Learning Complexity to Guide Light-Induced Self-Organized Nanopatterns

Eduardo Brandao, Anthony Nakhoul, Stefan Duffner, R Emonet, Florence Garrelie, Amaury Habrard, François Jacquenet, Florent Pigeon, Marc Sebban, Jean-Philippe Colombier

To cite this version:


HAL Id: ujm-04157829
https://ujm.hal.science/ujm-04157829
Submitted on 14 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Learning complexity to guide light-induced self-organized nanopatterns

Eduardo Brandao,1 Anthony Nakhoul,1 Stefan Duffner,2 Rémi Emonet,1 Florence Garrelie,1 Amaury Habrard,1,3 François Jacquenet,1 Florent Pigeon,1 Marc Sebben,1 and Jean-Philippe Colombier1

1Univ Lyon, UJM-Saint-Etienne, CNRS, IOGS, Laboratoire Hubert Curien UMR5516, F-42023 St-Etienne, France
2Univ Lyon, CNRS, INSA-Lyon, LIRIS, UMR5305, France
3Institut Universitaire de France (IUF), Paris, France

(Dated: April 30, 2023)

Ultrafast laser irradiation can induce spontaneous self-organization of surfaces into dissipative structures with nanoscale reliefs. These surface patterns emerge from symmetry-breaking dynamical processes that occur in Rayleigh-Bénard-like instabilities. In this study, we demonstrate that the coexistence and competition between surface patterns of different symmetries in two dimensions can be numerically unraveled using the stochastic generalized Swift-Hohenberg model. We originally propose a deep convolutional network to identify and learn the dominant modes that stabilize for a given bifurcation and quadratic model coefficients. The model is scale-invariant and has been calibrated on microscopy measurements using a physics-guided machine learning strategy. Our approach enables the identification of experimental irradiation conditions for a desired self-organization pattern. It can be applied generally to predict structure formation in situations where the underlying physics can be approximately described by a self-organization process and data is sparse and non-time series. Our work paves the way for supervised local manipulation of matter using timely-controlled optical fields in laser manufacturing.

The emergence of instabilities and symmetry breaking leading to the formation of coherent structures, is one of the most fascinating aspects of the complex dynamics governing light-surface interaction [1–3]. When a randomly rough surface is subjected to ultrafast laser pulses, it enters a far-from-equilibrium state due to the repeated absorption of pulsed optical fields. As a result, the surface exhibits spontaneous spatial organization, which is oriented by energy gradients generated by laser polarization, giving rise to laser-induced periodic surface structures (LIPSS) [4]. These structures form under far-from-equilibrium conditions and can be triggered by capillary waves, convection rolls, and thermoconvective instabilities, [5,6] which persist through dissipative structures [9]. Eliminating the prevailing laser polarization effects reveals puzzling patterns emerging from a sequence of instabilities, inducing different types of complex patterns, ranging from chaos to six-fold symmetries [10]. The photoexcited matter undergoes a transition from a disordered state to a more coherent one, referred to as a strange attractor in the phase space of nonlinear dynamics. This transition results in a metastable state, defining a self-organization structuring regime. Through this self-organization process, the material surface can be sculpted seamlessly, enabling nanoscale manufacturing [11]. Understanding the selection mechanisms involved in this morphogenesis to gain control over the uniformity, symmetry, and size of the resulting surface patterns is a major research theme in laser processing for photonics metasurfaces, biomimetics, or catalysis functionalization. [12,13]. To apply statistical inference approaches to complex systems and achieve generalizability, advanced physics-guided machine learning strategies are essential. Upon laser irradiation, a hazy boundary separates self-organized and organized surface patterns. When a material is exposed to sufficiently intense laser irradiation, it tends to organize along the stationary electromagnetic fields due to scattered/excited waves [4, 14] and self-organize in response to the random fluctuations of light absorption with a symmetry breaking with respect to polarization [15,16]. Light-oriented and self-assembled dynamical processes are inherently superimposed, and surface topographies evolve spatio-temporally towards equilibrium patterns that result from a complex competition between free energy dissipation imposing entropy production and spontaneous ordering. Consequently, any preexisting or transient organization can be disrupted by random perturbations, which can be amplified by positive feedback to lead the system towards new patterns. Ultrafast laser texturing has recently been used to obtain deep sub-wavelength periodic patterns, which raises questions about the relevant electromagnetic processes that drive the formation of these patterns well below the diffraction limit [17,18]. Various types of 2D surface patterning have been reported, including patterns with oriented, triangular, hexagonal, labyrinthine, or chaotic symmetries [19–22], featuring both positive and negative reliefs such as humps, bumps, peaks, and spikes [10]. To explain the remarkably uniform establishment of these patterns on the microscale independently from the oriented near-field optical effects on the random local nanotopography, a more global and collective perspective is required [10]. Nanoscale fluid flows were shown to be driven by a complex interplay between electromagnetic, internal and surface pressure forces which can become trapped due to the resolidification process [8, 21]. The deterministic approach to predict the underlying optical coupling processes is limited because it requires the ar-
Artificial integration of fluctuating conditions induced by surface roughness. Transiently formed structures can become unstable under nonlinear amplification and bifurcate into more complex patterns that are not accurately described by classical approaches like Navier-Stokes combined with Maxwell equations. Nonetheless, the complex pattern landscape has been experimentally explored and can be now compared with mathematical models dedicated to nonlinear system dynamics.

The Kuramoto-Sivashinsky approach has become a paradigm for describing pattern formation and spatiotemporal chaos on surfaces eroded by ion bombardment, which ultimately reproduces ripple formation and other organized patterns [22]. A similar approach was initially proposed for laser-induced nanopatterns, although a clear physical picture has yet to be established [21]. Along similar lines, the Swift-Hohenberg (SH) dynamics has been identified as a relevant candidate for representing the observed complexity of convective instabilities with spatiotemporal features, such as chaos, rolls, and hexagons [20, 25]. The SH approach has proven to be useful in identifying generic spatiotemporal dynamics of patterns in convective fluids [27, 28], as well as curvature- and stress-induced pattern-formation transition [29]. The SH approach was formally deduced from the Navier-Stokes equations in the Boussinesq approximation, with thermal fluctuation effects in a fluid near the Rayleigh-Bénard instability [30].

Along similar lines, the Swift-Hohenberg (SH) dynamical model that is variational in time and conservative in energy at fixed $\Delta t$ [36–40], but require abundant time-series data.

The purpose of this letter is to demonstrate that laser-induced pattern formation at the nanoscale can be efficiently characterized and predicted by a stochastic SH model that is variational in time and conservative in space. Our original strategy relies on the use of machine learning (ML) integrating partial physical information in the form of the SH model, which allows us to identify dominating stable modes for a set of parameters independently of initial roughness conditions. Incorporating data and prior knowledge is naturally expressed in terms of Bayesian inference, for which well-established domain-specific methods exist dating back to Laplace [31], but which cannot be applied in our experimental situation of few data and partial physical knowledge: in geophysics and climate science, where the physical process is well-understood, methods focus on state reconstruction, known as data assimilation [32]; in physics, since states can be prepared, model calibration was developed [33], with recent advancements using ML [34] to integrate the parameters of either the full model or a correction to incomplete physical knowledge from data [35]. However, solving the joint inverse problem of finding both state and model parameters is more challenging. In the climate sciences, sophisticated machine-learning techniques were recently proposed, integrating physical information via constraints, either during training or in model architecture itself [36, 10], but require abundant time-series data. Our original strategy allows us to solve the dual inverse problem using only one observed state — a scanning electron microscope (SEM) image — even with little data. Furthermore, our modelling is scale-invariant and can be applied to any laser process. By reducing experimental irradiation parameters to simple model coefficients, they can be optimized and extrapolated for surface pattern engineering.

![Figure 1](https://via.placeholder.com/150)

**Figure 1.** (a) Schematic illustration of experimental self-organization regimes induced by bursts of ultrafast laser (150 fs) double pulses. (b) Self-organized patterns of topography that develop varying time delays for a given $F$ and $N$ (AFM-3D mode). (c) Nanopattern variation with respect to laser fluence at fixed $\Delta t$ and $N$ (AFM-3D mode). (d) Nanostructure growth by feedback at different number of pulses (AFM-2D mode), for a fixed $\Delta t$ and $F$. The scale bars represent a length of 500 nm.

Tailoring nanotopographic features on a surface is a challenging task that has been successfully accomplished using ultrafast laser processes with time-controlled polarization strategies. Numerous regimes of LIPSS have been reported with various periodicities, heights, orientations, and symmetries depending on different polarization directions between the first $\vec{E}_1$ and second pulse $\vec{E}_2$, characterized by $\sin \alpha = (\vec{E}_1 \cdot \vec{E}_2)/(||\vec{E}_1|| \cdot ||\vec{E}_2||)$ in Fig.1(a) [10, 21, 41, 42]. Figs.1(b-d) present surface topographies measured by high resolution atomic force microscopy (AFM). A circular region with a diameter of 1 $\mu$m corresponding to the laser impact center was mapped in 3D (tilted) mode in Fig.1(b-c) and in 2D for Fig.1(d). To observe the significant role of temporal pulse splitting $\Delta t$ in nanopatterns control, laser peak fluence $F$ and $N$ were kept fixed at 0.18 J/cm$^2$ and 25 respectively, as shown in Fig.1(b). At $\Delta t = 8$ ps, organized nanopeak structures were observed with a high aspect
ratio, a height of \( \sim 100 \text{ nm} \) and a diameter of \( \sim 20 \text{ nm} \) \cite{11}. An extension of 2 ps in \( \Delta t \) modifies the observed patterns that turn into a different organization, a regime referred to as nanobumps \cite{10}. For \( \Delta t = 15 \text{ ps} \), a regime of nanohump generation is reached with a lower aspect ratio as the structures display a height of \( \approx 10 \text{ nm} \) and a diameter of \( \approx 30 \text{ nm} \).

The role of laser fluence is revealed by fixing \( \Delta t = 25 \text{ ps} \) and \( N = 25 \), as depicted in Figure \( \text{I(a)} \). At \( F = 0.18 \text{ J/cm}^2 \), a low-contrast nanopeak regime is formed, evolving into a nano-stripe pattern with a slight increase in laser fluence increase to 0.20 \text{ J/cm}^2. At \( F = 0.22 \text{ J/cm}^2 \), a transition region is established, combining both stripes and cavities. Finally, at \( F = 0.24 \text{ J/cm}^2 \), the surface is uniformly organized with hexagonally arranged nanocavities having a depth of \( \approx 25 \text{ nm} \) and a diameter of \( \approx 30 \text{ nm} \). Both nanohumps and nanovoids result from hydrothermal flows guided by surface tension and rarefaction forces, leading to thermoconvective instability at the nanoscale, similarly to well-known Rayleigh-Bénard-Marangoni instabilities \cite{8,10,21,43,54}. Laser dose also plays a role, as positive feedback regulates pulse-to-pulse topographical transformations. As shown in Figure \( \text{I(c)} \), pulse-to-pulse growth dynamics presented in Fig.1(a), three different surface organizations were observed. Pulse-to-pulse growth dynamics exhibits the transitions from convection cells \( (N = 15) \), to the creation of crests on the convection cells \( (N = 20) \). The nanopeaks grow on the edges of the crests to reach their optimal shape, concentration and organization at \( N = 20 \).

The adimensional form of the generalized Swift-Hohenberg equation (SH) used in this letter is (see derivation in Suppl. Mat.):

\[
\dot{u} = \epsilon \ddot{u} - (1 + \tilde{\nabla}^2)^2 \tilde{u} + \gamma \tilde{u}^2 - \tilde{u}^3. \tag{1}
\]

The SH model was introduced in \cite{30} as a model of Rayleigh-Bénard convection, modified by the inclusion of a \( u^2 \) nonlinearity allowing for small amplitude destabilization and the emergence of experimentally observed hexagonal patterns. With appropriate boundary conditions, the original SH equation exhibits a type-I's instability that is isotropic, invariant with respect to translations and to \( u \rightarrow -u \) \cite{28}. Perturbations of \( u_0 \) are selectively amplified depending on the norm of the wave number, leading to the formation of complex patterns with no preferential direction. The generalization of SH model has the Lyapunov functional \( \mathcal{L}[\tilde{u}] = \int_\Omega \frac{1}{2} (\nabla^2 \tilde{u} + 2 \nabla^2 \tilde{u} + \tilde{u}) + \frac{1}{4} \tilde{u}^4 - \frac{3}{32} \tilde{u}^3 - \frac{1}{2} \tilde{u}^2 dx \) and \( \tilde{u} = -\frac{\mathcal{L}}{\mathcal{M}} \), as can be readily verified. During the SH dynamics, the Lyapunov functional \( \mathcal{L} \) decreases in the same way as entropy decreases during the formation of physical patterns, and it converges asymptotically to a stable value \cite{29} (see Fig. 2). We numerically solve the SH equation using a second-order Strang splitting pseudo-spectral solver with an adaptive time step \cite{55,59}, offering a good compromise between accuracy and speed. Fig. 2a-c and Fig. 2d-f show evolution dynamics of pattern formation for two pairs of \( \epsilon, \gamma \).

A ML model is employed to learn the relationship between observed laser parameters \( \theta \) and patterns, using only few, non-time series data (I), assuming an approximately SH process, not explicitly given in terms of \( \theta \), parameterized by \( \varphi \) (consisting of a scale factor \( l \), the maximum wavenumber in a domain of side 224 pixels given as multiple of 2\( \pi \), the adimensional model parameters \( \epsilon \) and \( \gamma \), and \( \tilde{t} \), which can be seen as a stabilization time) (II), with unknown initial conditions \( u_0 \) (III). We motivate this choice by symmetry considerations (Suppl. Mat.) as well as the similarity between SEM images and SH solutions (Fig 3). Combining experimental information with that obtained via the ML model, we find that the timescale of the convective instability is consistent with that reported in \cite{10} (Suppl. Mat.), further supporting our choice. Learning the relationship between laser parameters and patterns consists in solving the dual inverse problem of estimating an unknown initial state and model parameters with severe constraints, which is
Figure 3. Each row shows a 224 by 224 pixel SEM experimental image, with 1 µm ≈ 237 pixels which was never seen by the ML model during learning, the corresponding ML-predicted image for the same laser parameters, and three nearest neighbors (NN) of the former among solver generated images; and a bar plot of $l^2$ distance in the image of the feature mapping $F_m$ between SEM image, ML-predicted image (ml) and NN (nn1, nn2, nn3). Image labels, left to right: $F, \Delta t, N$ (SEM); predicted SH parameters $l_p, \epsilon_p, \gamma_p, \tilde{t}_p$ (other). Bar plots: ML predictions are more accurate, as distance between SEM and ML predictions is smaller than to NN, the former integrating global information. On the first and second rows, NN with different length scales can be observed, suggesting concurrent multi-scale SH processes. The ML model, which integrates single-scale SH knowledge, can only predict one of these processes.

Figure 4. Each plot shows, as a heatmap, the ML model prediction of a single SH parameter (bottom to top: scale factor $l_p$; order parameter $\epsilon_p$; symmetry breaking parameter $\gamma_p$; simulation stabilization time $\tilde{t}_p$) as a function of laser fluence, time delay, and number of pulses (respectively, x-axis and y-axis, and column). Experimental points are overlaid on each plot.

A feature mapping $F_m$ is therefore defined that is simplifying (non injective) and discriminating (if $u_i, u_j$ have different patterns then $F_m(u_i) \neq F_m(u_j)$) such that the image of the data distribution under $F_m$ is conditionally independent of $u_0$ given the physical knowledge $\varphi$. This considerably simplifies the problem since the initial state no longer needs to be estimated. Learning $F_m$ from few data is impractical [61], (II) precludes deriving it on first principles, and using traditional image features would limit discriminating power for unknown patterns. $F_m$ is therefore chosen as a deep convolutional neural net-work (CNN) pretrained for a broad classification task on Imagenet [63], since CNNs are translation equivariant (making them suited for a pattern specification task). Their features are learned automatically from data, and retain scale information [64]. Given experimental data $\{\theta_i, u_i\}_{i=1...N}$, we learn $\tilde{\varphi}_\alpha$ that maximizes the log likelihood of the observed $F_m(u_i)$:

$$\hat{\alpha} = \arg \max_{\alpha} \sum_{i=1}^{N} \log p(F_m(u_i)|\tilde{\varphi}_\alpha(\theta_i))$$ (2)

Assuming that the distribution of $\varphi$ given $\theta$ is peaky, we label experimental $u^i$ with $\tilde{\varphi}^i$ the SH parameters of its nearest neighbor (NN), in the image of $F_m$, among a large number of $u$ pre-generated with the SH solver from random $u_0$. By integrating physical knowledge in this way, the problem of maximizing the likelihood above can be replaced with a lower bound. Explicitly, assuming data is generated i.i.d. from a Gaussian distribution,

$$\hat{\alpha} = \arg \min_{\alpha} \frac{1}{N} \sum_{i=1}^{N} ||\tilde{\varphi}^i - \tilde{\varphi}_\alpha(\theta^i)||^2$$ (3)

which is a low-dimensional problem that can be solved with few data [65] with a support vector regressor [66] $\tilde{\varphi}_\alpha$ parameterized by $\alpha$. 

a challenging task and cannot be tackled in general using only ML methods. However, for a self-organization process, stating initial conditions exhaustively is wasteful, since for random perturbations of the uniformly zero solution of SH most Fourier modes are attenuated. A feature mapping $F_m$ is therefore defined that is simplifying (non injective) and discriminating (if $u_i, u_j$ have different patterns then $F_m(u_i) \neq F_m(u_j)$) such that the image of the data distribution under $F_m$ is conditionally independent of $u_0$ given the physical knowledge $\varphi$. This considerably simplifies the problem since the initial state no longer needs to be estimated. Learning $F_m$ from few data is impractical [61], (II) precludes deriving it on first principles, and using traditional image features would limit discriminating power for unknown patterns. $F_m$ is therefore chosen as a deep convolutional neural net-work (CNN) pretrained for a broad classification task on Imagenet [63], since CNNs are translation equivariant (making them suited for a pattern specification task). Their features are learned automatically from data, and retain scale information [64]. Given experimental data $\{\theta_i, u_i\}_{i=1...N}$, we learn $\tilde{\varphi}_\alpha$ that maximizes the log likelihood of the observed $F_m(u_i)$:

$$\hat{\alpha} = \arg \max_{\alpha} \sum_{i=1}^{N} \log p(F_m(u_i)|\tilde{\varphi}_\alpha(\theta_i))$$ (2)

Assuming that the distribution of $\varphi$ given $\theta$ is peaky, we label experimental $u^i$ with $\tilde{\varphi}^i$ the SH parameters of its nearest neighbor (NN), in the image of $F_m$, among a large number of $u$ pre-generated with the SH solver from random $u_0$. By integrating physical knowledge in this way, the problem of maximizing the likelihood above can be replaced with a lower bound. Explicitly, assuming data is generated i.i.d. from a Gaussian distribution,

$$\hat{\alpha} = \arg \min_{\alpha} \frac{1}{N} \sum_{i=1}^{N} ||\tilde{\varphi}^i - \tilde{\varphi}_\alpha(\theta^i)||^2$$ (3)

which is a low-dimensional problem that can be solved with few data [65] with a support vector regressor [66] $\tilde{\varphi}_\alpha$ parameterized by $\alpha$. 

a challenging task and cannot be tackled in general using only ML methods. However, for a self-organization process, stating initial conditions exhaustively is wasteful, since for random perturbations of the uniformly zero solution of SH most Fourier modes are attenuated. A feature mapping $F_m$ is therefore defined that is simplifying (non injective) and discriminating (if $u_i, u_j$ have different patterns then $F_m(u_i) \neq F_m(u_j)$) such that the image of the data distribution under $F_m$ is conditionally independent of $u_0$ given the physical knowledge $\varphi$. This considerably simplifies the problem since the initial state no longer needs to be estimated. Learning $F_m$ from few data is impractical [61], (II) precludes deriving it on first principles, and using traditional image features would limit discriminating power for unknown patterns. $F_m$ is therefore chosen as a deep convolutional neural net-
Figure 3 demonstrates the remarkable accuracy of our ML strategy in predicting the shape and scale of experimental patterns, even for never-before-seen laser parameters. Our strategy is more efficient than local methods that rely on nearest neighbor information, since the distance to the SEM experimental patterns in the image of the feature mapping $F_m$ is smaller. As shown in Fig. 3, the complexity of the learned relationship between laser parameters and SH parameters grows with the number of experimental observations, with sharp boundaries of rapidly varying parameter values in regions of many data. The ML model can extrapolate to regions with few data for $N$ and $\Delta t$, but less so for $F$, which would require higher experimental resolution. Importantly, we find that predicted SH parameters are correlated, and the correlation sign changes with $N$ (Suppl. Mat.): $l_p$ for example, is inversely correlated with pattern characteristic size and increases with $N$ (the increase is not uniform, being greater for large $l_p$ regions). This parameter is particularly important as the characteristic size of a stable mode is of great interest for applications. Because $l_p$ and other parameters are correlated, it cannot be set freely, but as seen in Fig. 4, parameter isosurfaces are orthogonal at places: at e.g. $N=25$, a high-gradient transition regime for $l_p$, at $F = 0.18 \text{ J/cm}^2$, $\Delta t = 15$ ps in the $\Delta t$ direction is observed, while for $F, \Delta t$ in the same region, the other SH parameters remain roughly constant. Varying $l_p$ in the direction of $\nabla l_p$ thus allows adjusting the characteristic size of the particular stable mode defined by the other SH parameters. This opens the door to pattern optimization for specific applications.

Interestingly, $(\gamma_p)$ determines whether holes or bumps are observed; $F = 0.2 \text{ J/cm}^2$ separates a region of high and low $\gamma_p$, roughly independently of $N$; the sign of $\gamma_p$ appears to be determined by $F$ and $\Delta t$ only. Furthermore, for large bifurcation parameter $(\epsilon_p)$, many modes are non-attenuated and patterns are less ordered. Correspondingly, large $\epsilon_p$ patches are observed at high $F/\Delta t$ (highest energy coupling). For $N = 25$, superimposing the $\gamma_p = 0$ isosurface on the $\epsilon_p$ prediction, it can be seen that it is roughly perpendicular to isosurfaces of $\epsilon_p$. These abrupt transition regimes of $\epsilon_p$ are consistent with experimental observations where two patterns of different order are superimposed on the same SEM image. $(\tilde{l}_p)$ for constant $l_p, \epsilon_p, \gamma_p$, symmetry increases with $\tilde{l}_p$, as symmetrical patterns require large $\tilde{l}_p$ to stabilize from a uniformly random state. As can be seen in the bottom row, $\tilde{l}_p$ tends to increase with $N$, consistently with the physical view that a large $N$ increases the time the dissipative system is in a far from equilibrium state. This increase is not uniform across $F, \Delta t$ pairs, and the area of laser parameter space of relatively large $\tilde{l}_p$ decreases with $N$.

We show that ultrafast laser-irradiated surface nanoscale patterns can be numerically modelled by a scale-invariant generalized Swift-Hohenberg equation. A machine learning model is trained to learn the connection between the stochastic SH equation and laser parameters, independently of initial conditions, using a deep convolutional network to extract features and by incorporating physical information. Our original strategy can generally be applied to accurately predict the shape and scale of physical patterns generated by other self-organization processes, even if the underlying physical model is only approximate and experimental data is limited and non-time series. The ML model is able to identify regions of laser parameters that are relevant for applications and can even be used to predict novel patterns, since the convolutional neural network features are not learned from observed patterns. Regions where pattern superpositions are observed could be modeled more accurately via a mixing of SH processes, as a manifestation of superposed states of self-organization, providing new routes toward nanoscale surface manipulation by light.

This work has been funded by a public grant from the French National Research Agency (ANR) under the "France 2030" investment plan, which has the reference EUR MANUTECH SLEIGHT -ANR-17-EURE-0026.

**Appendix on experimental set-up.** In the proposed experiment, Mach-Zehnder interferometry was used to combine the effect of polarization mismatch with an adjustable inter-pulse delay $\Delta t$, enabling fine control of surface topography at the tens of nanometer scale [10]. By breaking the surface isotropy imposed by a single polarization state, a wide range of self-organization regimes was achieved on a nickel monocrystal oriented in the (001) direction. Specifically, using a cross-polarization strategy, setting a depolarization angle of $\alpha = 90^\circ$ and a range of time delays between 8 and 25 ps were set, as shown in Fig. 1(a). The pulse duration was fixed at 150 fs, and the laser dose was finely controlled by the number $N$ of double-pulse sequences. Prior to laser irradiation, the Nickel surface was mechanically polished with a $Ra < 5 \text{ nm}$ to ensure that the surface dynamics followed a hydrodynamics-governed process, smoothing the inhomogeneous electromagnetic response.

**Appendix on Image similarity.** It is important to note, regarding the ML strategy, that the problem is approached within the image of a feature mapping $F_m$, where the concept of similarity differs from visual similarity. In this feature space, an image is equivalent to any of its images in the orbit of the group of symmetries of $F_m$, such as translations, but also a variety of other symmetries that are learned from data automatically. Intuitively, similarity in feature space corresponds to similarity of patterns, which can be described in terms of e.g. "bumpiness," "roundness," etc.


