Dynamic Data Generation and Dynamic Portfolio Selection

Ahmad Aghapour[1] Erhan Bayraktar[1] Fengyi Yuan[2]

[1]University of Michigan [2]The Chinese University of Hong Kong, Shenzhen

QF Conference 2026, NUS Singapore

June 2026

Contents

Dynamic mean-variance selection

- Static Markowitz:

$$\sup_{w \in \mathcal{K}} \left\{ \mathbb{E}[w^\top R] - \frac{\gamma}{2} \mathrm{Var}(w^\top R) \right\}.$$

- Dynamic mean-variance:

$$v^*(\mathbb{P}) = \sup_{\vartheta \in \Theta_{\mathcal{K},\mathbb{P}}} \left\{ \mathbb{E}_{\mathbb{P}}[(\vartheta \cdot S)_T] - \frac{\gamma}{2} \mathrm{Var}_{\mathbb{P}}[(\vartheta \cdot S)_T] \right\}.$$

- A model-based solution assumes the law $\mathbb{P}$ of the full price path $S^{1:T}$ is known.
- In practice, we only have limited historical samples and need a model-free scenario generator.

Question

Can a learned generative model $\mathbb{Q}$ replace the unknown market model $\mathbb{P}$ in a dynamic portfolio problem?

More generally, can we rely on generative models to make *dynamic* decisions?

Static v.s. dynamic generation

- Time-series data are paths:

$$X = (X^1, \dots, X^T) \sim \mathbb{P}.$$

- A dynamic decision problem needs more than a *joint* sampler: at each date t, the optimizer needs conditional information:

$$\mathbb{P}_{x^{1:t}} = \mathcal{L}(X^{t+1} \mid X^{1:t} = x^{1:t}).$$

- Treating $X^{1:T}$ as a single vector in $\mathbb{R}^{dT}$ ignores *nonanticipativity*.
- The downstream (dynamic) decision problems are generally NOT stable under ordinary Wasserstein distance: Wasserstein bounds and convergence in existing literature is not enough!

Instability of ordinary Wasserstein

real model $\mathbb{P}$

perturbed model $\mathbb{P}_n$

Figure: A Wasserstein-small perturbation creates a tradable signal at $t = 1$.

We have $\mathcal{W}_2(\mathbb{P}, \mathbb{P}_n) \leq \frac{1}{n} \to 0$, by transporting the two split nodes $S_1 = 1 \pm 1/n$ back to $S_1 = 1$.

Instability of ordinary Wasserstein

Take initial wealth 1 and choose shares $\theta = (\theta_0, \theta_1)$.

Under $\mathbb{P}$. The first-period position is irrelevant, and

$$X_2^\theta = \begin{cases} 1 + \theta_1, & 1/4, \\ 1, & 1/2, \\ 1 - \theta_1, & 1/4. \end{cases}$$

Thus

$$J_{\mathbb{P}}(\theta) = 1 - \frac{\gamma}{4}\theta_1^2, \qquad v^*(\mathbb{P}) = 1,$$

with optimal second-period exposure $\theta_1^* = 0$.

Under $\mathbb{P}_n$. Use the signal-following policy

$$\theta_1 = k \operatorname{sgn}(S_1 - 1).$$

Then

$$\mathbb{E}_{\mathbb{P}_n}[X_2^\theta] = 1 + \frac{k}{2} - \frac{k}{n}, \quad \mathrm{Var}_{\mathbb{P}_n}(X_2^\theta) = \frac{k^2}{4} + \frac{\theta_0^2}{n^2}.$$

Hence

$$v^*(\mathbb{P}_n) \geq 1 + \frac{k}{2} - \frac{k}{n} - \frac{\gamma}{2}\left(\frac{k^2}{4} + \frac{\theta_0^2}{n^2}\right).$$

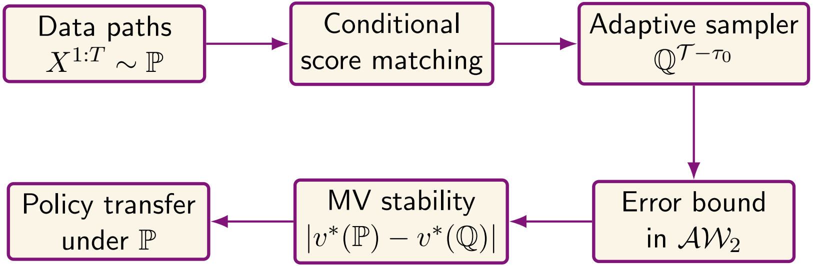
Conclusion

Taking $\theta_0 = 0$ and $k_n = 2(n-2)/(\gamma n)$ gives

$$v^*(\mathbb{P}_n) \geq 1 + \frac{(n-2)^2}{2\gamma n^2} \longrightarrow 1 + \frac{1}{2\gamma} > v^*(\mathbb{P}), \qquad k_n \to \frac{2}{\gamma}.$$

Even though $\mathcal{W}_2(\mathbb{P}, \mathbb{P}_n) \to 0$, we are always misled by the generative model to make a strictly suboptimal decision.

Roadmap



Data paths $X^{1:T} \sim \mathbb{P}$

Conditional score matching

Adaptive sampler $\mathbb{Q}^{\mathcal{T}-\tau_0}$

Error bound in $\mathcal{AW}_2$

MV stability $|v^*(\mathbb{P}) - v^*(\mathbb{Q})|$

Policy transfer under $\mathbb{P}$

Goals:

① Some dynamic generation methods which have $\mathcal{AW}_2(\mathbb{P}, \mathbb{Q})$ bounds.

② Some policy-transfer results using $\mathcal{AW}_2$ bounds.

Contents

Static score-based diffusion model

- Forward Ornstein–Uhlenbeck process:

$$\mathrm{d}X_\tau = -X_\tau\,\mathrm{d}\tau + \sqrt{2}\,\mathrm{d}B_\tau, \qquad X_0 \sim p_{\mathrm{data}}.$$

- For large τ, X_τ is close to a standard Gaussian.
- Reverse-time process:

$$\mathrm{d}\bar{X}_\tau = \left[\bar{X}_\tau + 2\nabla_x \log p(\mathcal{T} - \tau, \bar{X}_\tau)\right]\mathrm{d}\tau + \sqrt{2}\,\mathrm{d}\bar{B}_\tau.$$

- In practice, the unknown score $\nabla_x \log p$ is replaced by a trained score network s_θ.

Static guarantee

Small score error + large diffusion time $\longrightarrow$ a distribution close to the data law.

A dynamic version of score-based diffusion model

- For $t = 1, \dots, T-1$, condition on the observed history $x^{1:t}$.
- Conditional forward process:

$$\mathrm{d}X_\tau^{t+1} = -X_\tau^{t+1}\,\mathrm{d}\tau + \sqrt{2}\,\mathrm{d}B_\tau^{t+1}, \qquad X_0^{t+1} \sim \mathbb{P}_{x^{1:t}}.$$

- Conditional score:

$$s^{t+1}(\tau, x^{1:t}, x) := \nabla_x \log p_{t+1}(\tau, x \mid x^{1:t}).$$

- A score network s_θ^{t+1} should approximate this score in some sense.

Challenge

We do not observe samples directly from $\mathbb{P}_{x^{1:t}}$, only full historical trajectories from $\mathbb{P}$.

Conditional denoising score matching

Averaged score-matching error

$$\mathbb{E}_{X^{1:t}\sim\mathbb{P}_{1:t}}\mathbb{E}_{X_\tau^{t+1}\sim p_{t+1}(\tau,\cdot|X^{1:t})}\left|s_\theta^{t+1}(\tau,X^{1:t},X_\tau^{t+1})-\nabla_x\log p_{t+1}(\tau,X^{1:t},X_\tau^{t+1})\right|^2\le\epsilon_{\text{score}}^2.$$

Equivalent denoising objective

$$\min_\theta \mathbb{E}_{X^{1:t+1}\sim\mathbb{P}_{1:t+1}}\mathbb{E}_{Z\sim N(0,I)}\left|\sqrt{1-e^{-2\tau}}\,s_\theta^{t+1}\Big(\tau,X^{1:t},X^{t+1}e^{-\tau}\sqrt{1-e^{-2\tau}}Z\Big)+Z\right|^2.$$

- The second objective is trainable from joint samples.
- This is the conditional analogue of classical denoising score matching.

Adaptive sampling

The algorithm

1. Generate Y^1 from noise using s_θ^1.
2. Given the generated history $Y^{1:t} = y^{1:t}$, generate the next state via

$$\mathrm{d}\bar{Y}_\tau = \left[\bar{Y}_\tau + 2 s_\theta^{t+1}(\mathcal{T} - \tau, y^{1:t}, \bar{Y}_\tau)\right] \mathrm{d}\tau + \sqrt{2}\, \mathrm{d}\bar{B}_\tau.$$

3. Repeat for $t = 1, \ldots, T-1$.

- The resulting joint law is denoted by $\mathbb{Q}^{\mathcal{T}-\tau_0}$.
- The same procedure gives conditional samplers $\mathbb{Q}_{y^{1:t}}^{\mathcal{T}-\tau_0}$.
- Early stopping at $\tau_0 > 0$ avoids assuming that $\mathbb{P}$ has a joint density.

Adapted Wasserstein distance

Bicausal couplings

A coupling π of $\mathbb{P}$ and $\mathbb{Q}$ is causal in x if, for each t,

$$\pi(Y^t \in \mathrm{d}y^t \mid X^{1:T} = x^{1:T}) = \pi(Y^t \in \mathrm{d}y^t \mid X^{1:t} = x^{1:t}).$$

It is bicausal if the same condition also holds with x and y interchanged.

Metric

$$\mathcal{AW}_2^2(\mathbb{P}, \mathbb{Q}) = \inf_{\pi \in \Pi_{\mathrm{bc}}(\mathbb{P}, \mathbb{Q})} \int |x^{1:T} - y^{1:T}|^2 \, \pi(\mathrm{d}x^{1:T}, \mathrm{d}y^{1:T}).$$

- Π_{bc} only allows transport plans compatible with sequential information.
- Ordinary $\mathcal{W}_2$ can be small even when the dynamic decision problem changes.

Main theorem

Theorem: adapted-Wasserstein error bound

Under the data-distribution assumptions, score-network assumptions, and score-matching error bound, the adaptive sampler satisfies

$$\begin{aligned}\mathcal{AW}_2^2(\mathbb{P}, \mathbb{Q}^{\mathcal{T}-\tau_0}) \leq (K')^T \tau_0 + C_{\tau_0} \Big(& T^2 D^3_{\tau_0,\epsilon_{\text{score}},T} \mathcal{T} e^{-\mathcal{T}/2^{T-1}} \\ & + T^2 D^2_{\tau_0,\epsilon_{\text{score}},T} e^{-\frac{c'}{2^{T-1}}\sqrt{D_{\tau_0,\epsilon_{\text{score}},T}}} + T^2 D^3_{\tau_0,\epsilon_{\text{score}},T} \mathcal{T}^{1+1/2^T} \epsilon_{\text{score}}^{1/2^{T-1}} \Big).\end{aligned}$$

- τ_0 is the early-stopping level; $\mathcal{T}$ is the diffusion horizon.
- $D_{\tau_0,\epsilon_{\text{score}},T}$ is the cutoff/dissipativity scale induced by the network construction; it does not depend on the diffusion horizon $\mathcal{T}$.
- The right-hand side can be made arbitrarily small by choosing $\tau_0 \downarrow 0$, $\mathcal{T} \uparrow \infty$, and $\epsilon_{\text{score}} \downarrow 0$ on compatible scales.

Main theorem

- **Joint path generation:**

$$Y^{1:T} \sim \mathbb{Q}^{\mathcal{T}-\tau_0} \quad \text{is close to} \quad X^{1:T} \sim \mathbb{P} \quad \text{in } \mathcal{AW}_2.$$

- **Conditional generation:**

$$\mathbb{Q}^{\mathcal{T}-\tau_0}_{y^{1:t}} = \mathcal{L}(Y^{t+1} \mid Y^{1:t} = y^{1:t})$$

  is sampled by the same reverse SDE with $y^{1:t}$ plugged into s_θ^{t+1}.
- The $\mathcal{AW}_2$ metric matches nonanticipative decision rules, supplying the right error control for downstream dynamic decision problems.
- The learned model has similar *conditional kernals* to the data distribution.

Proof sketch

Conditional estimate

For histories $x^{1:t}$ and $y^{1:t}$,

$$
\begin{aligned}
\mathcal{W}_2^2(\mathbb{P}_{x^{1:t}}, \mathbb{Q}_{y^{1:t}}^{\mathcal{T}-\tau_0}) \leq \tau_0\big(2d + M_2^2(x^{1:t})\big) \\
+ C_{\tau_0}\Big(\alpha(\tau_0, \mathcal{T}, \epsilon_{\text{score}}, T) + D_{\tau_0,\epsilon_{\text{score}},T}\,\mathcal{E}(\tau_0, x^{1:t})^{1/2} \\
+ e^{-c'\sqrt{D_{\tau_0,\epsilon_{\text{score}},T}}}(1 + E_c(x^{1:t})) + D_{\tau_0,\epsilon_{\text{score}},T}\,\mathcal{T}^{1/2}(1 + M_2(x^{1:t}))|x^{1:t} - y^{1:t}|\Big).
\end{aligned}
$$

- TV/KL control comes from Girsanov for the reverse SDEs.
- Wasserstein control follows from TV plus tail bounds.
- Dissipativity of s_θ yields uniform tail estimates for generated conditional laws.

Proof sketch

① Build a bicausal coupling recursively:

$$\pi(\mathrm{d}x^{1:T}, \mathrm{d}y^{1:T}) = \pi_1(\mathrm{d}x^1, \mathrm{d}y^1) \prod_{t=1}^{T-1} \pi_{x^{1:t}, y^{1:t}}(\mathrm{d}x^{t+1}, \mathrm{d}y^{t+1}).$$

② At each step, choose $\pi_{x^{1:t}, y^{1:t}}$ close to an optimal coupling between $\mathbb{P}_{x^{1:t}}$ and $\mathbb{Q}_{y^{1:t}}^{\mathcal{T}-\tau_0}$.

③ Iterate the conditional estimate to control

$$\mathbb{E}_\pi |X^{1:T} - Y^{1:T}|^2.$$

When the horizon dependence improves

Additional contraction structure

If conditional kernels are Markovian and the learned kernels satisfy

$$\mathcal{W}_2(\mathbb{Q}_x^{\mathcal{T}-\tau_0}, \mathbb{Q}_y^{\mathcal{T}-\tau_0}) \le \kappa|x-y|, \qquad \kappa \in (0,1),$$

define the one-step conditional error

$$\Delta_{\tau_0,\epsilon_{\text{score}},T} = \mathcal{W}_2^2(\mathbb{P}_1, \mathbb{Q}_1^{\mathcal{T}-\tau_0}) + \max_t \mathbb{E}_{X^t \sim \mathbb{P}_t}\left[\mathcal{W}_2^2(\mathbb{P}_{X^t}, \mathbb{Q}_{X^t}^{\mathcal{T}-\tau_0})\right].$$

Then

$$\mathcal{AW}_2^2(\mathbb{P}, \mathbb{Q}^{\mathcal{T}-\tau_0}) \le C_\kappa T \Delta_{\tau_0,\epsilon_{\text{score}},T}.$$

- The iterated exponent disappears under genuine forgetting/contractivity.
- The general theorem should be read as a worst-case bound.

Contents

Mean-variance and quadratic hedging

- Terminal trading gain:

$$(\vartheta\cdot S)_T = \sum_{t=1}^{T-1} \vartheta_t^\top (S^{t+1} - S^t), \qquad \vartheta_t \in \mathcal{K}.$$

- Mean-variance value:

$$v^*(\mathbb{P}) = \sup_{\vartheta\in\Theta_{\mathcal{K},\mathbb{P}}} \left\{\mathbb{E}_{\mathbb{P}}[(\vartheta\cdot S)_T] - \frac{\gamma}{2}\operatorname{Var}_{\mathbb{P}}[(\vartheta\cdot S)_T]\right\}.$$

- Auxiliary quadratic-hedging value:

$$V(\mathbb{P}, c) = \min_{\vartheta\in\Theta_{\mathcal{K},\mathbb{P}}} \mathbb{E}_{\mathbb{P}} \left|c - (\vartheta\cdot S)_T\right|^2 .$$

Dual representation

$$v^*(\mathbb{P}) = \sup_{a\geq 0}\left\{-\frac{\gamma}{2}V\left(\mathbb{P}, \frac{1}{\gamma} + a\right) + \frac{1}{2\gamma} + a\right\}.$$

Dynamic programming for the dual problem

Conditional value

$$V(t,w,s^{1:t};c,\mathbb{P}) = \inf_{\vartheta\in\Theta^{t,s^{1:t}}_{\mathcal{K},\mathbb{P}}} \mathbb{E}_{\mathbb{P}_{s^{1:t}}}\left[\left|c-w-\sum_{\ell=t}^{T-1}\vartheta_\ell^\top \Delta S^\ell\right|^2\right].$$

DPP

$$V(T,w,s^{1:T}) = |c-w|^2,$$
$$V(t,w,s^{1:t}) = \inf_{\theta\in\mathcal{K}} \mathbb{E}_{\mathbb{P}_{s^{1:t}}}\left[V(t+1,w+\theta^\top\Delta S^t,(s^{1:t},S^{t+1}))\right].$$

- This avoids the time-inconsistency of the primal MV objective.
- Bounded convex constraints ensure compactness and existence of optimizers.

Stability of the dual value

Theorem: DPP stability

For any bicausal coupling $\pi \in \Pi_{\mathrm{bc}}(\mathbb{P}, \mathbb{Q})$,

$$\begin{aligned} |V(t, w, s^{1:t}; c, \mathbb{P}) - V(t, \tilde{w}, \tilde{s}^{1:t}; c, \mathbb{Q})| \le C\Big\{ & \big(|w - \tilde{w}| + |s^t - \tilde{s}^t|\big) B_t \\ & + \big(B_t + |w - \tilde{w}| + |s^t - \tilde{s}^t|\big) D_t^\pi \\ & + (D_t^\pi)^2 \Big\}, \end{aligned}$$

where D_t^π measures the conditional future path discrepancy under π. Consequantly,

$$|V(1, 0, s^1; c, \mathbb{P}) - V(1, 0, s^1; c, \mathbb{Q})| \le C\Big((1 + s^1 + c + M_2^{\mathbb{P}}(s^1))\mathcal{AW}_2(\mathbb{P}, \mathbb{Q}) + \mathcal{AW}_2^2(\mathbb{P}, \mathbb{Q}) \Big).$$

Stability of mean-variance values

Corollary

If $\mathcal{AW}_2(\mathbb{P},\mathbb{Q}) < 1$, then

$$|v^*(\mathbb{P}) - v^*(\mathbb{Q})| \le C\,\mathcal{AW}_2(\mathbb{P},\mathbb{Q}).$$

- The duality reduces MV stability to stability of $V(\cdot, c)$.
- A uniform growth bound confines the optimizing multiplier

  $$c = \frac{1}{\gamma} + a$$

  to a compact interval independent of $\mathbb{Q}$ when $\mathcal{AW}_2(\mathbb{P},\mathbb{Q}) < 1$.
- Therefore, the learned generative model can be used as a surrogate environment with controlled value loss.

Policy transfer

Regular feedback policies

Let $\phi = \{\phi_t\}_{t=1}^{T-1}$ be a Lipschitz feedback policy,

$$\phi_t : \mathbb{R} \times \mathbb{R}^{dt} \to \mathcal{K},$$

and define wealth pathwise by

$$(\phi \cdot s)_{t+1} = (\phi \cdot s)_t + \phi_t((\phi \cdot s)_t, s^{1:t})^\top (s^{t+1} - s^t).$$

Performance transfer

If ϕ^ϵ is ϵ-suboptimal under $\mathbb{Q}$, then

$$v(\mathbb{P}; \phi^\epsilon) \geq v^*(\mathbb{P}) - \epsilon - C\, \mathcal{AW}_2(\mathbb{P}, \mathbb{Q}).$$

- The generative error bound is translated into downstream portfolio performance control.
- Lipschitz regularity is the transfer condition for comparing a fixed feedback policy across models.

Contents

Conclusion

- **Dynamic generation.** Adaptive conditional score matching turns joint path data into a sequential conditional generator $\mathbb{Q}^{\mathcal{T}-\tau_0}$.
- **Error control.** The approximation is controlled in adapted Wasserstein distance, the metric compatible with nonanticipative decisions.
- **Decision stability.** Dynamic mean-variance values are stable in $\mathcal{AW}_2$ through the quadratic-hedging dual and DPP.
- **Empirical evidence.** Synthetic experiments show effective learning from sampled scenarios. On real industry-portfolio data, diffusion-based strategies outperform the S&P 500, equal weight, and historical Markowitz benchmarks in key risk-adjusted metrics; GenTD3 is more stable across risk aversion levels.

Conclusion

The diffusion model is useful for dynamic portfolio selection because it approximates the market law in the right dynamic sense, not merely because it produces plausible sample paths.

Future Directions

Theory

- Relax assumptions on the data distribution, especially the Lipschitz condition in the history variable $x^{1:t}$.
- Sharpen the $\mathcal{AW}_2$ error rate and understand whether the current bounds are optimal.

Methodology and Experiments

- Improve practical performance through larger-scale experiments, algorithm tuning, and richer covariates in the RNN encoder.
- Develop quantitative guarantees for the RL component, such as regret analysis.

Thank you

Questions?