming is another ethical concern GAN-BMC has, as generated images join art contests and bring fairness issues for artists.

Acknowledgements

References

Arjovsky, M. and Bottou, L. Towards principled methods for training generative adversarial networks. *International Conference on Learning Representations (ICLR)*, 2017.

Arnold, L., Crauel, H., and Wihstutz, V. Stabilization of linear systems by noise. *SIAM Journal on Control and Optimization*, 21(3):451–461, 1983.

Brock, A., Donahue, J., and Simonyan, K. Large scale gan training for high fidelity natural image synthesis. *arXiv preprint arXiv:1809.11096*, 2018.

Durugkar, I., Gemp, I., and Mahadevan, S. Generative multi-adversarial networks. *International Conference on Learning Representations (ICLR)*, 2017.

Farnia, F. and Ozdaglar, A. Do gans always have nash equilibria? In *International conference on machine learning*. PMLR, 2020.

Fedus, W., Rosca, M., Lakshminarayanan, B., Dai, A., Mohamed, S., and Goodfellow, I. Many paths to equilibrium: Gans do not need to decrease a divergence at every step. *International Conference on Learning Representations (ICLR)*, 2018.

Goldhirsch, I., Sulem, P.-L., and Orszag, S. A. Stability and lyapunov stability of dynamical systems: A differential approach and a numerical method. *Physica D: Nonlinear Phenomena*, 27(3):311–337, 1987.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. *Advances in Neural Information Processing Systems*, pp. 2672–2680, 2014.

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin1, V., and Courville, A. Improved training of wasserstein gans. *arXiv preprint arXiv:1704.00028*, 2017.

He, D., Shi, P., and Stone, L. Noise-induced synchronization in realistic models. *Physical Review E*, 67(2):027201, 2003.

Heusel, M., Ramsauer, H., Unterthiner, T., and Nessler, B. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *arXiv preprint arXiv:1706.08500*, 2017a.

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., and Hochreiter, S. Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30, 2017b.

Itô, K. *On stochastic differential equations*. Number 4. American Mathematical Soc., 1951.

Jiang, H., Chen, Z., Chen, M., Liu, F., Wang, D., and Zhao, T. On computation and generalization of generative adversarial networks under spectrum control. In *International Conference on Learning Representations*, 2019.

Karras, T., Aila, T., Laine, S., and Lehtinen, J. Progressive growing of gans for improved quality, stability, and variation. *International Conference on Learning Representations (ICLR)*, 2018.

Karras, T., Laine, S., and Aila, T. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 4401–4410, 2019.

Karras, T., Aittala, M., Hellsten, J., Laine, S., Lehtinen, J., and Aila, T. Training generative adversarial networks with limited data. *Advances in Neural Information Processing Systems*, 33:12104–12114, 2020a.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., and Aila, T. Analyzing and improving the image quality of stylegan. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 8110–8119, 2020b.

Karras, T., Aittala, M., Laine, S., Härkönen, E., Hellsten, J., Lehtinen, J., and Aila, T. Alias-free generative adversarial networks. *Advances in Neural Information Processing Systems*, 34:852–863, 2021.

Kloeden, P. E. and Platen, E. Stochastic differential equations. In *Numerical solution of stochastic differential equations*, pp. 103–160. Springer, 1992.

Kodali, N., Abernethy, J., Hays, J., and Kira, Z. On convergence and stability of gans. *arXiv preprint arXiv:1705.07215v5*, 70, 2017.

Krizhevsky, A., Hinton, G., et al. Learning multiple layers of features from tiny images. 2009.

Li, C.-L., Chang, W.-C., Cheng, Y., Yang, Y., and Póczos, B. Mmd gan: Towards deeper understanding of moment matching network. *Advances in neural information processing systems*, 30, 2017.

Lin, W. and Chen, G. Using white noise to enhance synchronization of coupled chaotic systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 16(1): 013134, 2006.

Lin, Z., Khetan, A., Fanti, G., and Oh, S. Pacgan: The power of two samples in generative adversarial networks. *arXiv preprint arXiv:1712.04086*, 2017.

Mao, X. A note on global solution to stochastic differential equation based on a semimartingale with spatial parameters. *Journal of Theoretical Probability*, 4(1):161–167, 1991.

Mao, X. Stochastic stabilization and destabilization. *Systems & control letters*, 23(4):279–290, 1994.

Mao, X. *Stochastic differential equations and applications*. Elsevier, 2007.

Mao, X. and Yuan, C. *Stochastic differential equations with Markovian switching*. Imperial college press, 2006.

Mao, X., Marion, G., and Renshaw, E. Environmental brownian noise suppresses explosions in population dynamic. *Stochastic Processes and their Applications*, 97, 2002.

Mathieu, M. F., Zhao, J. J., Zhao, J., Ramesh, A., Sprechmann, P., and LeCun, Y. Disentangling factors of variation in deep representation using adversarial training. *Advances in neural information processing systems*, 29, 2016.

Mescheder, L., Geiger, A., and Nowozin, S. Which training methods for gans do actually converge? In *International conference on machine learning*, pp. 3481–3490. PMLR, 2018.

Miyato, T., Kataoka, T., Koyama, M., and Yoshida, Y. Spectral normalization for generative adversarial networks. *arXiv preprint arXiv:1802.05957*, 2018.

Nowozin, S., Cseke, B., and Tomioka, R. f-gan: Training generative neural samplers using variational divergence minimization. *Advances in neural information processing systems*, 29, 2016.

Peng, S. and Wu, Z. Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM Journal on Control and Optimization*, 37(3):825–843, 1999.

Radford, A., Metz, L., and Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015a.

Radford, A., Metz, L., and Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015b.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. Improved techniques for training gans. *Advances in neural information processing systems*, 29, 2016.

Sastry, S. Lyapunov stability theory. In *Nonlinear systems*, pp. 182–234. Springer, 1999.

Scheutzow, M. Stabilization and destabilization by noise in the plane. *Stochastic Analysis and Applications*, 11(1): 97–113, 1993.

Shevitz, D. and Paden, B. Lyapunov stability theory of nonsmooth systems. *IEEE Transactions on automatic control*, 39(9):1910–1914, 1994.

Sønderby, C. K., Caballero, J., Theis, L., Shi, W., and Huszár, F. Amortised map inference for image super-resolution. *arXiv preprint arXiv:1610.04490*, 2016.

Sun, Z. and Yang, X. Generating and enhancing lag synchronization of chaotic systems by white noise. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 21(3): 033114, 2011.

Toral, R., Mirasso, C. R., Hernández-Garcıa, E., and Piro, O. Analytical and numerical studies of noise-induced synchronization of chaotic systems. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 11(3):665–673, 2001.

Veretennikov, A. J. On strong solutions and explicit formulas for solutions of stochastic integral equations. *Mathematics of the USSR-Sbornik*, 39(3):387, 1981.

Wang, J., Yang, C., Xu, Y., Shen, Y., Li, H., and Zhou, B. Improving gan equilibrium by raising spatial awareness. *arXiv preprint arXiv:2112.00718*, 2021.

Wu, F. and Hu, S. Suppression and stabilisation of noise. *International Journal of Control*, 82(11):2150–2157, 2009.

Xu, K., Li, C., Zhu, J., and Zhang, B. Understanding and stabilizing GANs' training dynamics with control theory. In *International conference on machine learning*. PMLR, 2020.

Yu, F., Seff, A., Zhang, Y., Song, S., Funkhouser, T., and Xiao, J. Lsun: Construction of a large-scale image dataset using deep learning with humans in the loop. *arXiv preprint arXiv:1506.03365*, 2015.

Zhao, J., Mathieu, M., and LeCun, Y. Energy-based generative adversarial network. *arXiv preprint arXiv:1609.03126*, 2016.

Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232, 2017.

Appendix A: Proof of Theorem 5.2

Under Assumption 5.1, for any initial value $X(0) = \xi \in \mathbb{R}^2$, if $\varrho_2 \neq 0$ and $\beta > 1$, then there a.s. exists a unique global solution $X(t)$ to system (8) on $t \in [0, \infty)$.

Proof. Under Assumption 5.1, then, we can calculate that

$$
\begin{aligned}
&X^{\mathrm{T}}(t) f(X(t)) \\
=&\phi(t) h_1^{'}(\phi(t)c)c + \phi(t) h_2^{'}(\phi(t)(\tilde{\theta}(t) + c))\tilde{\theta}(t) + \\
&\phi(t) h_2^{'}(\phi(t)(\tilde{\theta}(t) + c))c + \\
&\tilde{\theta}(t) h_3^{'}(\phi(t)(\tilde{\theta}(t) + c))\phi(t) \\
\leq&[(1 + \frac{1}{2}\alpha_1^2)c^2 + 2c + \frac{1}{2}]|X|^2 + (\alpha_2^2 + \frac{1}{2}\alpha_3^2)|X|^4.
\end{aligned}
$$

For any bounded initial value $X(0) \in \mathbb{R}^n$, there exists a unique maximal local strong solution $X(t)$ of system (8) on $t \in [0, \tau_e)$, where τ_e is the explosion time. To show that the solution is actually global, we only need to prove that $\tau_e = \infty$ a.s. Let k_0 be a sufficiently large positive number such that $|X(0)| < k_0$. For each integer $k \geq k_0$, define the stopping time

$$
\tau_k = \inf\{t \in [0, \tau_e) : |X(t)| \geq k\}
$$

with the traditional setting $\inf \emptyset = \infty$, where $\emptyset$ denotes the empty set. Clearly, τ_k is increasing as $k \to \infty$ and $\tau_k \to \tau_\infty \leq \tau_e$ a.s. If we can show that $\tau_\infty = \infty$, then $\tau_e = \infty$ a.s., which implies the desired result. This is also equivalent to prove that, for any $t > 0$, $\mathbb{P}(\tau_k \leq t) \to 0$ as $k \to \infty$. To prove this statement, for any $p \in (0, 1)$, define a C^2-function

$$
V(x) = |X(t)|^p.
$$

One can obtain that $X(t) \neq 0$ for all $0 \leq t \leq \tau_e$ a.s. Thus, one can apply the Itô formula to show that for any $t \in [0, \tau_e)$,

$$
\begin{aligned}
\mathrm{d}V(X(t)) =&LV(X(t))\mathrm{d}t + p\varrho_1|X(t)|^p\mathrm{d}B_1(t) \\
&+ p\varrho_2|X(t)|^{\beta+p}\mathrm{d}B_2(t),
\end{aligned}
$$

where LV is defined as

$$
\begin{aligned}
LV(X) =&p|X|^{p-2}X^{\mathrm{T}}f(X(t)) + \frac{p(p-1)\varrho_1^2}{2}|X|^p \\
&+ \frac{p(p-1)\varrho_2^2}{2}|X|^{2\beta+p}
\end{aligned}
$$

By Assumption 5.1, we therefore have

$$
\begin{aligned}
LV(X) \leq&\frac{p(p-1)\varrho_2^2}{2}|X|^{2\beta+p} + ((1 + \frac{1}{2}\alpha_1^2)c^2 + 2c + \frac{1}{2}) \\
&p|X|^{\alpha+p} + p\left(\frac{(p-1)\varrho_1^2}{2} + (\alpha_2^2 + \frac{1}{2}\alpha_3^2)\right)|X|^p.
\end{aligned}
$$

Noting that $p \in (0, 1)$ and $\beta > 1$ and $\varrho_2 \neq 0$, by the boundedness of polynomial functions, there exists a positive constant $\bar{H}$ such that $LV(x) \leq \bar{H}$. We therefore have

$$
\begin{aligned}
\mathbb{E}V(X(t \wedge \tau_k)) &\leq \mathbb{E}|\xi|^p + \mathbb{E}\int_0^{t\wedge\tau_k} LV(X(s))\mathrm{d}s \\
&\leq \mathbb{E}|\xi|^p + \bar{H}t \\
&=: \bar{H}_t,
\end{aligned}
$$

where $\bar{H}_t$ is independent of k. By the definition of τ_k, $|X(\tau_k)| = k$, so

$$
\begin{aligned}
\mathbb{P}(\tau_k \leq t)k^p &\leq \mathbb{P}(\tau_k \leq t)V(X(\tau_k))] \\
&\leq \mathbb{E}[l_{\{\tau_k \leq t\}}V(X(t \wedge \tau_k))] \\
&\leq \mathbb{E}V(X(t \wedge \tau_k)) \\
&\leq \bar{H}_t,
\end{aligned}
$$

which implies that

$$
\limsup_{k \to \infty} \mathbb{P}(\tau_k \leq t) \leq \lim_{k \to \infty} \frac{\bar{H}_t}{k^p} = 0,
$$

as required. $\square$

Appendix B: Proof of Theorem 5.3

Let Assumption 5.1 hold. Assume that $\varrho_2 \neq 0$ and $\beta > 1$. If

$$\frac{\varrho_1^2}{2} - \varphi > 0,$$

where

$$\varphi = \max_{x \geq 0} \left\{ -\frac{\varrho_2^2}{2} x^{2\beta} + (\alpha_2^2 + \frac{1}{2}\alpha_3^2)x^2 + [(1 + \frac{1}{2}\alpha_1^2)c^2 + 2c + \frac{1}{2}] \right\}, \tag{14}$$

then for any $X(0) = \xi$ with sufficiently small constant $\epsilon \in (0, \varrho_1^2/2 - \varphi)$, the global solution $X(t)$ of system (8) has the property that

$$\limsup_{t \to \infty} \frac{\log |X(t)|}{t} \leq -\left(\frac{\varrho_1^2}{2} - \varphi \right) + \epsilon, \quad a.s.$$

that is, the solution of system (8) is a.s. exponentially stable.

Proof. Applying Itô formula to $\log |X(t)|$ yields

$$\log |X(t)| = \log |X(0)| + \int_0^t \left[|X(t)|^{-2} X^{\mathrm{T}}(s) f(X(s)) \right.$$

$$\left. - \frac{\varrho_2^2}{2} |X(s)|^{2\beta} - \frac{\varrho_1^2}{2} \right] \mathrm{d}s + \int_0^t \varrho_1 \mathrm{d}B_1(t)$$

$$+ \varrho_2 \int_0^t |X(s)|^\beta \mathrm{d}B_2(s).$$

Letting $M(t) = \varrho_2 \int_0^t |X(s)|^\beta \mathrm{d}B_2(s)$, clearly $M(t)$ is a continuous local martingale with the quadratic variation

$$< M(t), M(t) >= \varrho_2^2 \int_0^t |X(s)|^{2\beta} \mathrm{d}s.$$

For any $\varepsilon \in (0, 1)$, choose $\theta > 0$ such that $\theta\varepsilon > 1$. Then for each integer $m > 0$, the exponential martingale inequality gives

$$\mathbb{P} \left\{ \sup_{1 \leq t \leq m} \left[M(t) - \frac{\varepsilon\varrho_2^2}{2} \int_0^t |X(s)|^{2\beta} \mathrm{d}s \right] \geq \theta\varepsilon \log m \right\} \leq \frac{1}{m^{\theta\varepsilon}}.$$

Since $\sum_{m=1}^{\infty} m^{-\theta\varepsilon} < \infty$, by the well-known Borel-Cantelli lemma, there exists an $\bar{\Omega}_0 \subseteq \Omega$ with $\mathbb{P}(\bar{\Omega}_0) = 1$ such that for any $\omega \in \bar{\Omega}_0$, there exists an integer $\bar{m}(\omega)$, when $m > \bar{m}(\omega)$, and $m - 1 \leq t \leq m$,

$$M(t) \leq \frac{\varepsilon\varrho_2^2}{2} \int_0^t |X(s)|^{2\beta} \mathrm{d}s + \theta\varepsilon \log(t + 1).$$

This, together with Assumption 5.1, yields

$$\log |X(t)| \leq \log |\xi| + \int_0^t \left[-\frac{\varrho_2^2(1 - \varepsilon)}{2} |X(s)|^{2\beta} \right.$$

$$+ (1 + (\frac{1}{2}\alpha_1^2)c^2 + 2c + \frac{1}{2})|X(s)|^\alpha$$

$$\left. + (\alpha_2^2 + \frac{1}{2}\alpha_3^2) - \frac{\varrho_1^2}{2} \right] \mathrm{d}s$$

$$+ \int_0^t \varrho_1 \mathrm{d}B_1(t) + \theta\varepsilon \log(t + 1).$$

14

Letting ϵ be sufficiently small, by the definition of φ in (11), for sufficiently small $\epsilon \in (0, \varrho_1^2/2 - \varphi)$, we have

$$\log |X(t)| \leq \log |\xi| - \left[(\frac{\varrho_1^2}{2} - \varphi) - \epsilon \right] t + \int_0^t \varrho_1 \mathrm{d}B_1(t)$$
$$+ \theta \varepsilon \log(t+1).$$

Applying the strong law of large number, we therefore have

$$\limsup_{t \to \infty} \frac{\log |X(t)|}{t} \leq -(\frac{\varrho_1^2}{2} - \varphi) + \epsilon \quad a.s.$$

$\square$

Appendix C: Experimental Setups

Table 4: Detailed Experiment Setups of StyleGANv2-ada and its BMC trails

Dataset	Batch Size	Learning Rate	Optimizer	GPUs
CIFAR-10	64	0.0025	Adam	2
LSUN-Cat (256x256)	64	0.0025	Adam	4
LSUN-Bedroom (256x256)	64	0.0025	Adam	4
FFHQ (1024x1024)	32	0.0002	Adam	4

Table 5: Detailed Experiment Setups of ProjectedGAN and its BMC trails

Dataset	CIFAR-10
Batch Size	64
Learning Rate	0.002
Optimizer	Adam
Projected	EfficientNet

Appendix D: Results on ProjectedGAN and ImageNet

Table 6: FID scores on CIFAR-10 for ProjectedGAN with respect to the number of real images seen

	200k	400k	600k	800k	1M
Projected-GAN	12.77	9.40	8.17	7.34	7.01
Projected-GAN-BMC ($\varrho_1 = 1, \varrho_2 = 0.1$)	11.20	9.11	7.28	6.67	6.28
Projected-GAN-BMC ($\varrho_1 = 0.1, \varrho_2 = 0.01$)	12.57	9.31	8.17	7.31	6.74
Projected-GAN-BMC ($\varrho_1 = 0.01, \varrho = 0.001$)	12.17	8.92	7.93	6.42	6.39

Table 7: FID scores on ImageNet 32×32 with respect to the number of real images seen

	1M	2M	3M	4M	5M	6M
StyleGANv2-ada	28.67	18.93	15.94	14.17	13.46	12.45
StyleGANv2-ada-BMC ($\varrho_1 = 1, \varrho_2 = 0.1$)	21.41	16.10	13.76	12.25	11.48	10.93
StyleGANv2-ada-BMC ($\varrho_1 = 0.1, \varrho_2 = 0.01$)	25.34	17.31	14.84	12.89	12.10	11.79
StyleGANv2-ada-BMC ($\varrho_1 = 0.01, \varrho_2 = 0.001$)	26.85	17.89	15.46	13.50	12.88	12.31