Statistical Results on Filtering and Epi-convergence for Learning-Based Model Predictive Control

Anil Aswani
Humberto Gonzalez
S. Shankar Sastry
Claire Tomlin

Electrical Engineering and Computer Sciences
University of California at Berkeley

December 17, 2011

Acknowledgement

Statistical Results on Filtering and Epi-convergence for Learning-Based Model Predictive Control

Anil Aswani, Humberto Gonzalez, S. Shankar Sastry, and Claire Tomlin

December 17, 2011

Abstract

Learning-based model predictive control (LBMPC) is a technique that provides deterministic guarantees on robustness, while statistical identification tools are used to identify richer models of the system in order to improve performance. This technical note provides a result that elucidates the reasons for the choice of measurement model used with LBMPC, and it gives proofs concerning the stochastic convergence of LBMPC. The first part of this note discusses simultaneous state estimation and statistical identification (or learning) of unmodeled dynamics, for dynamical systems that can be described by ordinary differential equations (ODE's). The second part provides proofs concerning the epi-convergence of different statistical estimators that can be used with the LBMPC technique. In particular, we prove results on the statistical properties of a nonparametric estimator that we have designed to have the correct deterministic and stochastic properties for numerical implementation when used in conjunction with LBMPC.

1   Introduction

This technical note is meant to be understood in the context of [3], and it consists of two distinct parts. Sections 2 and 3 concern simultaneous state estimation and statistical identification (or learning) of unmodeled dynamics, for dynamical systems that can be described by ordinary differential equations (ODE's). The second part is found in Section 4 and provides proofs concerning the epi-convergence of different statistical estimators that can be used with the learning-based model predictive control (LBMPC) technique.

For the results on estimation and learning, we assume that for state vector $x \in \mathbb{R}^p$, control input $u \in \mathbb{R}^m$, and output $y \in \mathbb{R}^q$, the system dynamics are given by the following ODE:

$$\begin{aligned} \dot{x} &= A_c x + B_c u + g_c(x, u) \\ y &= Cx, \end{aligned} \tag{1}$$

where A_c, B_c, C are matrices of appropriate size and $g_c(x, u)$ describes the unmodeled (possibly nonlinear) dynamics. We will assume that the control inputs generated by the model predictive control (MPC) schemes are piecewise constant

$$u(t) = u_n, \qquad \forall t \in \big[n T_u, (n+1) T_u \big), \tag{2}$$

1

where T_u is the sampling period of the input. Note that appropriately designed MPC can generate other control schemes, such as piecewise linear inputs.

2 Limitations on Filtering and Learning

We begin with a negative result about the inability of filters to estimate both the state and unmodeled dynamics, for a general system in which all states are not measured. (This result does not apply to systems with special structure, such as in [2].) This limitation applies to situations in which unmodeled dynamics are described by a series expansion with constant terms, and so it is relevant to a wide class of systems and filtering approaches.

Suppose the unmodeled dynamics are parameterized as $g_c(x, u) = \gamma(x, u; \theta) + K$, where K is a constant, non-zero vector and $\gamma(x, u; \theta)$ is a parametrized function such that $\gamma(x, u; \theta_0) \equiv 0$ for some parameter value θ_0. We again note that this includes the situation in which g_c is given by a series expansion (e.g., Taylor polynomial, Fourier series, etc.).

The intuition is that statistical identification (or learning) of the parameters θ, K and estimation of the state x can be cast into the framework of observability of an augmented dynamical system. The augmented system has $y = Cx$ and dynamics

$$
\begin{bmatrix} \dot{x} \\ \dot{K} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} A_c x + B_c u + \gamma(x, u; \theta) + K \\ 0 \\ 0 \end{bmatrix}.
\tag{3}
$$

When all states are not measured and there is no special structure on K, then this augmented system is not observable. This means that (x, K, λ) cannot be simultaneously estimated using measurements of the system output y. This is formalized by the following theorem.

Theorem 1. *A necessary condition for the observability (and detectability) of the system given in (3) with $y = Cx$ is that $rank(C) = p$.*

Proof. Suppose $\theta = \theta_0$, which makes $\gamma(x, u; \theta) \equiv 0$. Then the system is linear and time-invariant (LTI). Using the Popov-Belevitch-Hautus (PBH) test, the system is observable if and only if $rank(\phi) = p + p = 2p$, for all $s \in \mathbb{C} : Re(s) \geq 0$, where

$$
\phi = \begin{bmatrix} s\mathbb{I} - A_c & -\mathbb{I} \\ 0 & s\mathbb{I} \\ C & 0 \end{bmatrix}.
\tag{4}
$$

If $s = 0$, then the matrices ϕ and $s\mathbb{I}$ are both singular, and the block structure of ϕ implies that $rank(\phi) = p + rank(C)$. The system is not observable (and not detectable) when $rank(C) < p$, establishing necessity. $\qquad\square$

Remark. This result also applies to discrete time systems, and the proof is nearly identical.

In light of this negative result concerning filters, we require that C be full rank. Without loss of generality, we assume that the full state x is measured.

3 Nonparametric Filtering for Dynamical Systems

The design of a Kalman filter for systems with unmodeled dynamics can be complex, and so we propose a nonparametric regression approach for estimating the state. Available approaches include local polynomial regression (LPR) or spline-smoothing; the Savitzky-Golay filter [13] is technically a finite impulse response (FIR) filter implementation of LPR. We design a new nonparametric filter, and one advantage is of this filter is that it is easily computed because it is the weighted sums of measurements.

An important point to note is that the statistical guarantees provided by our filter are not the same as for a Kalman filter. The Kalman filter is defined to be consistent if its state estimates are unbiased and the true error covariance is smaller (covariance matrices are positive semi-definite, and so a partial order can be defined) than the estimated error covariance. In our method, consistency is defined with respect to the sampling period T_s of state measurements. As $T_s \to 0$, the estimates converge to the real values in probability. This philosophical change is necessary in order to use nonparametric statistics, otherwise we would be forced to use a parametric model of the unmodeled dynamics.

We begin with a lemma about the differentiability of the state trajectory $x(t)$ when the inputs are piecewise constant.

Lemma 1. *Suppose $g_c(x, u)$ is $Q-1$-times differentiable. For $n \in \mathbb{Z}$, the trajectory $x(t)$ which solves the ODE in (1) is once-differentiable everywhere, Q-times differentiable at $t \neq nT_u$, and not twice-differentiable at $t = nT_u$.*

Proof. The first time-derivative of $x(t)$ is given by (1), by definition. Because the inputs are piecewise constant (2), the input $u(t)$ is not differentiable at $t = nT_u$. Because the first time-derivative of $x(t)$ is a function of $u(t)$, this means that $x(t)$ is not twice-differentiable at $t = nT_u$. Recall that $u(t)$ is constant for $t \neq nT_u$. Thus, $u(t)$ is smooth at $t \neq nT_u$. This implies that $x(t)$ is Q-times differentiable at $t \neq nT_u$, because $g_c(x, u)$ is $Q - 1$-times differentiable. $\square$

Remark. These qualitative features mean that we cannot use LPR methods with order higher than zero (i.e., the Nadaraya-Watson estimator) without modifying the filtering scheme. This is an important point, because the differentiability of the trajectory $x(t)$ makes it tempting to use LPR. Yet, no theoretical convergence guarantees can currently be made in such a situation, and the behavior of these filters may be unpredictable.

In light of these restrictions, we propose a modified sampling scheme. Recall that T_u is the sampling time for control inputs, and we define T_s to be the sampling time for state measurements. We require that $kT_s = T_u$ for some $k \in \mathbb{Z}_+$, and this scheme is illustrated in Fig. 1 for the case of $k = 4$. The advantage of this sampling scheme is that the trajectory $x(t)$ is piecewise smooth (infinitely differentiable) in between the samples taken at nT_u, because the control input $u(t)$ is piecewise constant. This allows us to use LPR of order higher than zero (e.g., local linear regression), which can give significant improvements in estimation error over zeroth order LPR.

If the trajectory of the real system is $x(t)$, then consider a measurement model

$$\xi_i(jT_s + nT_u) = x_i(jT_s + nT_u) + \epsilon_i, \qquad j \in \mathbb{Z}_+ : jT_s \in [0, T_u], \tag{5}$$
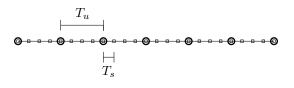
3

Figure 1: We use a sampling scheme with two sampling periods. The inputs change at every T_u units of time, and the states are measured every T_s units of time. In this example, $kT_s = T_u$ with $k = 4$.

where ϵ_i are independent and identically distributed (i.i.d.) random variables with zero mean and bounded values $l_\mu \leq [\epsilon_i]_\mu \leq s_\mu$. The notation $[\epsilon_i]_\mu$ indicates the μ-th component of the i-th noise vector. Suppose that we have made measurements for $n = 0, \ldots, N$. This measurement model corresponds to the sampling scheme seen in Figure 1.

3.1 Filter Design

Suppose $\kappa(\nu)$ is a kernel function, which is a bounded even function with finite support. We will use λ, ρ to denote left and right differentiability, and r is the polynomial order of the filter. Let $h_{\lambda;n,i}, h_{\rho;n,i} \in \mathbb{R}$ be bandwidth parameters. Next we define a diagonal matrix $R_{n,i}$ that is used to filter to the right side of the i-th entry of the measurement at $t = nT_u$; its entries are given by

$$R_{n,i} = \text{diag}\{\kappa(0), \kappa(T_s/h_{\lambda;n,i}), \ldots, \kappa(kT_s/h_{\lambda;n,i})\}. \tag{6}$$

Similarly, we define a diagonal matrix $L_{n,i}$ that is used to filter to the left side of the i-th entry of the measurement at $t = nT_u$:

$$L_{n,i} = \text{diag}\{\kappa(kT_s/h_{\rho;n,i}), \ldots, \kappa(T_s/h_{\rho;n,i}), \kappa(0)\}. \tag{7}$$

Note that the $R_{n,i}$ matrix uses the bandwidth $h_{\lambda;n,i}$, and $L_{n,i}$ uses bandwidth $h_{\rho;n,i}$. The reason is that filtering to the right of a measurement requires left differentiability, while filtering to the left of a measurement requires right differentiability. Lastly, we define the Vandermonde matrix

$$\Gamma = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 1 & T_s & \ldots & T_s^r \\ \vdots & \vdots & \vdots & \vdots \\ 1 & kT_s & \ldots & k^r T_s^r \end{bmatrix}. \tag{8}$$

We are now ready to design the filter. The filter coefficients are given by

$$\begin{aligned} w_{n,i} &= e_1'(\Gamma' L_{n,i}\Gamma)^{-1}\Gamma' L_{n,i} \\ v_{n,i} &= e_1'(\Gamma' R_{n,i}\Gamma)^{-1}\Gamma' R_{n,i}, \end{aligned} \tag{9}$$

and e_1 is the unit-vector with a 1 in the first position and zeros everywhere else. The idea is that $w_{n,i}$ filters on the left side of $t = nT_u$ and $v_{n,i}$ filters on the right side of $t = nT_u$. As time advances to $t = NT_u$, we first filter on the left side of $\xi(NT_u)$ (because there is

4

no right side). At the next point in time $t = (N+1)T_u$ we filter on both sides of $\xi(NT_u)$. Consequently, the filter is time-varying.

Let the number within the angled brackets $\langle \cdot \rangle$ denote the (discrete) time at which the filter is computed. The raw state estimates (for times $t = nT_u$, for $n = 0, \ldots, N$) computed at time $t = NT_u$ are given by

$$
\begin{aligned}
[\bar{x}_N]_i\langle N \rangle &= \sum_{j=0}^{k}[w_{N-1,i}]_j\xi_i(jT_s + (N-1)T_u) \\
[\bar{x}_{N-1}]_i\langle N \rangle &= 1/2 \cdot [\bar{x}_{N-1}]_i\langle N-1 \rangle + 1/2 \cdot \sum_{j=0}^{k}[v_{N-1,i}]_j\xi_i(jT_s + nT_u) \\
[\bar{x}_n]_i\langle N \rangle &= [\bar{x}_n]_i\langle N-1 \rangle, \qquad \forall n < N-1.
\end{aligned}
\tag{10}
$$

The state estimates are given by

$$
[\hat{x}_n]_i\langle N \rangle = \min\Big\{\xi_i(nT_u) - l_i, \max\big\{\xi_i(nT_u) - s_i, [\bar{x}_n]_i\langle N \rangle\big\}\Big\}, \qquad \forall n.
\tag{11}
$$

The operation in (11) maintains the bounds on the noise, and it makes sure that the filter saturates if it tries to exceed the bounds of the noise. This filtering is well-defined because of the piecewise continuity of the control input $u(t)$, and it is consistent in a pointwise sense, as the following theorem shows.

Theorem 2 (Ruppert and Wand, 1994)**.** *If T_u is fixed, r is the polynomial order of the filter, and $k \to \infty : kT_s = T_u$; then, the filter defined in (10)-(11) is consistent: $\|x_n - \hat{x}_n\langle N \rangle\| = O_p(k^{-(r+1)/(2r+3)})$.*

Proof. Strictly speaking, the result in [12] applies to the filter defined in (9)-(10). Consistency with respect to (11) is established by noting that the bounds on the noise imply that $\|x_n - \hat{x}_n\langle N \rangle\| \le \|x_n - \bar{x}_n\langle N \rangle\|$. $\qquad\square$

Remark. Because $k = T_u/T_s$, this theorem intuitively says that the filter performs well as long as T_s is much smaller than T_u.

We also have the following lemma which discusses the finite-sample properties of (11). The intuition is that if the measurement noise is bounded and all states are measured, then the filter preserves the property that the state estimates remain within a bounded distance of the true states. Note that the Minkowski sum [14] of two sets $\mathcal{U}, \mathcal{V}$ is defined as $\mathcal{U} \oplus \mathcal{V} = \{u + v : u \in \mathcal{U}; v \in \mathcal{V}\}$.

Lemma 2. *Under the assumptions delineated above, we have that $\hat{x}[n] \in x[n] \oplus \mathcal{E}$, where $\mathcal{E} = \{\epsilon : l_j \le [\epsilon]_j \le s_j\} \oplus (-\{\epsilon : l_j \le [\epsilon]_j \le s_j\})$.*

Proof. Note that (11) enforces that $\xi_j - s_j \le \hat{x}_j \le \xi_j - l_j$, which can be rewritten as $\hat{x} \in \xi \oplus (-\{\epsilon : l_j \le [\epsilon]_j \le s_j\})$. The bounds on the noise $x_i + l_i \le \xi_i \le x_i + s_i$ are equivalent to having $\xi \in x \oplus \{\epsilon : l_j \le [\epsilon]_j \le s_j\}$. The result follows from properties of $\oplus, \ominus$ [14]. $\qquad\square$

3.2 Filter Implementation

Because the filter is simply a weighted sum of measurements (10), the largest difficulty with implementation is in computing the filter coefficients (9). The first step in doing this is to

choose the order of the filter. Empirical results show that linear ($r = 1$) or quadratic ($r = 2$) LPR typically gives good results. For clarity of presentation, we focus here on the case of $r = 1$.

Having chosen the order of the filter, the next step is to compute the bandwidth parameters $h_{\lambda;n,i}, h_{\rho;n,i}$. To make the notation compact, let ? be a blank spot that is either replaced with ? $= \rho$ or ? $= \lambda$. Using results from [5], it can be shown that the optimal bandwidths for $r = 1$ are approximately given by

$$
h_{?;n,i} = \left(\frac{a\sigma^2 T_u}{2\ddot{x}_i(nT_u^?)k} \right)^{1/5}
$$
$$
a = \left. \int_{\mathbb{R}} \kappa^2(\nu)\, d\nu \middle/ \left(\int_{\mathbb{R}} \nu^2 \kappa(\nu)\, d\nu \right)^2 \right.,
\tag{12}
$$

and the second time-derivative $\ddot{x}_i(nT_u^?)$ is the left-sided derivative if ? $= \lambda$ (or right-sided derivative if ? $= \rho$). Unsimplified expressions for the cases $r > 1$ can be found in [5]. We can approximate the values of these second time-derivatives by using (1). More specifically, the estimated values are given by $\hat{\ddot{x}}_i(nT_u^\rho) = [A_c^2 \xi(nT_u) + A_c B_c u_n]_i$ and $\hat{\ddot{x}}_i(nT_u^\lambda) = [A_c^2 \xi(nT_u) + A_c B_c u_{n-1}]_i$.

Because it is time consuming to compute the filter coefficients (9), we suggest an implementation in which they are precomputed. Define a set $\mathcal{H} = \{h_1, \ldots, h_{max}\}$, and compute the filter coefficients for each value in $\mathcal{H}$. Then, when we would like to filter, we estimate the time derivatives $\hat{\ddot{x}}_i(nT_u^\rho)$ and $\hat{\ddot{x}}_i(nT_u^\lambda)$, and use these to compute $h_{?;n,i}$. The closest value in $\mathcal{H}$ is selected, and the corresponding set of precomputed filter coefficients are used to do the filtering as defined in (10)-(11).

4 Epi-convergence Proofs

We provide proofs of the theorems regarding convergence of the control law of LBMPC to an MPC that knows the unmodeled dynamics, for both the case where the oracle is parametric and the case where the oracle is the nonparametric oracle that we defined and call the $L2$-regularized Nadaraya-Watson (L2NW) estimator. The key for these results is that the system trajectory must have a property called sufficient excitation (SE), which intuitively means that all modes of the system are perturbed so that they can be identified. The theorem on convergence is trivial in parametric case, because it results from combining two existing theorems that are valid under SE. The proof for the L2NW case is more involved, since it requires showing epi-convergence of the L2NW estimator under the notion of SE.

4.1 Parametric Oracle

Proof of Theorem 4 in [3]. The proof simply requires application of existing theorems. If $\hat{\theta}_m$ converges in probability to θ_0, then the result is true by Proposition 2.1 of [17]. The required convergence in probability occurs under SE [7, 6, 9], and so the result trivially follows. $\square$

Remark. The situation in which the states are measured with noise requires the use of the continuous mapping theorem [15] taken in conjunction with Theorem 2. For the case where

the parameters enter linearly, the hypothesis of the continuous mapping theorem is satisfied because $\hat{\lambda}_m = (X'X)^{-1}X'Y$ is continuous with respect to X given SE [7]. For the nonlinear case, we need to explicitly assume that $\hat{\lambda}_m$ is continuous with respect to X in order to be able to apply the continuous mapping theorem.

4.2 Nonparametric Oracle

Showing that the L2NW estimator leads to convergence of the control law of LBMPC under an assumption of SE requires additional work. For ease of reference, we give one expression of the L2NW estimator defined in [3]. Let $X_i = [x_i' \ u_i']'$, $Y_i = x_{i+1} - (Ax_i + Bu_i)$, and $\Xi_i = \|\xi - x_i\|^2/h^2$, where $X_i \in \mathbb{R}^{p+m}$ and $Y_i \in \mathbb{R}^p$ are data and $\xi = [x' \ u']'$ are free variables. We define any function $\kappa : \mathbb{R} \to \mathbb{R}_+$ to be a kernel function if it has (a) finite support $\kappa(\nu) = 0$ for $|\nu| \geq 1$, (b) even symmetry $\kappa(\nu) = \kappa(-\nu)$, (c) positive values $\kappa(\nu) > 0$ for $|\nu| < 1$, (d) differentiability (i.e., the derivative $d\kappa$ exists), and (e) nonincreasing values of $\kappa(\nu)$ over $\nu \geq 0$. The $L2$-regularized NW (L2NW) estimator is defined as

$$\mathcal{O}_m(x, u; X_i, Y_i) = \frac{\sum_i Y_i \kappa(\Xi_i)}{\lambda + \sum_i \kappa(\Xi_i)}, \tag{13}$$

where $\lambda \in \mathbb{R}_+$. If $\lambda = 0$, then (13) is simply the NW estimator. The λ term acts to regularize the problem and ensures differentiability.

We begin by proving a uniform version of a theorem that is called either the continuous mapping theorem [15] or Slutsky's theorem [4], depending on the author.

Lemma 3. *Given random variables $V_k, V \in \mathcal{V}$, for all $k \in \mathbb{Z}_+$, such that $\|V_k - V\| = O_p(r_k)$; if $L(x,v) : \mathcal{X} \times \mathcal{V} \to \mathbb{R}$ is a continuous function and $\mathcal{X}, \mathcal{V}$ are compact sets, then $\sup_{x \in \mathcal{X}} |L(x, V_k) - L(x, V)| = O_p(r_k)$.*

Proof. The Heine-Cantor theorem (Theorem 4.19 in [11]) gives uniform continuity of $L(x,v)$ on $\mathcal{X} \times \mathcal{V}$, and this implies that for all x, $\|V_k - V\| > \delta > 0$ whenever $|L(x, V_k) - L(x, V)| > \epsilon > 0$. Proceeding analogously to [15], we have

$$\mathbb{P}(\sup_x |L(x, V_k) - L(x, V)| > \epsilon) = \mathbb{P}(\exists x : |L(x, V_k) - L(x, V)| > \epsilon) \leq \mathbb{P}(\|V_k - V\| > \delta). \tag{14}$$

The result is immediate. $\qquad\square$

We can now show the first convergence result for the L2NW estimator. Let $\hat{X}_i, \hat{Y}_i$ be defined in the same way as X_i, Y_i, with the change that $\hat{X}_i, \hat{Y}_i$ are defined using state estimates $\hat{x}$ instead of noiseless measurements of the state x. The intuition of why this result is true is that though noise in $\hat{Y}_i$ and $\hat{X}_i$ is correlated, our filtering defined in Section 3 makes this correlation asymptotically insignificant. This result can be interpreted in an instrumental variables context [1, 8].

Corollary 1. *If $\|\hat{X}_i - X_i\| = O_p(r_k)$, then $\sup_{x \in \mathcal{X}, u \in \mathcal{U}} \|\mathcal{O}_m(x, u; \hat{X}_i, \hat{Y}_i) - \mathcal{O}_m(x, u; X_i, Y_i)\| = O_p(r_k)$.*

7

Proof. Define a random variable $V_k = \begin{bmatrix} \hat{X}'_1 & \ldots & \hat{X}'_N & \hat{Y}'_1 & \ldots & \hat{Y}'_N \end{bmatrix}'$, and let V be the corresponding limiting vector. The definition of Y_i and the corollary's assumption imply that $\|\hat{Y}_i - Y_i\| = O_p(r_k)$, and so $\|V_k - V\| = O_p(Nr_k)$.

Now consider the functions $\eta = \sum_i \hat{Y}_i \kappa(\Xi_i)/N$, $\delta = \sum_i \kappa(\Xi_i)/N$, and $\rho = \eta/(\lambda/N + \delta)$. Applying Lemma 3 gives, $\sup_{x \in \mathcal{X}, u \in \mathcal{U}} \|\eta(\xi; V_k) - \eta(\xi; V)\| = O_p(r_k)$ and $\sup_{x \in \mathcal{X}, u \in \mathcal{U}} \|\delta(\xi; V_k) - \delta(\xi; V)\| = O_p(r_k)$. Another application of Lemma 3 gives $\sup_{x \in \mathcal{X}, u \in \mathcal{U}} \|\rho(\xi; V_k) - \rho(\xi; V)\| = O_p(r_k)$. The result follows by noting that $\mathcal{O}_m(x, u; \hat{X}_i, \hat{Y}_i) = \rho(\xi; V_k)$ and $\mathcal{O}_m(x, u; X_i, Y_i) = \rho(\xi; V)$. $\qquad\square$

Remark. The variance of the NW estimator in its typical setup is known to uniformly converge at a rate no faster than $n^{-4/(p+4)}$ [12]. Our result gives a nonstandard rate of convergence r_k, because we have a time-series setup with presmoothing to account for the errors in measurements.

Convergence of an estimator is often studied by decomposing the estimation error into a bias and variance term. For proving convergence of our regularized NW estimator, we have to be careful in defining the probabilistic framework before we can decompose the error into two terms. The X_i values are not independent variables drawn from some probability distribution. They are exactly the states of a deterministic system, as it evolves in time. In fact, if the control inputs u_n are (deterministically or statistically) known for each point in time, then X_i and X_j are dependent for all values where $i \geq j$.

For a nonlinear system, SE is usually defined using ergodicity or mixing, but this is hard to verify in general. Instead, we define SE as a finite sample cover (FSC) of $\mathcal{X}$. Let $\mathcal{B}_h(x) = \{y : \|x - y\| \leq h\}$ be a ball centered at x with radius h, then a FSC of $\mathcal{X}$ is a set $\mathcal{S}_h = \bigcup_i \mathcal{B}_{h/2}(X_i)$ that satisfies $\mathcal{X} \subseteq \mathcal{S}_h$. The intuition is that $\{X_i\}$ sample $\mathcal{X}$ with average, inter-sample distance less than $h/2$. Assuming SE in the form of a FSC with asymptotically decreasing radius h, we can show that the control law of LBMPC that uses L2NW converges to that of an MPC that knows the true dynamics.

Recall that $g(x, u)$ is the modeling error of the approximate linear system defined in [3]. We have the following result, which shows that the L2NW estimator with noiseless measurements and SE can approximate the unmodeled dynamics arbitrarily well.

Lemma 4. *If $g(x, u)$ is Lipschitz with constant L and $\mathcal{S}_h$ is a finite sample cover of $\mathcal{Z} \subseteq \mathcal{X} \times \mathcal{U}$, then $\sup_{(x,u) \in \mathcal{Z}} \|g(x, u) - \mathcal{O}_m(x, u; X_i, Y_i)\| \leq \mu M_g + (1 - \mu)Lh$, where $\mu = \lambda/(\lambda + \kappa(1/2))$ and $M_g = \max \|x\| : x \in \mathcal{X}$.*

Proof. Define $I = \{i : \Xi_i \leq 1\}$, and note that $\kappa(\Xi_j) = 0$ for $j \notin I$. An alternative characteristic of the L2NW estimator is as the positively weighted average: $\mathcal{O}_m(x, u; X_i, Y_i) = w_0 \cdot 0 + \sum_{i \in I} w_i \cdot Y_i$, where $w_0, w_i > 0$, $w_i = \kappa(\Xi_i)/(\lambda + \sum_j \kappa(\Xi_j))$, and $w_0 + \sum_i w_i = 1$. The finite sample cover property of $\mathcal{S}_h$ implies that $\sum_j \kappa(\Xi_j) \geq \kappa(1/2)$. Noting $w_0 < \lambda/(\lambda + \kappa(1/2))$ and $Y_i = g(X_i)$, the result follows from the triangle inequality. $\qquad\square$

Remark. The result shows that the regularized NW estimator in our setup has bias $O(\lambda + h)$, where $\lambda = O(h)$. This matches the bias of the NW estimator $O(h)$ in a standard setup at both interior and boundary points [10, 12].

Theorem 3. *Assuming $\mathcal{S}_{h_n}$ is a finite sample cover of $\mathcal{Z} \subseteq \mathcal{X} \times \mathcal{U}$, for some sequence $h_n \to 0$; $\lambda = O(h_n)$; and k is a sequence such that $T_u/k \to 0$ (see Theorem 2); then the regularized NW estimator is uniformly consistent on $\mathcal{Z}$ and converges at rate*

$$\sup_{(x,u)\in\mathcal{Z}} \|g(x,u) - \mathcal{O}_m(x,u;\hat{X}_i,\hat{Y}_i)\| = O(\lambda + h_n) + O_p(k^{-(r+1)/(2r+3)}). \qquad (15)$$

Proof. Using the triangle inequality, the left-hand side of (15) is bounded by

$$\sup_{(x,u)\in\mathcal{Z}} \|\mathcal{O}_m(x,u;\hat{X}_i,\hat{Y}_i) - \mathcal{O}_m(x,u;X_i,Y_i)\| + \sup_{(x,u)\in\mathcal{Z}} \|g(x,u) - \mathcal{O}_m(x,u;X_i,Y_i)\| \qquad (16)$$

This first term is controlled by Corollary 1 and Theorem 2, and the second is governed by Lemma 4. $\qquad\square$

Remark. Ideally, we would like $\mathcal{Z} = \mathcal{X} \times \mathcal{U}$, but this does not always happen. It requires that the trajectory of the system sufficiently explores the space in a manner formalized by the definition of finite sample cover. A set $\mathcal{Z}$ which meets the assumptions of Theorem 3 always exists, and this can be shown by construction: Given any $n > 0$, let $\mathcal{Z} = \cup_{i=1}^n X_i$. A better set $\mathcal{Z}$ is defined as the limit of a convergent subsequence of X_i, and its advantage is that the X_i visit a neighborhood of the limit infinitely often. Such a limit is guaranteed to exist by the Bolzano-Weierstrass theorem. These two constructions mean that there is always some set on which the nonlinear dynamics $g(x,u)$ can be learned, and this set corresponds to points which the trajectory visits.

We need the following theorem in order to show epi-convergence of $\tilde{\psi}_m$ as defined in [3], for the LBMPC problem that uses the L2NW estimator (13) as the oracle.

Theorem 4. *Let $\mathcal{X}_v, \mathcal{X}_w, \mathcal{R} \subset \mathbb{R}^n$ be closed and compact sets, and assume that we have a sequence of functions $V_k(x) : \mathcal{X}_v \to \mathcal{X}_w$ and $W_k(x) : \mathcal{X}_w \to \mathcal{R}$ which converge in probability to $V(x), W(x)$ as $\sup_{x\in\mathcal{X}_v} \|V_k(x) - V(x)\| = O_p(r_k)$ and $\sup_{x\in\mathcal{X}_w} \|W_k(x) - W(x)\| = O_p(s_k)$. If W is Lipschitz continuous with constant L_w, then $\sup_{x\in\mathcal{X}_v} \|W_k(V_k(x)) - W(V(x))\| = O_p(c_k)$, where $c_k = \max\{r_k, s_k\}$.*

Proof. Applying the triangle inequality gives

$$\mathbb{P}\Big(\sup_{x\in\mathcal{X}_v} |W_k(V_k(x)) - W(V(x))|/c_k \geq \epsilon\Big) \leq \mathbb{P}\Big(\sup_{x\in\mathcal{X}_v} |W_k(V_k(x)) - W(V_k(x))|/c_k \geq \epsilon\Big) +$$
$$\mathbb{P}\Big(\sup_{x\in\mathcal{X}_v} |W(V_k(x)) - W(V(x))|/c_k \geq \epsilon\Big). \quad (17)$$

The first term on the right in (17) can be bounded as

$$\mathbb{P}\Big(\sup_{x\in\mathcal{X}_v} |W_k(V_k(x)) - W(V_k(x))|/c_k \geq \epsilon\Big) \leq \mathbb{P}\Big(\sup_{x\in\mathcal{X}_w} |W_k(x) - W(x)|/c_k \geq \epsilon\Big), \quad (18)$$

and so the limit of (18) by assumption is $\lim \mathbb{P}(\sup_{x\in\mathcal{X}_v} |W_k(V_k(x)) - W(V_k(x))|/c_k \geq \epsilon) = 0$. The second term on the right in (17) is bounded using the Lipschitz constant as

$$\mathbb{P}\Big(\sup_{x\in\mathcal{X}_v} |W(V_k(x)) - W(V(x))|/c_k \geq \epsilon\Big) \leq \mathbb{P}\Big(\sup_{x\in\mathcal{X}_v} L_w|V_k(x) - V(x)|/c_k \geq \epsilon\Big), \quad (19)$$

and taking its limit gives by assumption that $\lim \mathbb{P}(\sup_{x\in\mathcal{X}_v} |W(V_k(x)) - W(V(x))|/c_k \geq \epsilon) = 0$. The result follows by taking the limit of (17) and observing that the limit is equal to zero. $\qquad\square$

Remark. The theorem shows that convergence in probability is preserved under composition, but the one subtlety in the result and subsequent proof is the issue of domains of convergence. We are composing two functions $W_k(V_k(x))$, and convergence occurs as long as the range of the function on the inside $V_k(\cdot)$ lies within the domain of convergence of the function on the outside $W_k(\cdot)$. By Theorem 5 of [3], the L2NW estimator has the required range.

With the previous theorem, we can now show that $\tilde{\psi}_m$ epi-converges. Consequently, the control law of LBMPC with L2NW as oracle converges to that of MPC that knows the unmodeled dynamics, when there is SE as defined by the appropriate FSC.

Proof of Theorem 6 in [3]. Note that equality constraint in LBMPC

$$\tilde{x}_{n+1} = A\tilde{x}_n + B\check{u}_n + \mathcal{O}_m(\tilde{x}_n, \check{u}_n) \tag{20}$$

recursively defines $\tilde{x}_{m+i}$, for $i = \{0, \dots, N\}$, as functions of only $\tilde{x}_m$ and c. For example, the equation for $\tilde{x}_{m+1}$ is given by

$$\begin{aligned}
\tilde{x}_{m+2}(\tilde{x}_m, c; \mathcal{O}_m) = {} & A^2\tilde{x}_m + AB(K\tilde{x}_m + c_m) + A\mathcal{O}_m(\tilde{x}_m, K\tilde{x}_m + c_m) \\
& + B(K(A\tilde{x}_m + B(K\tilde{x}_m + c_m) + \mathcal{O}_m(\tilde{x}_m, K\tilde{x}_m + c_m)) + c_{m+1}) \\
& + \mathcal{O}_m(A\tilde{x}_m + B(K\tilde{x}_m + c_m) + \mathcal{O}_m(\tilde{x}_m, K\tilde{x}_m + c_m), K(A\tilde{x}_m \\
& \qquad + B(K\tilde{x}_m + c_m) + \mathcal{O}_m(\tilde{x}_m, K\tilde{x}_m + c_m)) + c_{m+1}). \tag{21}
\end{aligned}$$

By Theorem 4 we have that $\sup_{x_m:\phi(x_m)\neq 0} \|\tilde{x}_{m+i}(x_m, \mathcal{O}_m) - \tilde{x}_{m+i}(x_m, g)\| = O_p(r_m)$, where r_m is the convergence rate from Theorem 3. Since ψ is continuous, we can compose it with $\tilde{x}[m+i]$ using Theorem 4. This gives that $\sup_{x_m:\phi(x_m)\neq 0} \|\tilde{\psi}_m - \tilde{\psi}_0\| = O_p(r_m)$.

The last step requires showing that this condition is equivalent to lower semicontinuity in probability. Because $\tilde{\psi}_0$ is continuous, given $\epsilon > 0$ and a point x_0, c, there exists a neighborhood $U\{x_0, c\}$ such that

$$|\tilde{\psi}_0(\zeta) - \tilde{\psi}_0(x_0, c)| < \epsilon/2, \tag{22}$$

for all $\zeta \in U\{x_0, c\}$. Now consider the expression

$$\alpha = \mathbb{P}\left(\inf_{\zeta \in U\{x_0,c\}} \tilde{\psi}_m < \tilde{\psi}_0(x_0, c) - \epsilon \right) \leq \mathbb{P}\left(\sup_{\zeta \in U\{x_0,c\}} |\tilde{\psi}_m - \tilde{\psi}_0(x_0, c)| > \epsilon \right). \tag{23}$$

Using (22), we can further bound the expression above by

$$\alpha \leq \mathbb{P}\left(\sup_{\zeta \in U\{x_0,c[\cdot]\}} |\tilde{\psi}_m - \tilde{\psi}_0| > \epsilon/2 \right). \tag{24}$$

Taking the limit, we have that $\lim \alpha = 0$, and so the result follows by applying Proposition 5.1 of [16]. $\qquad\square$

5 Acknowledgments

References

[1] K. Åström and P. Eykhoff. System identification–a survey. *Automatica*, 7(2):123–162, 1971.

[2] A. Aswani, P. Bouffard, and C. Tomlin. Extensions of learning-based model predictive control for real-time application to a quadrotor helicopter. Submitted.

[3] A. Aswani, H. Gonzalez, S. Sastry, and C. Tomlin. Provably safe and robust learning-based model predictive control. Submitted.

[4] Peter Bickel and Kjell Doksum. *Mathematical Statistics: Volume 1*. Prentice-Hall, 2 edition, 2007. Updated printing.

[5] J. Fan and I. Gijbels. Data-driven bandwidth selection in local polynomial fitting: Variable bandwdith and spatial adaptation. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(2):371–394, 1995.

[6] R. Jennrich. Asymptotic properties of non-linear least squares estimators. *Annals of Mathematical Statistics*, 40:633–643, 1969.

[7] T. Lai, H. Robbins, and C. Wei. Strong consistency of least squares estimates in multiple regression ii. *Journal of Multivariate Analysis*, 9:343–361, 1979.

[8] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, 1987.

[9] E. Malinvaud. The consistency of nonlinear regressions. *Annas of Mathematical Statistics*, 41(3):956–969, 1970.

[10] H. Müller. Weighted local regression and kernel methods for nonparametric curve fitting. *Journal of the American Statistical Association*, 82:231–238, 1987.

[11] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, 2 edition, 1964.

[12] D. Ruppert and M. Wand. Multivariate locally weighted least squares regression. *Annals of Statistics*, 22(3):1346–1370, 1994.

[13] A. Savitzky and M. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8):1627–1639, 1964.

[14] R. Schneider. *Convex bodies: the Brunn-Minkowski theory*. Cambridge University Press, 1993.

[15] A. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 2000.

[16] S. Vogel and P. Lachout. On continuous convergence and epi-convergence of random functions. part i: Theory and relations. *Kybernetika*, 39(1):75–98, 2003.

[17] S. Vogel and P. Lachout. On continuous convergence and epi-convergence of random functions. part ii: Sufcient conditions and applications. *Kybernetika*, 39(1):99–118, 2003.