

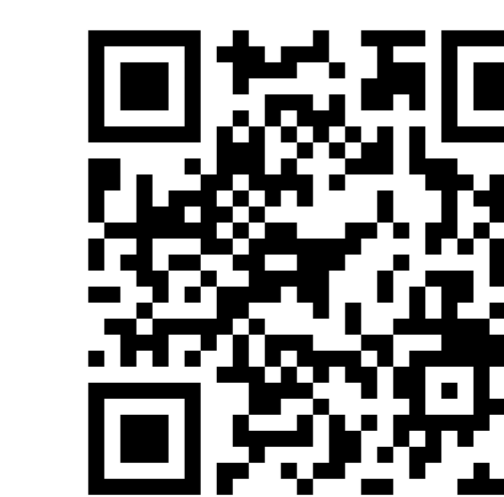


# Wasserstein Distributionally Robust Regret-Optimal Control Over an Infinite-Horizon

Taylan Kargin\*, Joudi Hajar\*, Vikrant Malik\*, Babak Hassibi



arXiv:2312.17376



arXiv:2406.07248

## 1. Introduction

➤ **Setting:** Full-information linear time-invariant (LTI) control

• Linear plant dynamics:  $\mathbf{x} = \mathcal{F}\mathbf{u} + \mathcal{G}\mathbf{w}$

• Disturbance feedback control:  $\mathbf{u} = \mathcal{K}\mathbf{w}$

• Quadratic cost:  $\text{COST}(\mathbf{u}, \mathbf{w}) := \|\mathbf{x}\|^2 + \|\mathbf{u}\|^2$

➤ **Goal:** design a *computationally efficient* robust controller under *distributional uncertainty* of exogenous disturbances

➤ **Performance metric:** regret against non-causal controller

$\text{REGRET}(\mathbf{u}, \mathbf{w}) := \text{COST}(\mathbf{u}, \mathbf{w}) - \min \text{COST}(\mathbf{u}_{\text{nc}}, \mathbf{w})$

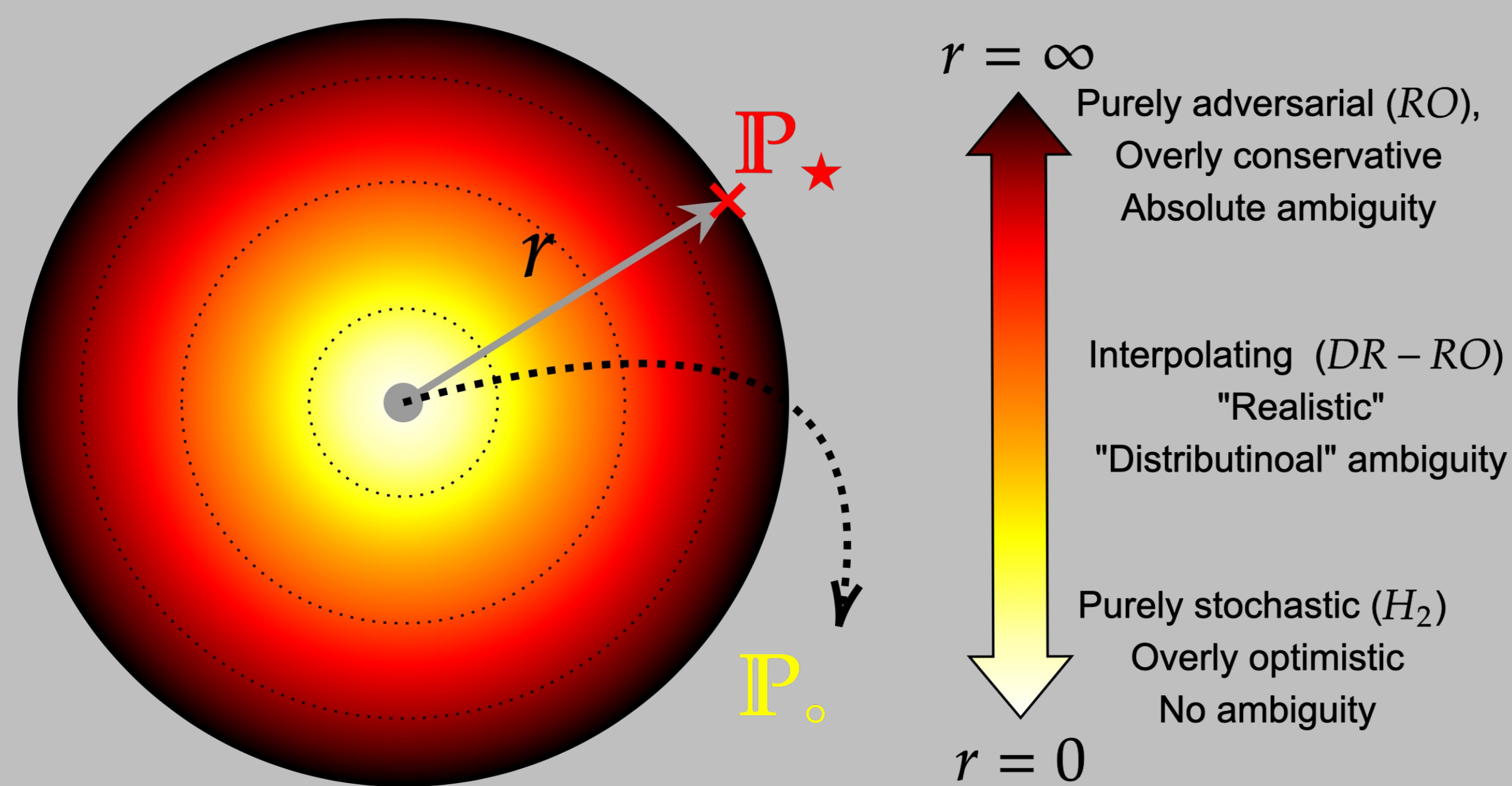
$$\mathbf{u}_{\text{nc}} = \underbrace{-(\mathcal{I} + \mathcal{F}^\dagger \mathcal{F})^{-1} \mathcal{F}^\dagger \mathcal{G}}_{\text{"best" non-causal controller } \mathcal{K}_{\text{nc}}} \mathbf{w}$$

➤ **Disturbances:**

• Unknown disturbance dist.:  $\mathbf{w} \sim \mathbb{P}$ ,  $\text{Cov}[\mathbf{w}] = \mathcal{M}$

• Known nominal dist.:  $\mathbf{w}_o \sim \mathbb{P}_o$ ,  $\text{Cov}[\mathbf{w}_o] = \mathcal{M}_o$

• Known Wasserstein-2 distance:  $W_2(\mathbb{P}, \mathbb{P}_o) \leq r$



➤ **Prior works:** Taskesen et al. 2023 restrict to time-independent disturbances, Taha et al. 2023 have similar setup but only for the finite-horizon setting

➤ **Challenge:** *Finite-horizon DR-RO controller requires solving an SDP scaling with time-horizon.* Therefore, we seek an infinite-horizon DR-RO controller.

➤ **Approach:** Take the limit of finite-horizon problem and formulate as *convex-concave optimization over Toeplitz, positive-definite, autocovariance operators.*

## 2. Main Results

➤ **Primal (P) problem:**

$$(P) \quad \inf_{\text{causal } \mathcal{K}} \sup_{\mathbb{P}: W_2(\mathbb{P}, \mathbb{P}_o) \leq r} \mathbb{E}_{\mathbf{w} \sim \mathbb{P}} [\text{REGRET}(\mathcal{K}\mathbf{w}, \mathbf{w})]$$

➤ (P) can be simplified since regret admits a quadratic form:

$$\text{REGRET}(\mathcal{K}\mathbf{w}, \mathbf{w}) = \mathbf{w}^\dagger \underbrace{(\Delta\mathcal{K} - \Delta\mathcal{K}_{\text{nc}})^\dagger (\Delta\mathcal{K} - \Delta\mathcal{K}_{\text{nc}})}_{\text{regret operator } \mathcal{R}_{\mathcal{K}}} \mathbf{w}$$

where  $\Delta^\dagger \Delta = \mathcal{I} + \mathcal{F}^\dagger \mathcal{F}$  is the *spectral factorization*.

**Theorem 1.** Problem (P) is equivalent to problem (D) below

$$(D) \quad \sup_{\substack{\mathcal{M} \succ 0, \\ \text{BW}(\mathcal{M}, \mathcal{M}_o) \leq r}} \inf_{\text{causal } \mathcal{K}} \text{Tr}(\mathcal{R}_{\mathcal{K}} \mathcal{M})$$

The saddle point  $(\mathcal{K}_*, \mathcal{M}_*)$  satisfies:

$$i) \quad \mathcal{K}_* = \underbrace{\mathcal{K}_o}_{\text{Nominal Policy}} + \underbrace{\Delta^{-1} \{ \{ \Delta\mathcal{K}_{\text{nc}} \mathcal{L}_o \}_- \mathcal{L}_o^{-1} \mathcal{L}_* \}_+ \mathcal{L}_*^{-1}}_{\text{Additive Corrective Term}}$$

$$ii) \quad \mathcal{M}_* = \underbrace{(\mathcal{I} - \gamma_*^{-1} \mathcal{R}_{\mathcal{K}_*})^{-1}}_{\text{Optimal Transport Map}} \underbrace{\mathcal{M}_o}_{\text{Nominal Autocovariance}} \underbrace{(\mathcal{I} - \gamma_*^{-1} \mathcal{R}_{\mathcal{K}_*})^{-1}}_{\text{Worst-case Autocovariance}}$$

where  $\mathcal{M}_* = \mathcal{L}_* \mathcal{L}_*^\dagger$  and  $\mathcal{M}_o = \mathcal{L}_o \mathcal{L}_o^\dagger$  are Wiener-Hopf spectral factorizations and  $\gamma_* > 0$  is such that  $\text{BW}(\mathcal{M}_*, \mathcal{M}_o) = r$ .

➤ **Worst-case disturbance:**  $\mathbf{w}_* = (\mathcal{I} - \gamma_*^{-1} \mathcal{R}_{\mathcal{K}_*})^{-1} \mathbf{w}_o$

**Theorem 2.** The worst-case covariance  $\mathcal{M}_*$  and the optimal DR-RO controller  $\mathcal{K}_*$  are *non-rational*. Thus, optimal DR-RO controller does *not* admit a finite-order state-space realization.

## 3. Algorithm

Current PSD (eg. autocovariance) sets the disturbance distribution.

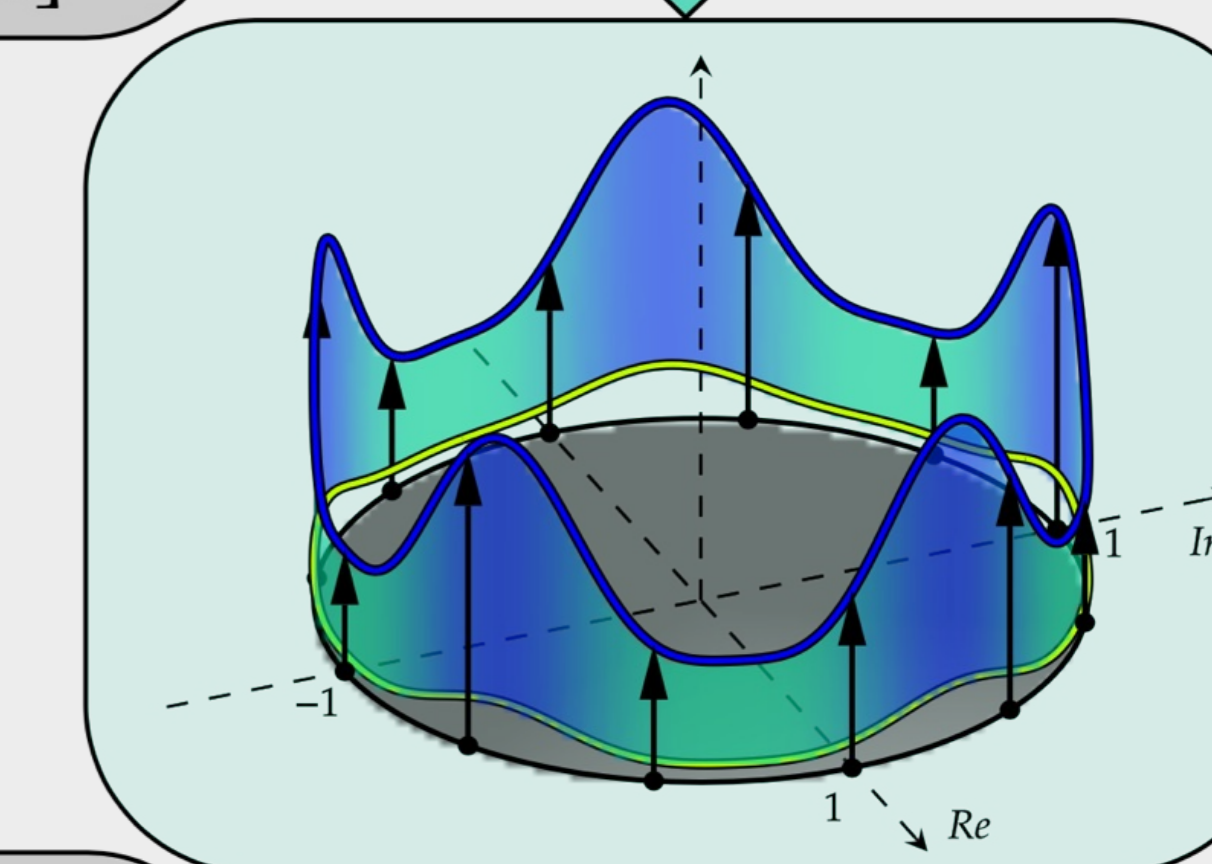
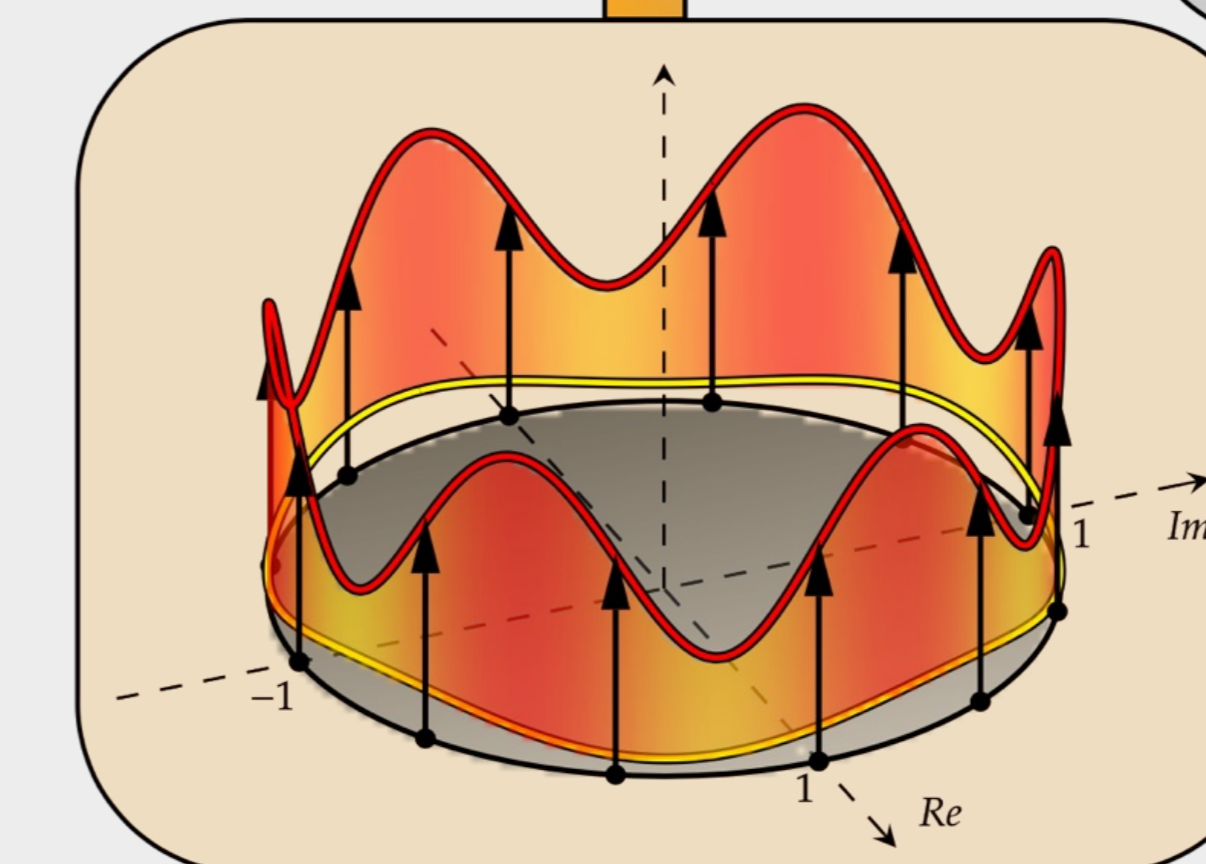
Wiener-Hopf Spectral Factorization

$$M_k = L_k L_k^\dagger$$

Controller's (optimal) response to current disturbance PSD

$$M_k(e^{j\omega_n})$$

$$R_k(e^{j\omega_n})$$



First-Order (Frank-Wolfe) Update

$$M_{k+1}(e^{j\omega_n})$$

Next iterate enters the loop again.

Regret operator propagates gradient information.

Rational Approximation

$$\min_{\substack{p_0, \dots, p_m \in \mathbb{R}, \\ q_0, \dots, q_m \in \mathbb{R}}} \max_{z \in \mathbb{T}_N} \left| \frac{P(z)}{Q(z)} - M_k(z) \right|$$

$$s.t. \quad i) P(z) = \sum_{i=-m}^m p_m z^{-i} > 0, \forall z \in \mathbb{T},$$

$$ii) Q(z) = \sum_{i=-m}^m q_m z^{-i} > 0, \forall z \in \mathbb{T}$$

Best order-m state-space controller

$$\hat{\mathcal{K}}: \begin{cases} \xi_{t+1} = \hat{F}\xi_t + \hat{G}w_t \\ u_t = \hat{H}\xi_t \end{cases}$$

Final (non-rational) iterate is approximated by rational functions.

## 4. Numerical Simulations

