

Vibration Signal-Assisted Endpoint Detection for Long-Stretch, Ultraprecision Polishing Processes

Shilan Jin

Department of Industrial & Systems Engineering,
Texas A&M University,
College Station, TX 77843
e-mail: jin0541@tamu.edu

Satish Bukkapatnam

Department of Industrial & Systems Engineering,
Texas A&M University,
College Station, TX 77843
e-mail: satish@tamu.edu

Sean Michael Hayes

Lawrence Livermore National Laboratory,
Livermore, CA 94550
e-mail: hayes53@llnl.gov

Yu Ding¹

Department of Industrial & Systems Engineering,
Texas A&M University,
College Station, TX 77843
e-mail: yuding@tamu.edu

The research reported in this article is concerned with the question of detecting and subsequently determining the endpoint in a long-stretch, ultraprecision surface polishing process. While polishing endpoint detection has attracted much attention for several decades in the chemical-mechanical planarization of semiconductor wafer polishing processes, the uniqueness of the surface polishing process under our investigation calls for novel solutions. To tackle the research challenges, we develop both an offline model and an online detection method. The offline model is a functional regression that relates the vibration signals to the surface roughness, whereas the online procedure is a change-point detection method that detects the energy turning points in the vibration signals. Our study reveals a number of insights. The offline functional regression model shows clearly that the polishing process progresses in three states, including a saturation phase, over which the polishing action could be substantially shortened. The online detection method signals in real-time when to break a polishing cycle and to institute a follow-up inspection, rather than letting the machine engage in an overpolishing cycle for too long. When implemented properly, both sets of insights and the corresponding methods could lead to substantial savings in polishing time and energy and significantly improve the throughput of such polishing processes without inadvertently affecting the quality of the final polish. [DOI: 10.1115/1.4056809]

Keywords: change-point detection, energy turning point, functional features, functional linear regression, saturated polishing effect, surface roughness prediction, inspection and quality control, process engineering, production systems optimization

1 Introduction

In early Aug. 2021, Lawrence Livermore National Laboratory (LLNL) announced sensational news that their National Ignition Facility (NIF) succeeded in the experiment of using a laser shot to spark a fusion explosion from a peppercorn-size fuel capsule [1] that resulted in an igniting implosion with over a megaJoule fusion yield. That was the first controlled fusion experiment, among the worldwide attempts for more than five decades, to achieve a sufficiently high fusion heating power that overcame all the cooling processes and created a positive thermodynamic feedback loop with rapidly increasing temperature [2]. More than one year later, LLNL conducted another breakthrough experiment on Dec. 5, 2022 [3]. This time, the fusion energy output is 1.5 times greater than the input. This is the first time ever that scientists were able to produce more energy than the immediate laser energy input in a controlled fusion process. The success of these experiments took a major step forward in demonstrating a route of using a small amount of energy to engender limitless fusion energy. Not surprisingly, this news quickly appeared on nearly all major media outlets in the US.

One centerpiece in the NIF experiments is the fuel capsule, which is a silicon ball (diameter of ~ 2 mm) with a multilayer coating, and the outermost is diamond coating; for its detailed design, please refer to Ref. [4]. Stringent surface roughness is required for these diamond shells to ensure efficiency and effectiveness in the ignition

experiments. Even minor irregularities reduce the efficiency and suppress ignition [5]. The specific requirement is to have the shell surface roughness down to the nanoscale, meaning that if measured using the surface roughness metrics [6], like R_a (for one-dimensional profiles) or S_a (for two-dimensional areas), their values should be in tens of nanometers at most.

To achieve the nanoscale roughness of the shell surface, the researchers in LLNL carry out the polishing processes in multiple long-stretch polishing stages. For each polishing stage, they use a specialized, proprietary machine to polish the shells for 24 h, then pause the polishing and remove the shells to take surface roughness measurements (using a Confocal microscope, model Keyence VK-X1100). After the measurement-taking, some cleaning up (like washing) of the shells and machine is undertaken. Then the polishing action is resumed for another 24 h. The time-series vibration data are collected during the entirety of the polishing process. This aforementioned process is repeated multiple times (seven and eight times in the LLNL polishing experiments). The whole polishing process could take more than a week to complete. Figure 1 illustrates the polishing and data acquisition process. For how to calculate S_a based on the measured surface roughness data, one may refer to Ref. [7].

Such long-stretch polishing processes are rooted in two primary reasons: (a) the shell surfaces are very hard and also brittle (diamond coating), and as a result, the material removal rate cannot be too high; and (b) such surface polishing is rather different from many existing polishing processes in the traditional manufacturing, including semiconductor manufacturing, and there is very little knowledge available for advising how long one should polish and when to stop. LLNL experimenters chose to err on the side of overpolishing over underpolishing. While understandable for the current setup, a research question confronting us is—can

¹Corresponding author.

Manuscript received September 30, 2022; final manuscript received January 14, 2023; published online February 20, 2023. Assoc. Editor: Qiang Huang.

This work is in part a work of the U.S. Government. ASME disclaims all interest in the U.S. Government's contributions.

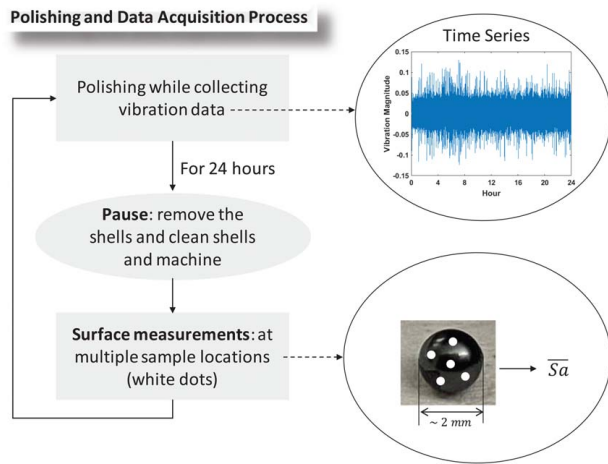


Fig. 1 The illustration of the polishing and data acquisition process. For the data acquisition, both the sensor data in time domain and the S_a values of the surface sample locations are acquired.

LLNL experimenters shorten the polishing process? For example, is the 24-h polishing stretch per stage necessary? Are there any opportunities for the polishing process to stop during the long stretch between two pauses? When is a good point for the polishing process to stop?

For polishing endpoint determination, there are two major schools of thought. One is product-oriented, and the other is process-oriented [7,8]. The product-oriented approach is to take the surface measurements of the product under polishing and compute its surface roughness metric like S_a or build surface roughness characterization. Such procedures are done at the pauses after each polishing stage, as described above. Also, please refer to Refs. [7,8] and the literature cited therein for the latest development on this front. The advantage of the product-oriented methods is their accurate detection of the endpoint as they determine those directly based on the surface condition. However, as the surface roughness evaluations take place in an offline fashion and through intermittent measurements, the endpoint decision can only be made at the end of each stage. As such, they are incapable of deciding how long each stage should last and when to stop during a stage.

The limitation of product-oriented approaches motivated researchers to develop process-oriented approaches, also known as sensor data-assisted approaches. The continuous sensing signal enables the online monitoring of the polishing process and provides the potential to allow detecting endpoints that may occur at any time during the polishing process. Interestingly, there are more process-oriented studies than product-oriented studies in the literature. We delay a detailed literature review on process-oriented studies to Sec. 2.

This second school of thought is most relevant to the research questions posed earlier. In the diamond shell polishing process, in addition to the measurements of the product surface taken intermittently between stages, an accelerometer is installed on the polishing machine, collecting the vibration signals continuously at a 10 kHz frequency. Our specific research aim then becomes whether it is possible and, if so, how we can make use of the vibration signals for detecting/suggesting the polishing endpoint. The “endpoint” includes not only the final stop point but also other machine stop points, such as changing polishing pads or adjusting spindle rotation speeds or downforce.

After carefully reviewing the relevant literature (details in Sec. 2), we conclude that the existing ideas and methods in the process-oriented approaches do not address the research questions posed earlier for the long-stretch ultraprecision polishing process. The reasons can be understood as follows. The process-oriented approaches can be further partitioned into two main categories:

one is through mechanistic modeling, which relies on some process conditions to hint when to stop, and the second is to build a data-driven model to let the in-process signals foretell what the product surface roughness is going to be, should the polishing stop right now or shortly after. For the first approach category, while the general concept is good, there is no existing knowledge directly applicable to the specific polishing process we have at hand. For the second category of methods to work, one needs numerous training data to build a data-driven model. A valid training data set requires the polishing stage to be short (a few minutes) and have numerous stages. In contrast, that is the luxury we do not have—we only have 7–8 stages, each lasting for a whole day.

To tackle the research challenges for this ultraprecision polishing process with long-stretch stages, we develop an offline model and an online detection method. The offline model is a functional regression that relates the vibration signals to the surface roughness, whereas the online procedure is a change-point detection method that detects the energy turning points in the vibration signals. Our study reveals a number of insights that advance the knowledge frontier for this important polishing process. We believe that our research makes the following contribution to the literature:

- Our study is one of the first quantitative studies looking into when to stop a diamond-coating shell polishing process, and the modeling frameworks and the insights into the long-term polishing effect are generalizable to other long-stretch, ultraprecision polishing processes.
- Our offline functional regression model shows clearly that the polishing process progresses in three phases, including a saturation phase, over which the polishing action could be substantially shortened.
- The online detection method signals in real-time when to break a polishing cycle and institute a follow-up inspection rather than letting the machine engage in an overpolishing cycle for too long.

Making proper use of the knowledge and methods produced by our research could lead to substantial savings in polishing time and energy and greatly improve the throughput of such polishing processes without inadvertently affecting the quality of the final polish.

The rest of the paper is organized as follows. Section 2 discusses the relation of our work with the existing studies about endpoint detection. Section 3 describes the feature extraction process from vibration signals, which provides the input into offline modeling and online detection. Sections 4 and 5 present the offline model and the online detection method, respectively. Section 6 concludes the study and discusses future work.

2 Relation to the Literature

As aforementioned in Sec. 1, there are two schools of thought on polishing endpoint detection: product- and process-oriented methods. For the product-oriented approaches, please refer to the latest literature review in Ref. [8]. This section aims to provide a literature analysis of the process-oriented methods.

The process-oriented polishing endpoint detection has been studied extensively in the chemical-mechanical planarization (CMP) of semiconductor wafer polishing processes [9]. The polishing object in our research is not a semiconductor wafer but the purposes, i.e., polishing endpoint detection, in our research and semiconductor CMP bear similarity. For this reason, we review the CMP endpoint detection methods that are relevant to the solution to our problem.

There are two categories of CMP polishing processes due to their different polishing objectives: the back-end-of-line (BEOL) and front-end-of-line (FEOL) polishing. The endpoint detection methods are so divided into two categories as shown in Fig. 2 and Secs. 2.1 and 2.2 provide further analysis and discussion,

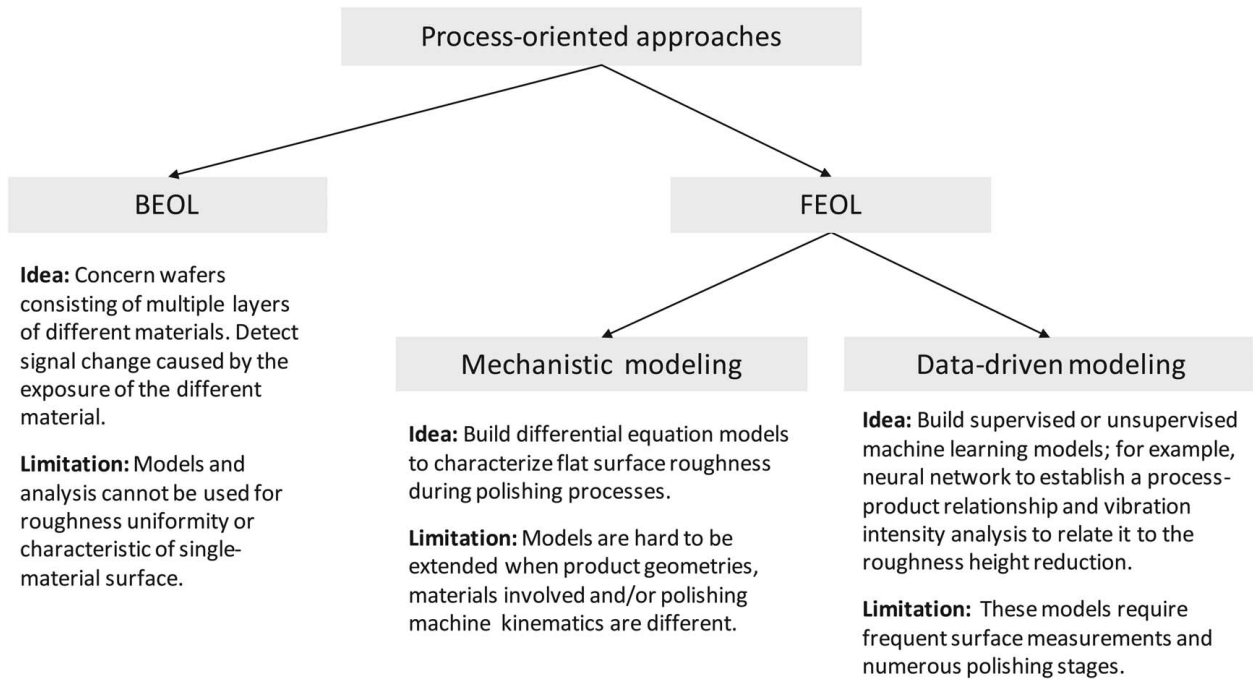


Fig. 2 A summary of the process-oriented approaches for polishing endpoint detection

respectively. Under the branch of FEOL, the methods are further categorized into mechanistic modeling and data-driven modeling methods.

2.1 BEOL Methods. Numerous methods [10–17] rely on different sensor data to determine the polishing endpoint in the so-called BEOL polishing. The BEOL polishing processes concern wafers consisting of multiple layers of different materials. The endpoint of BEOL is used to prevent overpolishing a layer where the desired material is lost or underpolishing where the desired material is not exposed for the next layer. Chris and Sandhu [10] monitor the transition from one film to another during the planarization of multi-layered stacked film by the varying speed of the reflected acoustic signals from the different film materials. Salugsugan [12] detects the change in the audio signals generated by rubbing contact between a polishing pad and a soft surface versus a hard material after the soft layer is removed. Kojima et al. [17] monitor the trend of vibration signals collected during Cu-CMP processes and detect the polishing endpoint at the exposure of an inter-level dielectric layer.

These methods are not suited to the FEOL polishing in CMP that matches the polishing objective in our research [18]. The wafers in FEOL consist of a single material (e.g., silicon). The endpoint of FEOL is determined based on the finish or uniformity of the wafer or the condition of the CMP tool (e.g., pad wear, slurry pH) to achieve a certain quality or productivity criterion.

2.2 FEOL Methods. The literature on process-oriented methods in FEOL endpoint detection is therefore reviewed in detail. These process-oriented methods can be further categorized into two branches: models based on the mechanistic theory of the polishing processes and models driven by the data that interpret the connection or relationship between the process conditions and product surface roughness.

2.2.1 Mechanistic Modeling Methods. The mechanistic modeling framework models the relationship between the product property and the process conditions based on the physical principles of the polishing processes [19–21]. For instance, Rao et al. [19] modeled the process machine interaction (PMI) in the CMP

processes to relate the vibration signals with several product conditions, such as the polishing pad asperities, bulk-pad structural dynamics, and machine kinetics. They built a set of deterministic differential equations based on the first principles of flat surface polishing processes. Botcha et al. [21] extended a three degrees-of-freedom lumped mass model to establish the relationship between the high frequency signatures of vibration signals and the surface roughness in a cylindrical plunge grinding process.

The main limitation of mechanistic models is that one has to develop a new model when any key parameter is different, such as product geometry, pad asperity level, underlying materials, and polishing machine kinematics. No known mechanistic model in the literature is applicable to our nanoscale ultraprecision shell polishing process.

2.2.2 Data-Driven Modeling Approaches. The other category of process-oriented models, arguably easier to extend, is based on machine learning models (or data science models, data-driven models), either unsupervised learning or supervised learning. Under this branch, relatively fewer works use the idea of following the magnitude change of certain process signals when the surface uniformity desired has not been achieved. That is an unsupervised learning method of the feature input only to find a pattern of the input when a certain output is desired. Hetherington et al. [22] use the vibration signal to monitor a continuous layer dielectric wafer surface from 475 μm to 20 μm within 160 s. They claimed that the attenuation in the vibration intensity is related to roughness reduction. In principle, the polishing endpoint is detected when vibration intensity reaches the minimum level or goes below a certain threshold. In reality, even if the decreasing trend could be established, deciding the threshold for stopping is not trivial. A bigger challenge of using this idea is that Hetherington et al.’s claim is not valid for nanoscale surface polishing. We will present data to support our findings in later sections. While it is intuitive that surface roughness affects the kinetic energy of polishing and hence the use of vibration signals, we found that the vibration signal trending on the nanoscale surfaces is much more complicated. As a result, there is a need for new ideas and detection methods.

Researchers have come to realize the limitations of using the process condition data alone. The more recent research aims to

build a relationship between process conditions and product surface roughness through a supervised learning process. The idea is that with such models, one can anticipate a product's roughness, given the in-process measurements, and therefore decide when to stop the polishing action. For instance, Bukkapatnam et al. [23] attempted to use linear and nonlinear regression models to relate process parameters and vibration signals to the process condition (i.e., material removal rate (MRR)). Kong et al. [24] used nonlinear sequential Bayesian models to predict the process state (i.e., vibration) and then used a nonlinear regression (e.g., neural network) model to relate the process state with the process condition. Similarly, García et al. [25] and Segreto et al. [26] used the feature extracted from vibration signals by wavelet packet decomposition to predict Ra . García and López [27] fit the vibration-surface roughness relationship with a polynomial regression predictive model.

While the machine learning process-product modeling approaches sound reasonable and attractive, their success depends on their model's capability to predict the surface roughness in the next stage of operation. For such models to have a good enough prediction for any time in the future, the polishing process needs to have relatively short intervals between product surface measurements and many enough stages with sufficient variation. Consider the process in Ref. [26]; there are 5580 training stages of different lengths in the polishing process. By contrast, each stage of our long-stretch nanoscale polishing process lasts 24 h, and there are usually just seven or eight training stages altogether. For this type of long-stretch polishing process with so few stages, the resulting machine learning model (such as a neural network) will not be

accurate enough to guide the in-process polishing process and inform its endpoint.

In this paper, our research undertaking is related to two sets of ideas reviewed above, but in both cases, we have to conduct new research. Under an unsupervised setting using the process input alone, our work makes use of the kinetic energy of vibration signals, but our data show that we have to do something different from following a simple monotonic trend in vibration intensity. Under a supervised setting where the process-product data-driven model can be established, we build a statistical model, but considering the limitation in our data, our model serves a different purpose (estimation of surface roughness and process understanding) than making an in-process prediction.

3 Feature Extraction

The inputs to subsequent modeling and analysis are features in frequency and time domains extracted from the vibration signals. The raw vibration signal reflects the vibration magnitude in the time domain, whereas the time-frequency analysis is desired in this work to understand the temporal variation of frequency as the surface polishing progresses.

The wavelet transform [28,29] provides the ability for the time-frequency analysis. We employ the discrete wavelet transform, specifically the wavelet packet decomposition (WPD), for the multi-resolution discrete-time signal decomposition. WPD decomposes the signal into a combination of low-frequency coefficients (i.e., *approximate coefficients*) and high-frequency coefficients (i.e.,

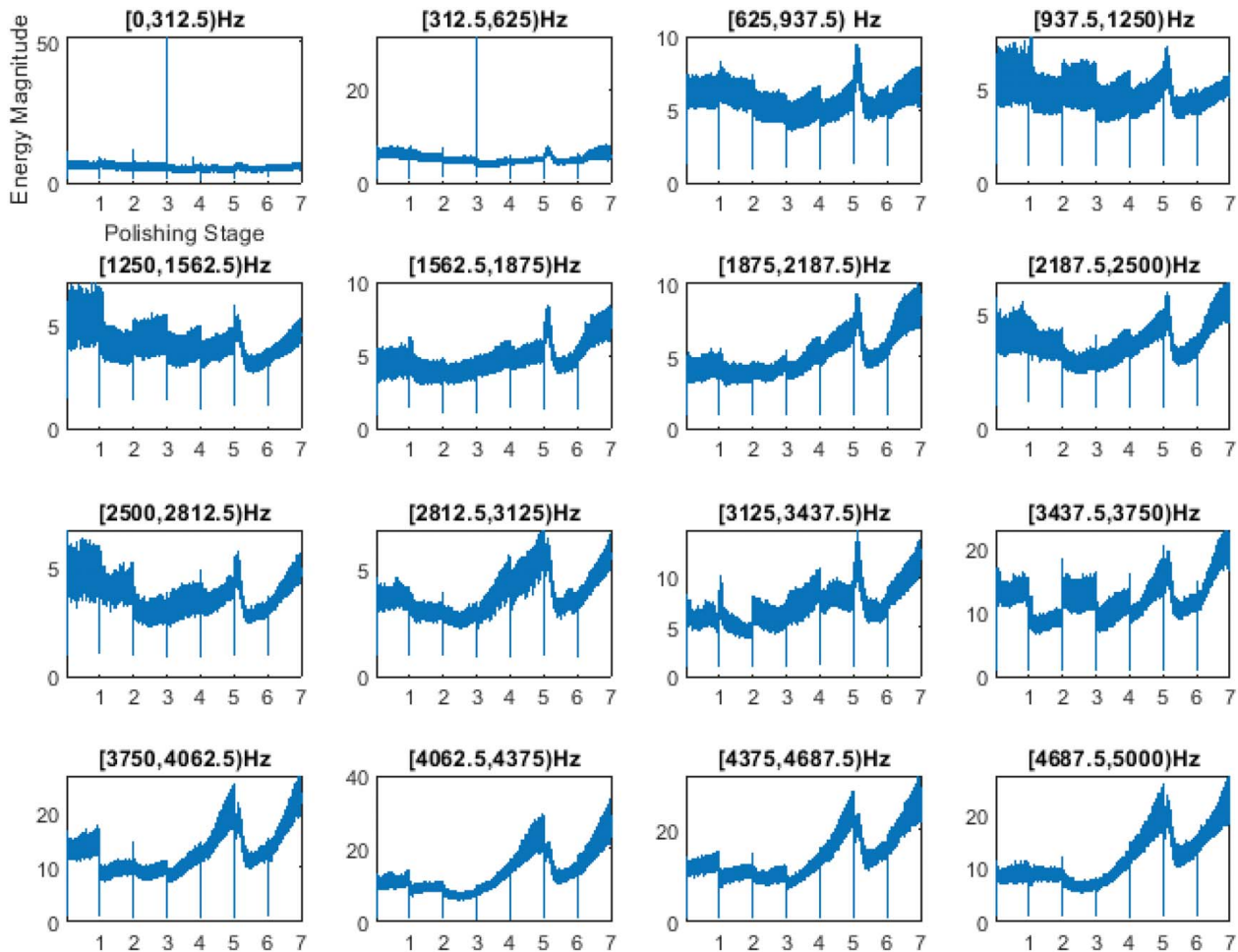


Fig. 3 Vibration signals over the seven polishing stages (x-axis) are decomposed into 16 frequency bands using wavelet packet decomposition. Each subfigure depicts the energy magnitude (y-axis) for one of the 16 frequency bands.

detail coefficients) by passing the original signal through a set of scaled and shifted wavelets. The same decomposition on both approximate and detail coefficients can continue to the next level until the preset level l is achieved. For the l levels decomposition, WPD produces 2^l sets of coefficients corresponding to signal in 2^l disjoint frequency bands, respectively.

García and López [27] provide a guideline on how to choose the mother wavelet and level of decomposition, which should result in the largest adjusted determination coefficient and smallest mean percentage error. Following their selection guideline, we use the biorthogonal 4.4 mother wavelet and decompose the raw signal in the time domain into 16 frequency bands (i.e., $l=4$). Figure 3 exhibits the change of the spectrum energy of each frequency band over seven polishing stages.

Instead of dealing with the 16 frequency bands, we employ the principal component analysis (PCA) method [30] to derive the single functional feature that possesses the largest variance in the signals. PCA is applied to the discretized spectra of the 16 frequency bands. In discretization, the energy of the continuous spectrum is aggregated into the average energy of a certain polishing period. The discretization on the spectra of 16 frequency bands yields a matrix of dimension $|T| \times 16$, where $|T|$ is the number of data points in the time domain after aggregation per stage and $|T|$ is the number of data points for the whole process. In our

analysis, we use $|T|=600$ and will provide more justification for such a choice in Sec. 4.4.

Figure 4(a) displays the scree plot of the principal component analysis using the vibration data of Experiment #1. The same insights are extended to the second experiment as well. The first PC dominates over the rest, whereas the second PC, while several folds smaller than the first PC, is noticeably larger than the third PC and onwards. The third PC and others have small eigenvalues of similar magnitude and can be safely assumed to be noise.

We use PC1 as the feature input fed into the subsequent modeling and analysis. Figure 4(b) shows the change of energy magnitude (scaled to $[0, 1]$) of PC1 over the polishing process, and Fig. 4(c) displays the scores of PCs, which are the coefficients of the 16 frequency bands in PC1. The magnitude of the coefficients suggests how much each frequency band contributes to PC1—the larger coefficient, the greater contribution. If more than one PC is significant, a weighted average of the PCs by their respective eigenvalues, could be used as the feature input. The feature input, hence, is still a one-dimensional functional curve that represents the largest variance of the signals. For our application, we tested using both PC1 and PC2 for the subsequent modeling and analysis versus the option of using PC1 only and did not find any appreciable difference.

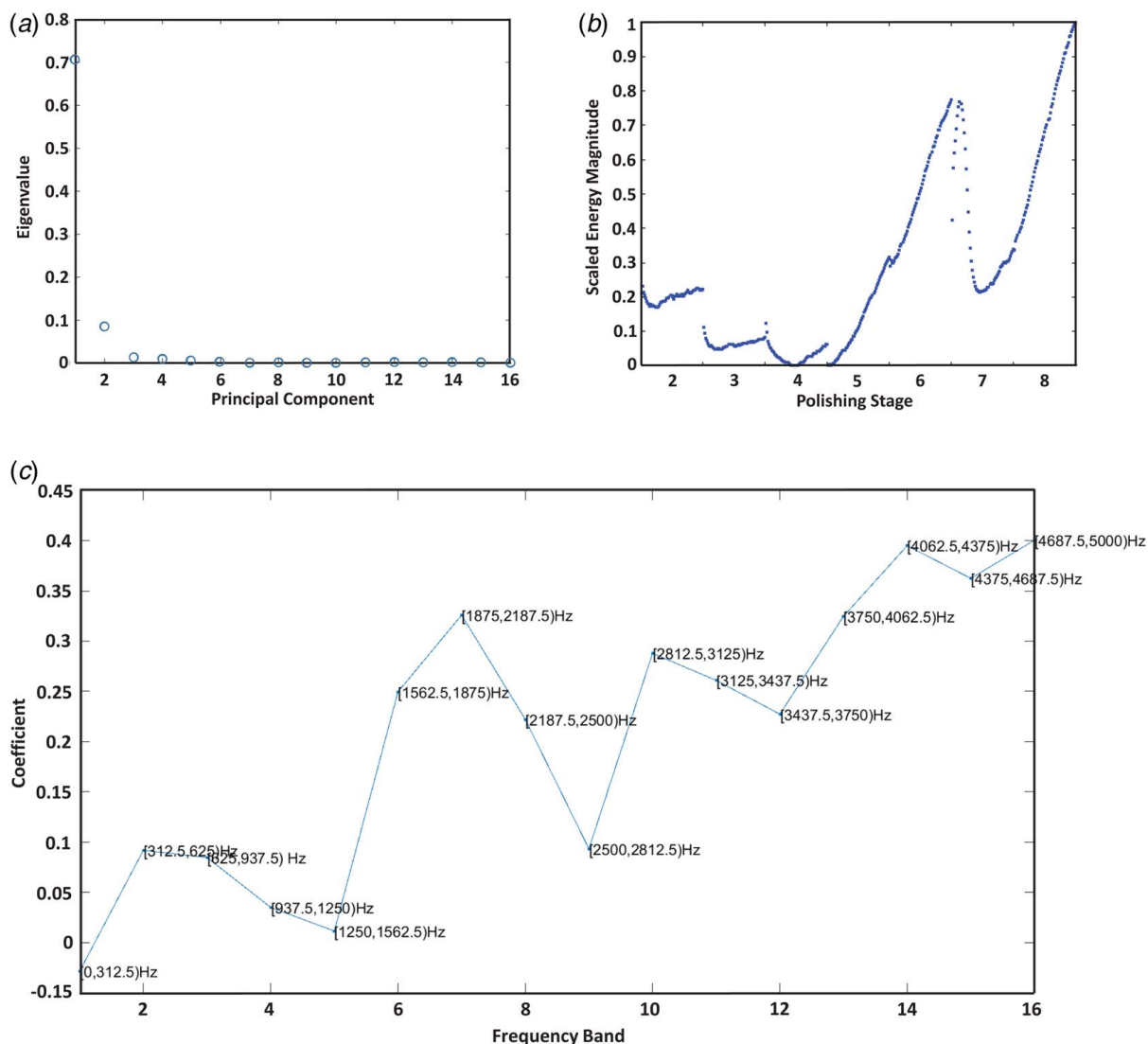


Fig. 4 PCA results on the 16 frequency bands of the vibration signal: (a) the scree plot, (b) the energy spectra of PC1 over the polishing process, and (c) the coefficients of the 16 frequency bands in PC1

Looking at the energy curve in Fig. 4(b), one immediately notices that the energy curve does not show a simple monotonic trend as noted by Hetherington et al. [22] in their polishing process. Toward the later part of the polishing process when the surface roughness is under 30 nm, there is a considerable amount of fluctuation in the energy curve, which leads us to believe that the energy pattern in ultraprecision polishing is simply different from what was observed in the micro-scale polishing.

4 Offline Study Through Functional Regression Model

The modeling objective of the offline study is to explain the average surface roughness (a scalar response) with the continuous signals collected over time (functional input).

The need to connect the scalar response, y , with a functional input, X , naturally points to the use of a functional regression technique [31–34], which was developed for the very purpose of regressing a scalar response on an input function. Neural networks are not a functional regression technique but can take in functional inputs and were indeed used for the same purpose [26]. As reviewed in Sec. 2.2, the characteristics of our polishing process, i.e., long-

stretch and few stages, make it challenging to build neural network models. On top of this difficulty, neural network models lack interpretability, which is the key in this offline study for providing critical insights about the polishing process.

The specific functional regression model we selected is indeed the modeling technique in Ref. [34], which trains a functional slope parameter to reveal the effect of each $x \in X$ on the scalar response. We will show later that this functional slope parameter plays an important role in interpreting the effect of vibration signals on the change of surface roughness.

In Sec. 4.1, we explain in more detail the polishing experiments. In the subsequent subsections, we apply the functional regression model of [34] and discuss the implication of its finding in our polishing problem.

4.1 Experiments. Two polishing experiments are conducted in LLNL for seven or eight stages and 24 h for each stage. Prior to the polishing action of each stage, an assessment of the surface roughness is conducted by measuring multiple inspection locations (white dots in Fig. 1). Throughout the polishing process, the same number of locations are measured on each shell's surface, but

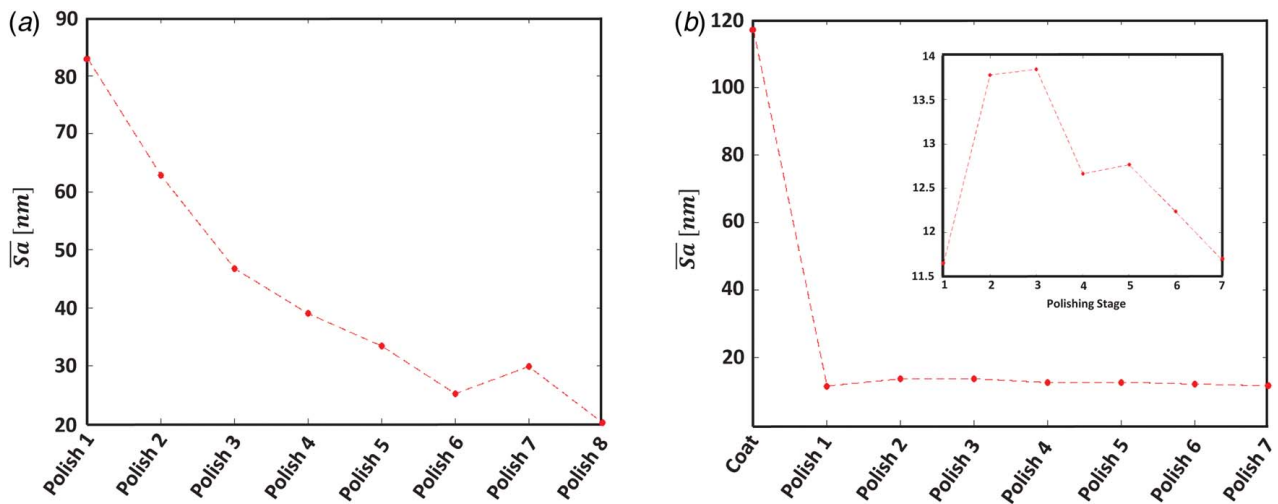


Fig. 5 \bar{S}_a is evaluated at the end of each polishing stage. Panel (a) shows that in Experiment #1 the doped surface roughness gradually evolves during polishing and panel (b) shows that in Experiment #2 the undoped surface roughness significantly decreases during the first polishing stage and fluctuates afterward.

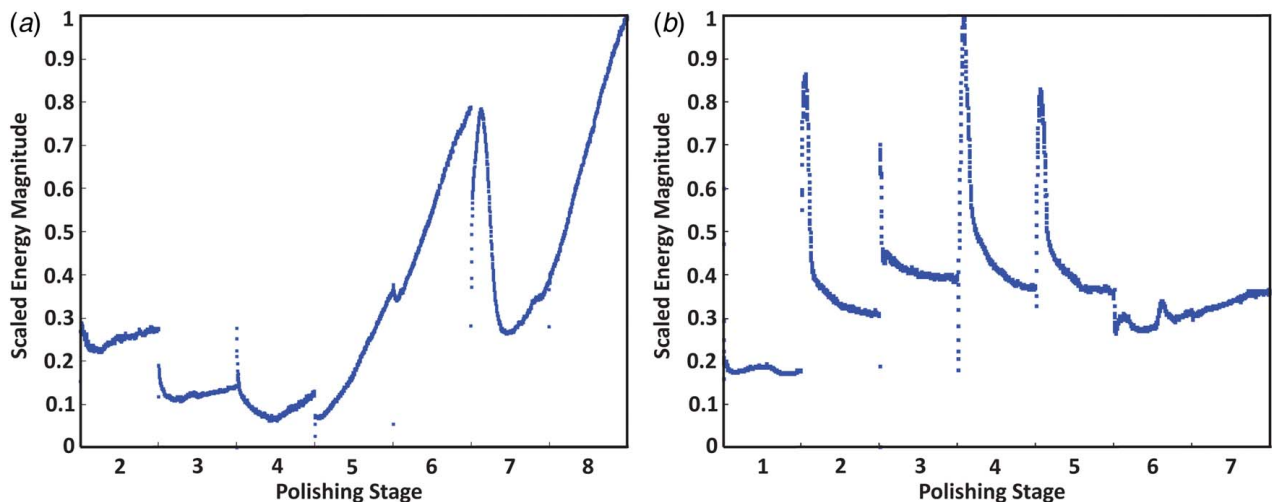


Fig. 6 The PC1 spectra are generated from the PCA on the vibration signal of the two experiments, respectively: (a) Experiment #1 and (b) Experiment #2

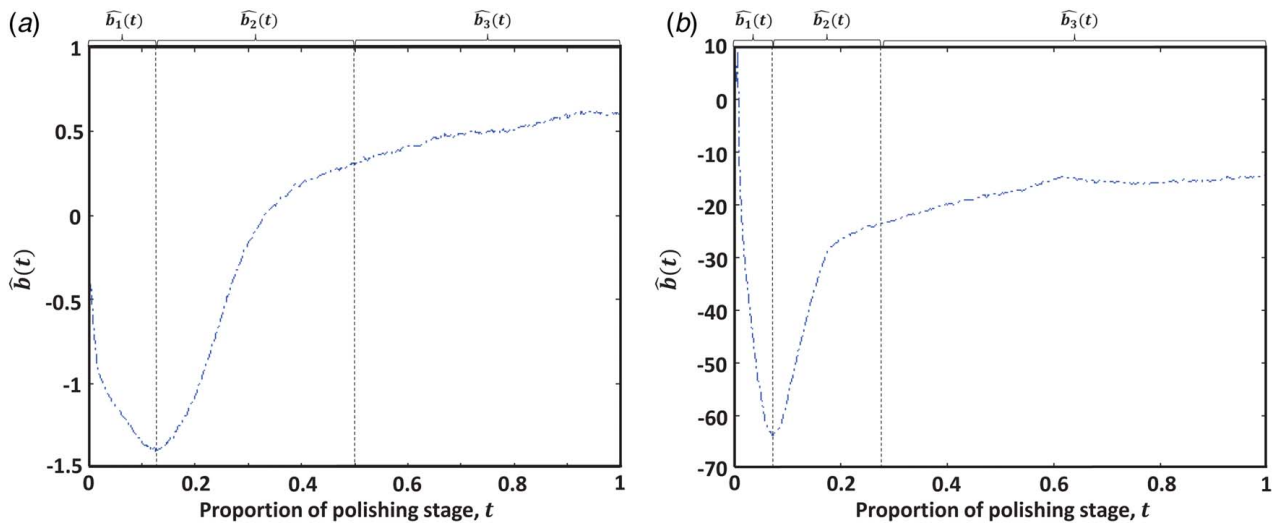


Fig. 7 Estimated coefficient curve, $\hat{b}(t)$, of two experiments: panel (a) is the estimate for the doped surface polishing in Experiment #1 and panel (b) is for the undoped surface polishing in Experiment #2

they are not necessarily the same locations measured at the previous or subsequent stages. The roughness of each location is evaluated by \overline{Sa} . The median value of the Sa values at stage i , denoted by \overline{Sa}_i , is used to represent the roughness of the surface at that stage.

A surface pre-polishing action is undertaken to preserve certain good properties of the shell surface and hence promote its usefulness for the inertial confinement fusion experiments in LLNL [35]. The surface pre-polishing action differs for the two experiments. In Experiment #1, the surface is first coated with diamond and doped with tungsten uniformly on the diamond coating. The doping step makes the polishing progression steady over the subsequent stages, as demonstrated in Fig. 5(a). In Experiment #2, the surface is diamond coated but not doped (or called undoped). The undoped surface roughness can be quickly brought down to around 10 nm even with a single polishing stage. For the undoped shells, the surface roughness fluctuates within a narrow range after the first stage.

By “a stage,” we refer to the polishing action between two pauses. The surface roughness measurements are taken at the end of each polishing stage. The surface roughness difference at the end of two consecutive stages (Stage $i-1$ and Stage i) reflects the performance of the polishing action in Stage i . We denote the surface roughness difference by $\Delta\overline{Sa}_i = \overline{Sa}_{i-1} - \overline{Sa}_i$. In Experiment

#1, the surface roughness measurement before the first polishing action is lost, so the first available surface measurement is obtained after one polishing action and denoted as “Polish 1” in Fig. 5(a). In Experiment #2, the first available surface measurement is after coating but before polishing and denoted as “Coat” in Fig. 5(b). For both experiments, we fit the functional regression model on the seven-stage data.

The vibration signals of each stage are collected while polishing is ongoing. We extract PC1 from the vibration signals following the procedures in Sec. 3. As discussed in that section, PC1 of each stage is the weighted average spectral curve of the 16 frequency bands, i.e., a univariate continuous function of time t . We denote the PC1 of Stage i by $X_i(t)$, $t \in [0, 24]$ and in the unit of hour. For generality, we can scale t such that $t \in [0, 1]$ with each decimal representing a temporal proportion of the polishing stage. Figure 6 illustrates the $X_i(t)$, $i = \{2, \dots, 8\}$, in Experiment #1 and that for $i = \{1, \dots, 7\}$ in Experiment #2.

4.2 Modeling. The functional linear regression model in Cai and Hall’s work provides a way to connect a scalar response with a functional input [34]. For our polishing monitoring and control problem, the functional input, $X_i(t)$, is the first principal component of the vibration signals in the time-frequency domain, and the scalar

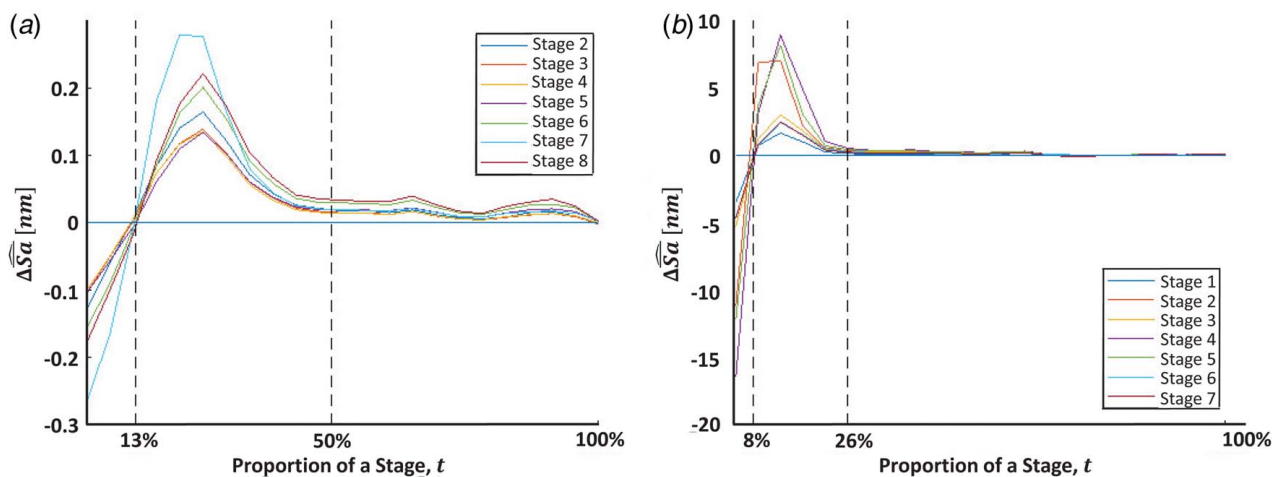


Fig. 8 The estimated marginal Sa over 24 h of all and each of the stages in the two experiments. Along the x-axis, x% of a stage corresponds to the end of the x%-24th h. For example, 50% of a stage corresponds to the end of the 12th hour: (a) Experiment #1 and (b) Experiment #2.

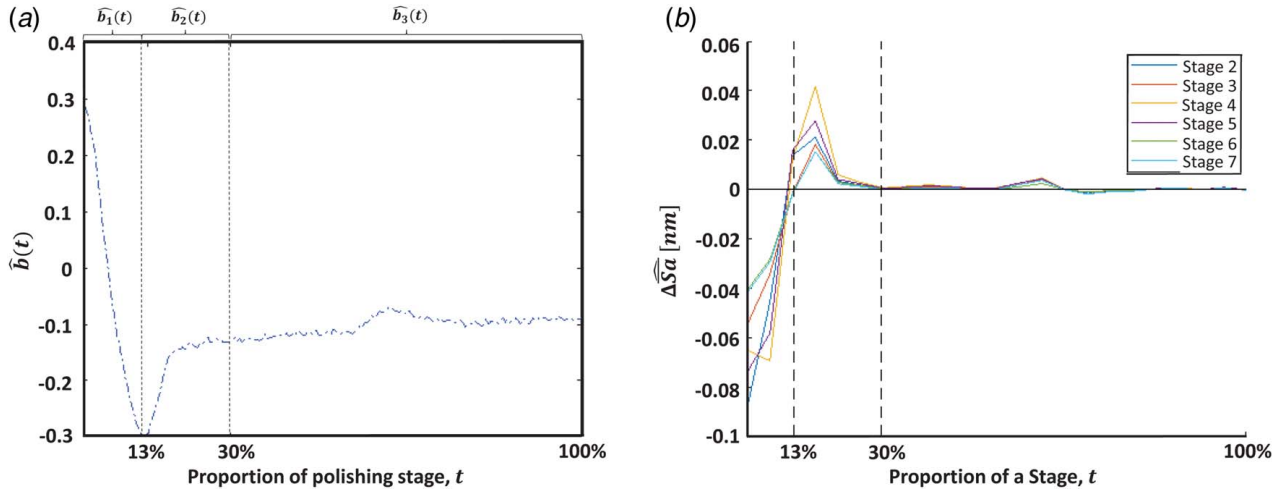


Fig. 9 The estimation results for Experiment #2 with Stage 1 removed for fitting the functional regression model: (a) estimated $\hat{b}(t)$ and (b) estimated marginal \hat{S}_a

response, y_i , is the change in surface roughness (scalar response), ΔS_{a_i} , after Stage i . The model in Ref. [34] is represented in our notation as

$$y_i = a + \int_T b(t)X_i(t) dt + \varepsilon_i \quad (1)$$

where T denotes the domain of the independent variable t . A constant parameter a and a functional coefficient parameter $b(t)$ are to be estimated from the model. The noise term, ε_i , is assumed to follow a normal distribution. The task of model fitting is to estimate the parameters a and $b(t)$; their estimates are denoted by \hat{a} and $\hat{b}(t)$, respectively.

The estimation of a is straightforward, that is $\hat{a} = \bar{y} - \int_T \hat{b}(t)\bar{X}(t)dt$, assuming that $\hat{b}(t)$ is already estimated. But the interpretation to a is not trivial, although it is not a function of time. The functional regression model (in Eq. (1)) assumes a constant a throughout the stages of the whole polishing process. In practice, it sensibly represents a constant polishing effect of each stage, given a set of polishing process parameters.

The estimation of $b(t)$ is an infinite-dimensional problem. The idea in Ref. [34] is to apply the functional PCA for dimension reduction. Let K denote the covariance function of X that is positive definite and equipped with a spectral decomposition of (θ_j, ϕ_j) , i.e., the (eigenvalue, eigenfunction) pairs of the linear operator of kernel K . The estimation of K with its spectral decomposition is

$$\begin{aligned} \hat{K}(u, v) &= \frac{1}{n} \sum_{i=1}^n \{X_i(u) - \bar{X}(u)\} \{X_i(v) - \bar{X}(v)\} \\ &= \sum_{j=1}^{\infty} \hat{\theta}_j \hat{\phi}_j(u) \hat{\phi}_j(v) \end{aligned} \quad u, v \in T \quad (2)$$

where $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ and $\hat{\theta}_j$ is ordered such that $\hat{\theta}_1 \geq \hat{\theta}_2 \geq \dots$. Cai and Hall [34] proposed to estimate the function b with its Fourier series, i.e.,

$$\hat{b} = \sum_{j=1}^m \hat{b}_j \hat{\phi}_j \quad (3)$$

where b_j is estimated by $\hat{b}_j = \hat{\phi}_j \hat{g}_j$ and g_j denotes the j th Fourier coefficient of $g(u) = \int_T K(u, v)b(v) dv$ and $\hat{g}_j = \int_T \hat{g} \hat{\phi}_j$. If one evaluates $\hat{b}(v)$ with a consistent estimator $\hat{b}(v) = \frac{1}{n} \sum_{i=1}^n \{X_i(v) - \bar{X}(v)\}^{-1} (y_i - \bar{y})$, then $\hat{b}(t) = \sum_{j=1}^{\infty} \hat{\theta}_j \hat{\phi}_j(t) \hat{g}_j \approx \sum_{j=1}^m \hat{\theta}_j \hat{\phi}_j(t) \hat{g}_j$, where m is the number of functional principal components.

The choice of the cutoff value for the number of principal components, m , is related to the cutoff for the eigenvalue $\hat{\theta}_j$. Let us

denote the latter cutoff by δ . Cai and Hall proposed a number of potential approaches for choosing δ . The variety of methods includes least squares, bootstrap, and cross validation. In our work, we use the least square method to choose δ in the set of $\{0.001, 0.01, 0.05, 0.1, 0.15, 0.2\}$, such that the square root of the mean squared error (RMSE), i.e., $\sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2/n}$, is minimal. For more derivation of the equations, one may refer to Ref. [34].

Applying Cai and Hall's functional regression model to the two experiments data, we obtained the resulting $\hat{b}(t)$ curves, as shown in Figs. 7(a) and 7(b). The coefficient curve indicates the leverage of the input variable in contributing to the output variable over t . In our case, the coefficient curve reflects the average polishing effect over time during a given polishing stage. Here we say "average" because the resulting $\hat{b}(t)$ is estimated using data across all seven stages. We label this analysis an offline study because such an average model is not stage-specific and is not intended for prediction purpose in real-time decision-making.

The $\hat{b}(t)$ curves estimated from both the experiments present a similar pattern over time: the $\hat{b}(t)$ curve quickly decreases in the first portion of the polishing stage, followed by a rapid increase in the second portion, and then followed by a flat third portion. The specific time durations for the two experiments are different. If denoted the three portions by $\hat{b}_1(t)$, $\hat{b}_2(t)$, $\hat{b}_3(t)$, respectively, Experiment #1 has $\hat{b}_1(t)$ up to $t = 13\%$, $\hat{b}_2(t)$ for $13\% < t \leq 50\%$,

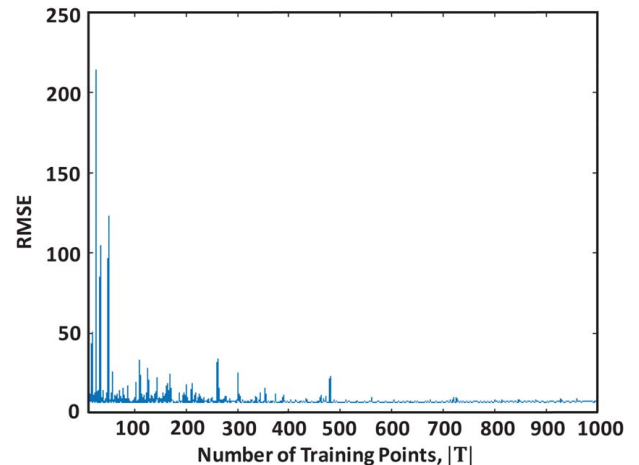


Fig. 10 The RMSE remains at the low value as the number of training points, $|T|$, is greater than 500

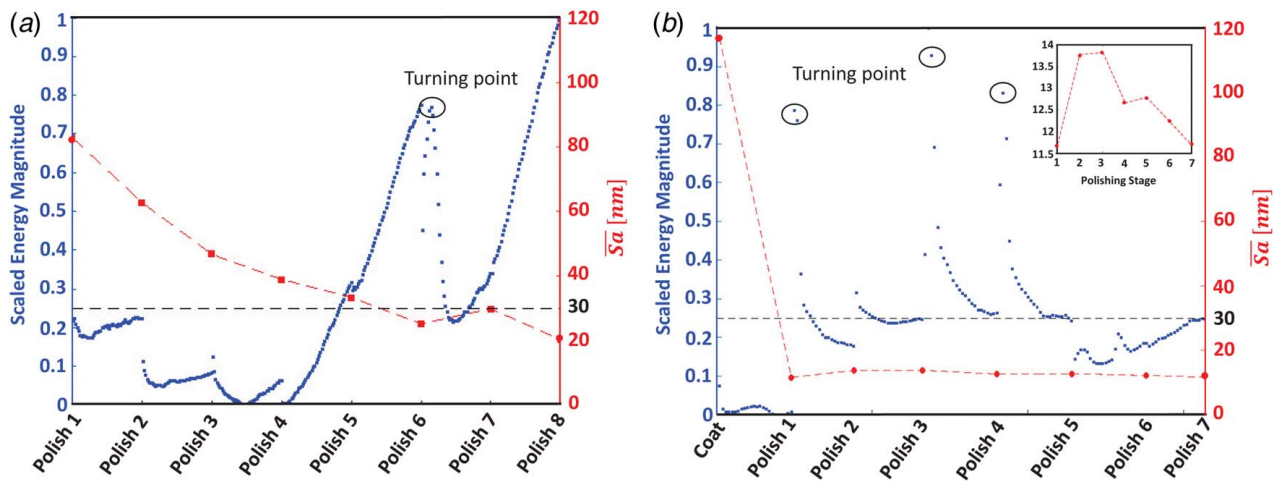


Fig. 11 The scaled PC1 spectra of vibration and \overline{Sa} for both polishing experiments. The small insert in the right panel is the zoom-in view of Polish 1 through Polish 7 where \overline{Sa} fluctuates in a narrow range between 11.5 nm and 14 nm: (a) Experiment #1 and (b) Experiment #2. (Color version online.)

and $\hat{b}_3(t)$ for the rest half of the duration, whereas Experiment #2 has $\hat{b}_1(t)$ up to $t = 8\%$, $\hat{b}_2(t)$ for $8\% < t \leq 26\%$, and $\hat{b}_3(t)$ for the rest three-quarter of the stage. The constant \hat{a} has very different values in the two experiments: $\hat{a} = 8.9$ nm in Experiments #1 and $\hat{a} = 112.2$ nm in Experiment #2. The difference in \hat{a} of two experiments stems from the different polishing progressions. The large \hat{a} in Experiment #2 is due to the steep reduction in Sa after its first stage.

4.3 Results and Discussions. The offline model nonetheless offers useful insights. In Eq. (1), there are two sets of parameters, the constant regression term, a , and the functional coefficient, $b(t)$. The term a represents the average effect of the polishing action (like how much the Sa has been reduced), whereas the $\int_0^t \hat{b}(t)X(t)dt$ term reflects the dynamic effect after removing the constant effect, which is how much it is varying overtime in the rest of the signals. Based on the $b(t)$ estimation and the input vibration $X(t)$, we plot the estimated incremental surface roughness, $\Delta\widehat{Sa}$, for the two experiments in Fig. 8. A positive $\Delta\widehat{Sa}(= \widehat{Sa}_t - \widehat{Sa}_{t-1})$ on the vertical axis indicates a surface quality improvement, while a negative $\Delta\widehat{Sa}$ suggests a surface quality deterioration aside the constant effect. The absolute value of $\Delta\widehat{Sa}$ indicates how fast the surface quality evolves (either improves or deteriorates) with time. A close to zero $\Delta\widehat{Sa}$ apparently means that the surface quality barely changes.

Experiment #1 has a steady and even progression over stages, as shown by a similar change in Sa , which is a 10 nm or so difference of every two consecutive stages (see Fig. 5(a)). The constant effect, $\hat{a} = 8.9$ nm, close to the above observation, explains the average effect in Experiment #1. Aside from the constant effect, the residuals of Sa for all the stages are reflected by the dynamic effect in a small magnitude in Fig. 8(a). In Experiment #2, Sa is considerably reduced in the first stage (see Fig. 5(b)), reflected by the large constant effect, $\hat{a} = 112.2$ nm. After removing this large constant effect, the dynamic effect, $\Delta\widehat{Sa}$, is shown in Fig. 8(b). The dynamic effect of Stage 1 is not pronounced. From Stage 2 onward, Sa remains at the same level, but its fluctuation varies and the dynamic effect of the latter stages could be greater than that of Stage 1.

The \widehat{Sa} curves in both experiments (in Fig. 8) show a distinct pattern that delineates the polishing process in each stage into three phases, which is consistent with the three portions observed in the $\hat{b}(t)$ curves. The first phase, corresponding to $\hat{b}_1(t)$, is the “running-in” phase, showing that the surface quality deteriorates,

but the deterioration is in deceleration. The second phase, consisting of positive $\Delta\widehat{Sa}$ and corresponding to $\hat{b}_2(t)$, indicates a progressive surface quality improvement with the maximum effect at the peak of the curve. This second phase is known as the “steady” phase. The third phase, known as the “saturation” phase, shows an apparent pattern of small fluctuations without any significant change in \widehat{Sa} , meaning that during the saturation phase, the surface roughness does not change considerably.

We deem the running-in and steady phases together as the effective phase, where the effective polishing takes place before it gets into the saturation phase. The proportions of the effective phase in the two experiments are different (see Fig. 8): 50% for Experiment #1 and 26% for Experiment #2. The different length of the effective phase in the two experiments is consistent with the polishing progression on the two different types of surfaces. On the doped surface, roughness (in Fig. 5(a)–Experiment #1) gradually evolves over eight stages, and the effective phase of the polishing lasts longer (up to 12 h), whereas the effective phase in Experiment #2 when polishing the undoped surface shortens considerably, down to 6 h.

The revelation of the three phases is hardly a surprise. There were previous studies alluding to a similar understanding. For instance, Tang et al. [15] named three phases in terms of MRR: the loading phase, self-accommodation phase, and equilibrium phase. The time scale in Ref. [15] was rather different. The loading phase lasts for about 2–3 s when the sensor signal rises while the product surface initially contacts the polishing tool and achieves its local peak value. After the loading phase, it is the self-accommodation phase during which the physical-mechanical characteristics of the product surface may undergo considerable changes due to the set-down of the polishing tool on the product surface. This phase may last from several seconds to several tens of seconds. The last phase claimed in Ref. [15], i.e., the equilibrium phase, is a stabilization period of the MRR.

The phases of the polishing effect revealed in our study show consistency with the understanding discussed in Ref. [15], but the phases are not one-to-one correspondences. In fact, the loading and self-accommodation phases in MRR help explain the running-in phase in polishing, and the equilibrium phase in MRR explains the steady and saturation phases in polishing.

In every frequency band shown in Fig. 3, at the beginning of each stage, the signal shows as a vertical line. This vertical line presents evidence of the loading and self-accommodation phases. The set-down of the polishing tool seems to create a profound effect, due to which defects may happen, and the defect-

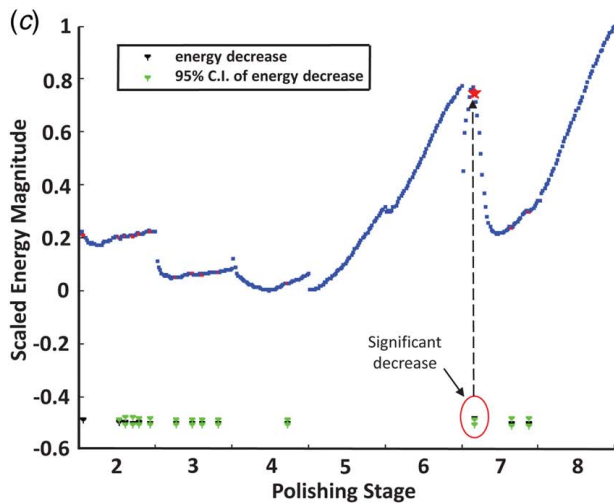
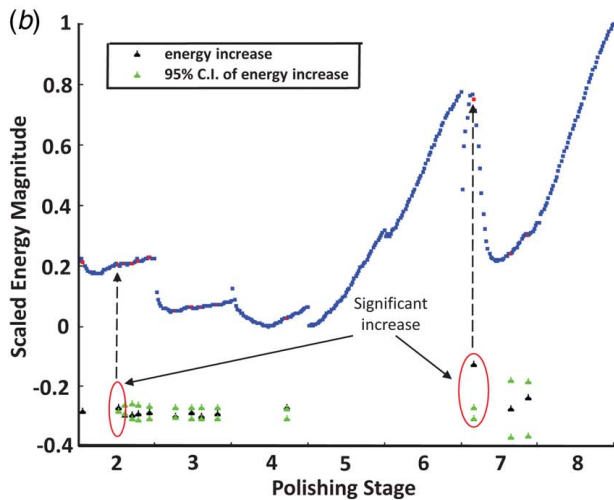
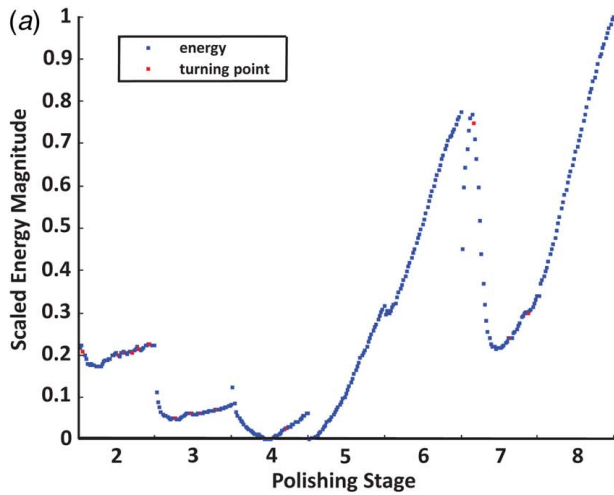


Fig. 12 Energy turning point detection for Experiment #1. Panel (a) shows the turning points (red dots) that pass the derivative checks in Step 2 of Algorithm 1, panel (b) plots the energy increase (black up-pointing triangle) and the 95% confidence interval (green up-pointing triangles), and panel (c) exhibits the energy decrease (black down-pointing triangle) and their 95% confidence intervals (green down-pointing triangles). The significant turning point is marked by the red star. (Color version online.)

correcting period may last for some time. That explains why the running-in phase of the shell polishing extends rather long (about 1–3 h).

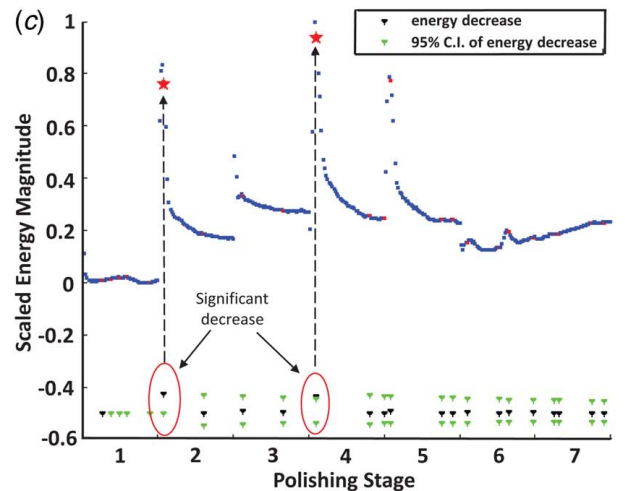
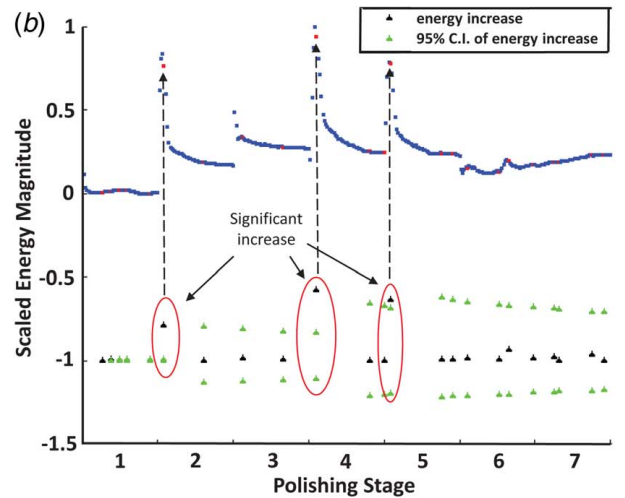
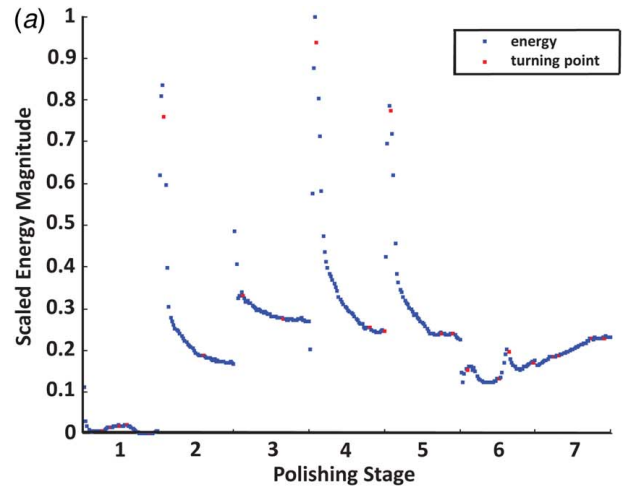


Fig. 13 Energy turning point detection for Experiment #2. Interpretation of panels (a), (b), and (c) is the same as in Fig. 12.

During the equilibrium phase in MRR, the polishing effect exhibits both steady progression and the saturation phenomenon. Polishing is a cycling process in which the pad abrasives, in an iterative fashion, destruct the surface's uniformity, remove the higher peaks and flatten the local surface. The polishing effect shows a steady progression when the pad abrasives have deep contact with surface heights. When the surface's high peaks are polished off, the pad abrasives barely touch the surface heights, and only a small portion of heights can be flattened. That is

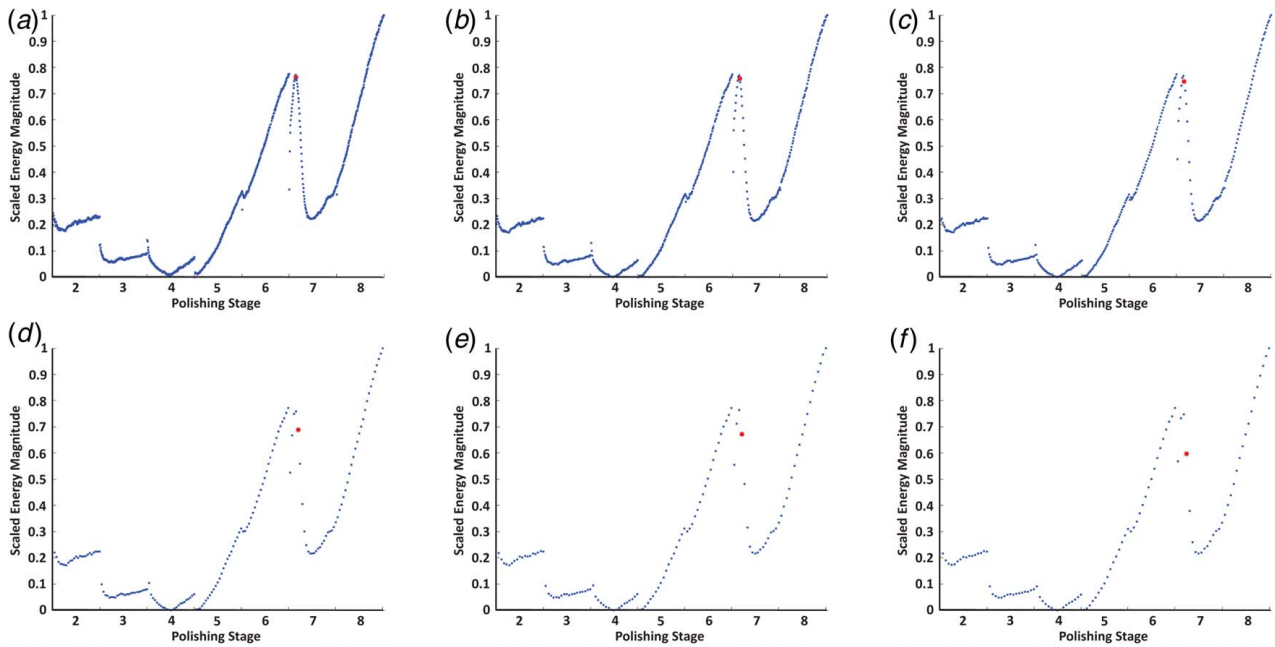


Fig. 14 The sensitivity of the detection algorithm to data aggregation time intervals in Experiment #1. In (a)–(f), the signals are aggregated every 10, 20, 30, 60, 80, and 90 min, respectively.

when the fluctuation in roughness, or the saturation in the polishing effect, occurs.

While the general understanding of these polishing exists, to our best knowledge, we are the first team who provides a mathematical model for quantifying the time duration of each respective phase. The significance of this quantification is the finding of a significantly long saturation phase. Its existence provides a great opportunity for shortening each polishing stage 50–75% of the time. Even considering some model inaccuracy and uncertainty, we believe it is rather safe to start some experiments by shortening the polishing time by 25%, which is 6 h. This suggests that polishing up to 18 h for seven (or eight) stages would have produced more or less the same surface roughness outcomes at the end. That would lead to substantial time and cost savings on such an operation.

4.4 Robustness and Sensitivity Analysis

4.4.1 Model Robustness. The polishing effect of Stage 1 in Experiment #2 is much different than the other stages. Does its existence dominate or distort the understanding garnered in the above analysis revealing three polishing phases? We therefore analyze the robustness of the offline study by removing the first stage when fitting the functional regression model. The estimation results are presented in Fig. 9, which show a similar pattern as in Figs. 7(b) and 8(b) where data from all seven stages are used. When using the last six stages, the estimated constant effect, $\hat{a} = -0.44$ nm, meaning that on average, the last six stages do not reduce S_a . This value is very different from \hat{a} value when using all seven stages, confirming that the previous large, positive a value reflects primarily the effect of the first stage. The dynamic polishing effect still shows clearly the three polishing phases but the saturation phase is now about 70% of the polishing stage, slightly shorter than the previous estimate at 75% for the saturation phase.

4.4.2 Sensitivity of Training Points. The fitting performance of the functional linear regression model depends on the feature function, $X(t)$. As the model takes in the discretized version of $X(t)$, how much of the original signal can be preserved is primarily determined by the number of training points along the time dimension, $|I|$. Using $|I|$, one can also compute the time intervals by which we divide the vibration signal in each polishing stage. Considering

that it is 24 h per polishing stage and suppose that the vibration signal is aggregated at equal-distanced intervals, $|I| = 10$ corresponds to aggregating the vibration signal every 144 min, and $|I| = 1,000$ is for aggregation every 1.44 min.

Figure 10 presents the fitting performance in terms of RMSE versus the choice of $|I|$. RMSE decreases as the number of training points increases and retains a level of stability around $\text{RMSE} \approx 6.7$ when the number of training points is greater than 500, which corresponds to aggregating the vibration signal every 2.88 min. Therefore, we recommend setting $|I| > 500$ for this type of polishing processes. For a new polishing process, practitioners can run the same analysis to determine the appropriate $|I|$.

5 Online Detection

The offline model in Sec. 4 reveals the potential of shortening the polishing process per stage by 50% to 75% of the time. But the offline model is not intended to guide polishing decision-making in real-time. In this section, we discuss an algorithm that provides the online, real-time detection capability for nano-scale polishing.

Recall that at the end of Sec. 3, we noted the disagreement between our data and the observation made by Hetherington et al. [22], who found that the vibration energy intensity decreases as the surface roughness reduces. Such a monotonic pattern certainly does not happen in our nano-scale polishing process. Figures 11(a) and 11(b) plot the vibration energy (the blue dotted curves scaling on the left-hand vertical axis) and $\overline{S_a}$ (the red dashed lines scaling on the right-hand vertical axis) for Experiments #1 and #2, respectively. The patterns are evidently more complicated. In Experiment #1, Stages 5, 6, and 7 experience higher vibration energy but lower $\overline{S_a}$ than Stages 2 and 3. In Experiment #2, from the first polishing stage onward, $\overline{S_a}$ is brought down to around 12 nm from the initial 118 nm but the vibration energy is, although varying, nevertheless higher than that in the first stage. We conjecture that the discrepancy in our data and Hetherington et al.'s observations could be due to the scale difference—our polishing is in nano-scale (the surface roughness goes below 30 nm), whereas the polishing process observed by them is in micro-scale, from 475 μm to 20 μm [22]. That is three to four orders of magnitude higher than the surface roughness of our nano-scale polishing

process. We believe the physics behind is dissimilar at such drastically different surface roughness scales.

5.1 Turning Point Detection Algorithm. We have to look for different clues to enable online change-point detection than using simple monotonicity. When we take a closer look at Figs. 11(a) and 11(b), we notice the appearance of peaks on the energy curves right before the polishing goes into the saturation stage. The occurrence of the peaks may suggest the onset of saturation and indicate a polishing endpoint. In Experiment #1, such a phenomenon happens after Stage 6 and during Stage 7 when the surface roughness is under 30 nm. In Experiment #2, this pattern repeatedly occurs after the first polishing stage since the surface roughness has been under 30 nm for all remaining stages.

In order to identify the right peaks, the first line of action is to detect the maximum energy points within a polishing stage. Mathematically speaking, a maximum energy point is a stationary or turning point on a concave curve. More specifically, the turning point should satisfy two conditions: its first derivative at t is equal to zero (for being a stationary point), and its second derivative at t is negative (for being a maximum rather than a minimum). By checking these two conditions, we can identify the set of turning points, which constitutes the candidate set for us to identify the peaks.

In a noisy data set, there are inevitably numerous turning points—some trivial and others significant. Only the significant ones warrant attention. To determine which turning point is significant, we need to test the statistical significance of the energy increase leading to that point and the energy decrease right after. This needs two actions. The first is to calculate the energy increase and decrease associated with a candidate point, and the second is to estimate the energy fluctuation within the current polishing stage and up to the candidate point.

The calculation of the energy increase leading to a candidate point is the accumulation of energy increase since the previous stationary point. This may involve multiple points if they are on a persistently increasing trend. The calculation of the energy decrease after the candidate point is, in principle, done in the same way, but there is a practical constraint to consider. In order to flag the significant turning point as soon as possible, one does not want to collect too many points after the candidate point; doing so will delay the decision of detection. We, in the end, decide to use one single point immediately after the candidate point to calculate the energy decrease (more discussion in the next paragraph). We acknowledge that this action sounds greedy, but our test using actual polishing data shows it is rather effective without causing any more delay in detection than what is already part of the algorithm. For estimating the energy fluctuation, we set a time window, which is from the beginning of a stage up to the current candidate point (and includes the point right after), and then calculate the energy increases/decreases associated with all turning point candidates in that time window. We construct the 95% confidence interval of the energy fluctuation for the current candidate and use that to check whether the current candidate is a significant one. Algorithm 1 summarizes the above-described procedure. $E(\cdot)$ and $\sigma(\cdot)$ in the algorithm represent the mean and standard deviation, respectively.

We want to add a note about the delay in detection mentioned above. In the online monitoring problem, a turning point will not be identified until the next point is available. This means that we can only remit the check on those aforementioned mathematical and statistical conditions for the signal at time t until after the data point at $t + 1$ becomes available. This causes a slight delay in detection. The actual delay depends on how often the vibration signals are processed into energy points. The default setting in our algorithm is every 30 min, suggesting a half-hour delay; such a delay is considered well tolerable for these 24-h long-stretch polishing stages. For calculating the energy decrease, when we only use one point right after the candidate point, we do not add any additional delay, as the next point is needed for checking derivatives anyway.

Algorithm 1 The turning point detection based on tracing significant energy change

Initialize $t = 0, X(0) = X'(0) = X''(0) = 0$, where X' and X'' are the first and second derivatives of X , respectively.

$\mathcal{N}_{up} = \mathcal{R}_{up} = \mathcal{R}_{down} = \emptyset. t = t + 1.$

At each stage $t, t \in T,$

- (1) **Track and record the trend of energy increases.**

Record $\mathcal{N}_{up} = \mathcal{N}_{up} \cup \{X(t-1), X(t)\}$, if $X(t) - X(t-1) \geq 0$

The calculation of energy decreases is easier and done in Step 3.

- (2) **Detect the turning point on a concave curve.**

Calculate $X'(t-1)$ and $X''(t-1)$.

If $X'(t-1) = 0$ and $X''(t-1) < 0$, go to Step 3; otherwise, go to Step 4.

- (3) **Determine if the turning point is significant.**

Calculate $R_{\mathcal{N}_{up}} = \max(\mathcal{N}_{up}) - \min(\mathcal{N}_{up})$ and reset $\mathcal{N}_{up} = \emptyset$

$R_{\mathcal{N}_{down}} = X(t) - X(t-1)$

Record $\mathcal{R}_{up} = \mathcal{R}_{up} \cup \{R_{\mathcal{N}_{up}}\}$

$\mathcal{R}_{down} = \mathcal{R}_{down} \cup \{R_{\mathcal{N}_{down}}\}$

Check if both $R_{\mathcal{N}_{up}}$ and $R_{\mathcal{N}_{down}}$ exceed the upper 95% confident interval of \mathcal{R}_{up} and \mathcal{R}_{down} , respectively, i.e.,

$$R_{\mathcal{N}_{up}} \geq E(\mathcal{R}_{up}) + 1.96 \times \sigma(\mathcal{R}_{up}) \quad (4)$$

and

$$R_{\mathcal{N}_{down}} \geq E(\mathcal{R}_{down}) + 1.96 \times \sigma(\mathcal{R}_{down}) \quad (5)$$

If both Eqs. (4) and (5) hold, **suggest a stop/check point at time t** ; otherwise, go to Step 4.

- (4) **Continue polishing to the next stage.**

Set $t = t + 1$. Go to Step 1.

5.2 Experimental Results. The vibration signals are collected continuously and processed every 30 min in our analysis. The signals are decomposed into 16 frequency bands using the wavelet packet decomposition and further converted into the most significant PC of these frequency bands (or a weighted average of the PCs if multiple PCs are significant). Then the PC of the signals over the 30 min duration is transformed into an energy point, $X(t)$; for instance, $X(t)$ at $t=1$ is the average point of the signals in the first 30 min segment.

The detection results for the two experiments are shown in Figs. 12 and 13, respectively. In Experiment #1, only one turning point exceeds the 95% confidence interval of both energy increases and decreases. The significant turning point is at the beginning of Stage 7 (roughly at the fourth hour). The red star, marked in Fig. 12, indicates the detection time point, which is one point after the actual significant turning point. This is the 30 min delay we talked about earlier.

Should one stop and check the polishing process at this point, as Algorithm 1 suggests, one would most likely find the surface satisfactory. It would save the rest of Stage 7 and the entire Stage 8, which amounts to more than 44 h of polishing.

In Experiment #2 (Fig. 13), the turning point at the beginning of Stage 2 is detected. This point captures the start of the saturation phase, soon after the surface roughness Sa decreases below 30 nm. Had this turning point been detected and used for decisions, it would provide an early suggestion for stopping the polishing and avoiding the subsequent inefficient polishing operation. As shown by the flattened Sa line (the red dashed line) in Fig. 11(b), the polishing process continues for 120 h from Stage 2 onward but barely improves Sa . In case the first significant turning point is missed, the detection method would pick up another stopping point in Stage 4 of a similar pattern. Of course, had the polishing process been stopped after the first turning point, the subsequent stages of operation would not happen anymore, and thus the second turning point would not exist. We want to note that not all seemingly peaks in the

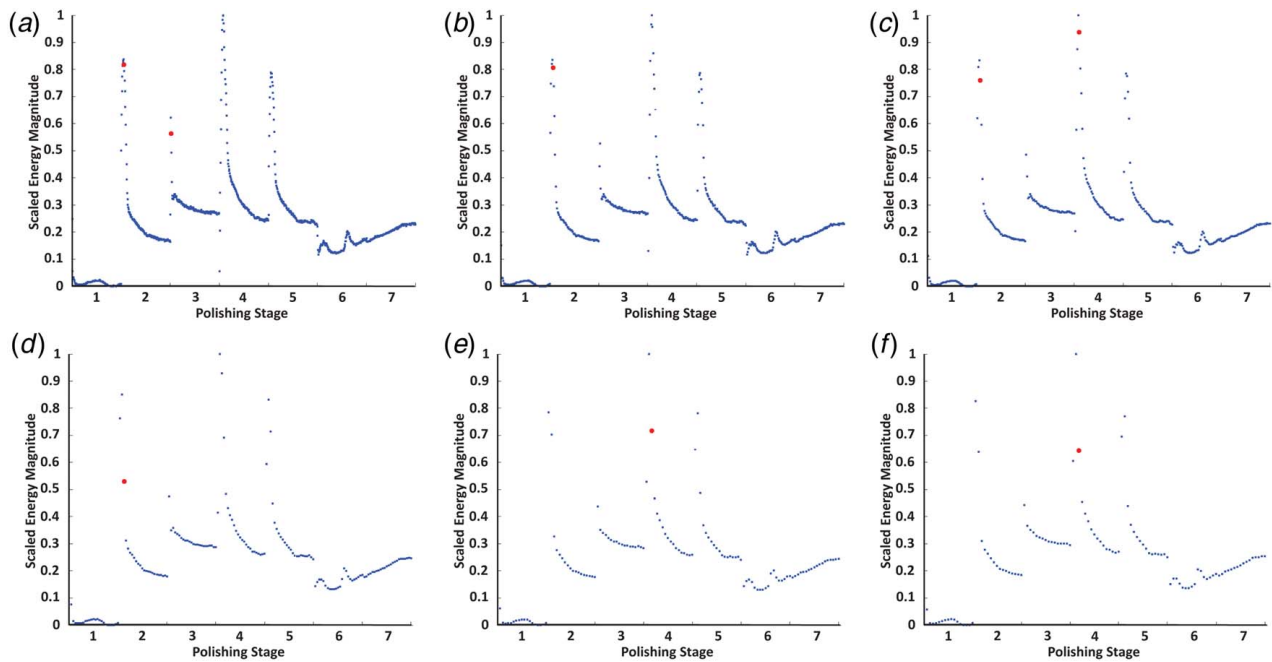


Fig. 15 The sensitivity of the detection algorithm to the data aggregation time interval in Experiment #2. In (a)–(f), the signals are aggregated every 10, 20, 30, 60, 80, and 90 min, respectively.

plots meet our standard of significant turning points. Consider the peak at the beginning of Stage 3. This point was not deemed significant because neither the energy increase nor the decrease is found significant. For the peak at the beginning of Stage 5, its energy increase is significant but its energy decrease is not. The statistical significance check based on the 95% confidence interval is a consideration of the signal-to-noise ratio. Showing a peak (the signal) itself is not enough for detection but needs to be large enough relative to the noise that is estimated through the data up to the current time point.

5.3 Sensitivity of Signal Aggregation Frequency. The detection results in Sec. 5.2 are based on the energy points aggregated every 30 min. We refer to this 30 min as the data aggregation time interval. Here, we conduct a sensitivity analysis of the detection results when the data aggregation time interval varies from 10 min to 90 min. Figures 14(a)–14(f) plot, respectively, the detection results of the significant turning point in Experiment #1 when the data aggregation time interval is 10, 20, 30, 60, 80, and 90 min. The figure shows a consistent detection outcome. With different data aggregation time intervals, the delay in detection, caused by the need to use the energy point after the significant turning point, also ranges from 10 min to 90 min. Even with a 90-min delay, a good detection is still a benefit, considering that such detection could shorten the polishing process by 44 h. Our default choice of 30 min is for alleviating such delay while also maintaining a good margin for robustness.

For Experiment #2, as shown in Fig. 15, the earliest turning point in Stage 2 is detected consistently when the data aggregation time interval is set from 10 min to 60 min. When the data aggregation time interval is 80 min and 90 min, the algorithm misses detecting the first turning point but still detects the turning point in Stage 4. If the significant turning point is detected in Stage 2, the time saving would be 140 h, whereas if it is detected in Stage 4, the saving would be 90 h. In either case, the saving is significant. Overall, we consider the detection outcome in Experiment #2 also fairly robust.

Upon the analysis and discussion above, we suggest setting the data aggregation time interval between 10 min and 60 min, with the default setting at 30 min. Longer than 60-min aggregation

time intervals not only further delay the detection but could also average out certain effects in the signals, causing missed detection like in Experiment #2. Shorter than 10-min aggregation time intervals are not advisable either. Short aggregation is more prone to fluctuation in data. Also, recall that the polishing machine usually takes a few minutes to settle at the beginning of each stage. Too short an aggregation time interval could pick up too much of the transition effect.

6 Concluding Remarks and Future Work

Two methods are proposed in this work to understand the surface roughness change and inform more efficient polishing operations for long-stretch nano-scale polishing in which only infrequent offline surface measurements are conducted. To our best knowledge, it is the first work to address the endpoint detection problem for such a unique yet important polishing process.

The offline study reveals the presence of a significantly long saturation phase of this type of polishing processes, unknown or at least unsuspected previously. Such insights lay the ground for considerably shortening the polishing duration per stage, potentially saving half of the original polishing time. The online detection method uses the in-process vibration signal and, with only a slight delay, flags the real-time stopping point during the operations. Equally significant is the finding that the vibration energy curve exhibits remarkably different and, in fact, more complicated patterns in nano-scale polishing than in micro-scale polishing. We believe that we are among the very first to report this finding.

The modeling framework is general enough and can be readily used for other polishing processes of the same nature, i.e., having long-stretch stages and with nano-scale precision. We believe that the existence of the three phases and the vibration intensity pattern are also generalizable, although the length of the saturation phase and the timing of the turning point would naturally vary when different materials are used, or other processing parameters are set.

At least two lines of future work are in need of pursuit. The first is to validate the effective/saturation phenomena learned from our offline study and the efficiency of polishing guided by our online detection method. Due to the high cost and other priorities of LLNL where the original polishing was conducted, such a

validation experiment is being discussed but has yet to be scheduled. We call for other researchers with a similar capability to test our recommendations if possible and interested. The second line of future work is to further the understanding of the physical differences between polishing at the nano-scale and those at other macro-scales. Does the material medium also matter when the polishing scale changes? For instance, will the energy curve differences, as we observed, persist even if the coating material is not diamond? Physical insights into such questions, combined with data science models as proposed, could lead to even better guidance for polishing decision-making.

Acknowledgment

The Lawrence Livermore National Laboratory portion of this work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The authors also acknowledge the generous support from National Science Foundation under Grant No. IIS-1849085 and Texas A&M Office of President's X-grant Program.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The datasets generated and supporting the findings of this article are obtainable from the corresponding author upon reasonable request.

References

- [1] Clery, D., 2021, "Laser-Powered Fusion Effort Nears 'Ignition'," *Science (New York, NY)*, **373**(6557), p. 841.
- [2] Abu-Shawareb, H., Acree, R., Adams, P., Adams, J., Addis, B., Aden, R., Adrian, P., Afeyan, B., Aggleton, M., and Aghaian, L., 2022, "Lawson Criterion for Ignition Exceeded in an Inertial Fusion Experiment," *Phys. Rev. Lett.*, **129**(7), p. 075001.
- [3] Lawrence Livermore National Laboratory, 2022, "National Ignition Facility Achieves Fusion Ignition". <https://www.llnl.gov/news/national-ignition-facility-achieves-fusion-ignition>. LLNL News Release on December 14, 2022.
- [4] Kritcher, A., Young, C., Robey, H., Weber, C., Zylstra, A., Hurricane, O., Callahan, D., et al., 2022, "Design of Inertial Fusion Implosions Reaching the Burning Plasma Regime," *Nat. Phys.*, **18**(3), pp. 251–258.
- [5] Biener, J., Wittstock, A., Zepeda-Ruiz, L., Biener, M., Zielasek, V., Kramer, D., Viswanath, R., Weissmüller, J., Bäumer, M., and Hamza, A., 2009, "Surface-Chemistry-Driven Actuation in Nanoporous Gold," *Nat. Mater.*, **8**(1), pp. 47–51.
- [6] ISO 4287, 1997, Geometrical Product Specifications (GPS), Surface Texture: Profile Method—Terms, Definitions and Surface Texture Parameters. International Organization for Standardization, Geneva, Switzerland.
- [7] Jin, S., Iquebal, A., Bukkapatnam, S., Gaynor, A., and Ding, Y., 2020, "A Gaussian Process Model-Guided Surface Polishing Process in Additive Manufacturing," *ASME J. Manuf. Sci. Eng.*, **142**(1), p. 011003.
- [8] Jin, S., Tuo, R., Tiwari, A., Bukkapatnam, S., Aracne-Ruddle, C., Lighty, A., Hamza, H., and Ding, Y., 2022, "Hypothesis Tests With Functional Data for Surface Quality Change Detection in Surface Finishing Processes," *IISE Trans.*
- [9] Hocheng, H., and Huang, Y. L., 2003, "A Comprehensive Review of Endpoint Detection in Chemical Mechanical Planarization for Deep-Submicron Integrated Circuits Manufacturing," *Int. J. Materials Product Technol.*, **18**(4-6), pp. 469–486.
- [10] Chris, C. Y., and Sandhu, G. S., 1993, "Chemical Mechanical Planarization (CMP) of a Semiconductor Wafer Using Acoustical Waves for In-Situ End Point Detection, Aug. 31. US Patent 5,240,552.
- [11] Yu, C. C., 1994, "Acoustical Method and System for Detecting and Controlling Chemical-Mechanical Polishing (CMP) Depths Into Layers of Conductors, Semiconductors, and Dielectric Materials," *J. Acoust. Soc. Am.*, **95**(5), pp. 2791–2791.
- [12] Salugsugan, I., 1994, "Audio End Point Detector for Chemical-Mechanical Polishing and Method Thereof," *J. Acoust. Soc. America*, **95**(5), pp. 2792–2792.
- [13] Fukuroda, A., Nakamura, K., and Arimoto, Y., 1995, "In Situ CMP Monitoring Technique for Multi-layer Interconnection," Proceedings of International Electron Devices Meeting, Washington, DC, Dec. 10–13, IEEE, pp. 469–472.
- [14] Fayolle, M., Sicurani, E., and Morand, Y., 1997, "W CMP Process Integration: Consumables Evaluation—Electrical Results and End-Point Detection," *Microelectron. Eng.*, **37–38**, pp. 347–353.
- [15] Tang, J., Dornfeld, D., Pangrle, S. K., and Dangca, A., 1998, "In-Process Detection of Microscratching During CMP Using Acoustic Emission Sensing Technology," *J. Electron. Mater.*, **27**(10), pp. 1099–1103.
- [16] Springer, G., 1999, "Dependence of Wafer Carrier Motor Current and Polish Pad Surface Temperature Signal on CMP Consumable Conditions and Ti/TiN Liner Deposition Parameters for Tungsten CMP Endpoint Detection," Proceedings of the Fourth International Chemical-Mechanical Planarization for ULSI Multilevel Interconnection Conference, Santa Clara, CA, Feb. 11–12, pp. 45–51.
- [17] Kojima, T., Miyajima, M., Akaboshi, F., Yogo, T., Ishimoto, S., and Okuda, A., 2000, "Application of CMP Process Monitor to Cu Polishing," *IEEE Trans. Semiconductor Manuf.*, **13**(3), pp. 293–299.
- [18] Zantye, P. B., Kumar, A., and Sikder, A., 2004, "Chemical Mechanical Planarization for Microelectronics Applications," *Materials Sci. Eng.: R: Reports*, **45**(3-6), pp. 89–220.
- [19] Rao, P. K., Bhushan, M. B., Bukkapatnam, S. T., Kong, Z., Byalal, S., Beyca, O. F., Fields, A., and Komanduri, R., 2013, "Process-Machine Interaction (PMI) Modeling and Monitoring of Chemical Mechanical Planarization (CMP) Process Using Wireless Vibration Sensors," *IEEE Trans. Semiconductor Manuf.*, **27**(1), pp. 1–15.
- [20] Rao, P. K., 2013, "Sensor-Based Monitoring and Inspection of Surface Morphology in Ultraprecision Manufacturing Processes," Ph.D. Dissertation, Oklahoma State University, Stillwater, OK.
- [21] Botcha, B., Rajagopal, V., Babu, R. N., and Bukkapatnam, S. T., 2018, "Process-Machine Interactions and a Multi-sensor Fusion Approach to Predict Surface Roughness in Cylindrical Plunge Grinding Process," *Procedia Manuf.*, **26**, pp. 700–711.
- [22] Hetherington, D. L., Stein, D. J., Lauffer, J. P., Wyckoff, E. E., and Shingledecker, D. M., 1999, "Analysis of In-Situ Vibration Monitoring for End-Point Detection of CMP Planarization Processes," Proceedings of the In-Line Characterization, Yield Reliability, and Failure Analyses in Microelectronic Manufacturing, Edinburgh, Scotland, May 19–21, pp. 89–101.
- [23] Bukkapatnam, S. T., Rao, P. K., and Komanduri, R., 2008, "Experimental Dynamics Characterization and Monitoring of MRR in Oxide Chemical Mechanical Planarization (CMP) Process," *Int. J. Mach. Tools. Manuf.*, **48**(12–13), pp. 1375–1386.
- [24] Kong, Z., Beyca, O., Bukkapatnam, S. T., and Komanduri, R., 2011, "Nonlinear Sequential Bayesian Analysis-Based Decision Making for End-Point Detection of Chemical Mechanical Planarization (CMP) Processes," *IEEE Trans. Semiconductor Manuf.*, **24**(4), pp. 523–532.
- [25] García, P. E., López, P. N., and Beamud Gonzalez, E., 2018, "Multi-sensor Data Fusion for Real-Time Surface Quality Control in Automated Machining Systems," *Sensors*, **18**(12), p. 4381.
- [26] Segreto, T., Karam, S., and Teti, R., 2017, "Signal Processing and Pattern Recognition for Surface Roughness Assessment in Multiple Sensor Monitoring of Robot-Assisted Polishing," *Int. J. Adv. Manuf. Technol.*, **90**(1), pp. 1023–1033.
- [27] García, P. E., and López, P. N., 2018, "Application of the Wavelet Packet Transform to Vibration Signals for Surface Roughness Monitoring in CNC Turning Operations," *Mech. Syst. Signal. Process.*, **98**, pp. 902–919.
- [28] Mallat, S. G., 1989, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. Pattern. Anal. Mach. Intell.*, **11**(7), pp. 674–693.
- [29] Cohen, A., Daubechies, I., and Feauveau, J. C., 1992, "Biorthogonal Bases of Compactly Supported Wavelets," *Commun. Pure Appl. Math.*, **45**(5), pp. 485–560.
- [30] Jolliffe, I. T., 2002, *Principal Component Analysis*, 2nd ed., Springer, New York.
- [31] Cardot, H., Ferraty, F., and Sarda, P., 1999, "Functional Linear Model," *Statist. Probabil. Lett.*, **45**(1), pp. 11–22.
- [32] Ferraty, F., and Vieu, P., 2004, "Nonparametric Models for Functional Data, With Application in Regression, Time Series Prediction and Curve Discrimination," *Nonparametric Statist.*, **16**(1-2), pp. 111–125.
- [33] Cardot, H., and Sarda, P., 2006, "Linear Regression Models for Functional Data," *The Art of Semiparametrics*, S. Sperlich, W. Härdle, and G. Aydinli, eds., Springer, Heidelberg, Germany, pp. 49–66.
- [34] Cai, T. T., and Hall, P., 2006, "Prediction in Functional Linear Regression," *Ann. Statist.*, **34**(5), pp. 2159–2179.
- [35] Berzak, H., Divol, L., Weber, C., Le Pape, S., Meezan, N., Ross, J., Tommasini, R., Khan, S., Ho, D., and Biener, J., 2018, "Increasing Stagnation Pressure and Thermonuclear Performance of Inertial Confinement Fusion Capsules by the Introduction of a High-Z Dopant," *Phys. Plasmas*, **25**(8), p. 080706.