High-fidelity physics-based simulation of 3D materials and natural phenomena has become a critical component in motion pictures, visual effects (VFX), animation, and video games over the past few decades. Beyond computer graphics, physics-based simulation is also essential in many computational science domains such as structural engineering and vehicle/aircraft design. As modern architecture progressed towards exascale systems supporting exaFLOPS and accessible hundred-teraFLOP consumer level workstations, numerical simulations have further found an increasing breadth of new applications such as real-time VFX previews, virtual reality games, interactive surgical training, predictive soft robotics, and computational fabrication. While theoretical computation capacity is now less of an impediment, a timely opportunity emerges for innovations in designing new numerical algorithms that mathematically resolve complex geometry and multi-physics with high accuracy, and can best utilize new computational platforms with plausible scalability.

I believe that the ultimate goal of physics-based simulation is to model and recreate the entire physical reality in a virtual world to the extent that humans cannot distinguish between the two. With the central theme being physics-based modeling and simulation, we perform research on numerical optimization, numerical methods for PDEs, continuum mechanics, high performance computing, and machine learning.

My research activity has a pronounced focus on drastically enhancing (1) the scale (in terms of both detail and resolution), (2) scope (including supported materials/phenomena and geometrical flexibility), (3) reliability (accuracy, stability, robustness, and the ability of predicting the real world), and (4) efficiency (with exploration of modern high performance computational architecture) of simulations leveraging computational modeling of elastoplastic, viscoelastic, and other complex materials in multi-physics simulations. Since the start of my academic career, much of my research has been centered around these aspects of the hybrid Lagrangian/Eulerian Material Point Methods (MPM) [1, 2] (on which I have published the greatest amount of research accomplishments in the computer graphics community). I also do research on various other Lagrangian and Eulerian schemes for numerically solving physical Partial Differential Equations.

After joining the University of Pennsylvania as an Assistant Professor (June 2017-present) I have further published more than 30 peer-reviewed papers at top venues, including 18 articles in ACM SIGGRAPH / Transaction on Graphics (the most prestigious computer graphics venue, in which I published 3 articles before becoming a professor). More broadly speaking, I am passionate about and have actively done research in (1) computer graphics: high-fidelity physics-based modeling and simulation [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35], (2) computational science and engineering: scientific computing for computational physics, geo-mechanics, and medical/bio-engineering [36, 37, 38, 39, 40, 41, 42, 43, 44, 45], and (3) artificial intelligence: numerical methods thriving together with AI, ML, vision, and robotics [46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56]. Although targeted at different research communities or application areas, I consider these research directions heavily overlapping with each other and mutually facilitating each other, and all of them are centralized on doing useful scientific computations.

The rest of this statement details my selected research and describes some ongoing and future research directions. I split my discussion into topics roughly focusing on 4 different (but highly related) aspects:

- (§1) Numerical discretization scheme innovations for simulations;
- (§2) Accommodating versatile physical behaviors in simulations;
- (§3) Enabling high-performance, massively parallel and scalable simulations;
- (§4) Scientific computation striving together with AI, Machine Learning, Vision, and Robotics.

A link to another historical document is provided here¹ to summarize my earlier research during when I was a PhD student (2010-2015, UCLA CS) and a postdoc (2015-2017, UCLA Math+CS). Those work shaped the direction of my career and built solid foundations for many exciting investigations after them.

1 Numerical Discretization Scheme Innovations

Over the past few years, I have demonstrated consistent efforts in identifying and resolving various computational challenges as well as innovating algorithmic designs centered around the Material Point Method (MPM) [57]. MPM, as a hybrid spatial discretization method, benefits from the advantages of both Lagrangian and Eulerian views. It uses Lagrangian particles to carry material states including mass, position, velocity, volume, strain, and stress, etc. The grid acts as an Eulerian scratch pad for computing stress divergence and performing spatial-temporal numerical integration. Grid nodes represent the actual degrees of freedom, which store mass, position and velocity. In each time step, particles transfer mass and velocities to the grid. Grid velocities are updated using

¹https://www.seas.upenn.edu/~cffjiang/research_statement_phd_and_postdoc.pdf
forces computed through particle quadratures. The velocities and strain increments are then interpolated back to the particles, followed by particle advection.

1.1 Adaptivity.

While MPM is well suited for capturing topology-changing shapes, the need for adaptive discretization is particularly pronounced for objects with internal geometric features and heterogeneous characters. For example, in the case of materials with intense anisotropy (such as muscles, fiber-reinforced concretes, plant leaves), sticking with a uniform MPM background grid is no longer the best choice.

Collaborating with colleagues, we proposed a spatially adaptive variant of the Generalized Interpolation Material Point (AGIMP) method [23] to alleviate these limitations. By only refining regions of particular interest, the AGIMP scheme can resolve fine-grained self-contact and object collision features with significantly reduced computational cost. The framework was built with all operations implemented with the use of efficient, uniform grid operations as building blocks using SPGrid [58] (Figure 2). On the other hand, a crucial need for temporal adaptivity in MPM is also emerging due to the potentially substantial variation of material stiffness and velocities in multimaterial scenes. To this end we worked on a temporally adaptive symplectic Euler scheme for MPM with regional time stepping (RTS): asynchronous MPM integration (AsyncMPM) [20], where different time steps are used in different regions (Figure 1). A time stepping scheduler operates at the granularity of small blocks to maintain a natural consistency with the hybrid particle/grid nature of MPM.

1.2 Spatial discretization.

In a recent paper [21] we discovered that the formulation of my earlier works the Affine Particle-In-Cell (APIC) [29] and Polynomial Particle-In-Cell (PolyPIC) [24] can be expressed in a way that is consistent with a Galerkin-style weak form discretization of the governing equations. We used Moving Least Squares to replace the shape functions in the stress divergence term, leading to a new force computation scheme that does not require evaluating the gradients of nodal shape functions. Compared to traditional MPM, the resulting Moving Least Squares Material Point Method (MLS-MPM) provides almost identical visual results, enables high order spatial discretizations, and results in an effortless $2 \times$ speed up with easier implementation. MLS-MPM immediately became a new MPM standard in academic codebases and industrial softwares [59]. It also enables a simple approach for two-way coupling between granular media and rigid bodies. In computational mechanics, we further developed an approach $\textsc{MPM-SDEM}$ (Figure 3) that couples MPM and the spheropolygon discrete element method (SDEM) through the exchange of contact force information. It combines the advantage of MPM for accurately simulating elastoplastic continuum materials and the high efficiency of DEM for calculating the Newtonian dynamics of discrete near-rigid objects.
1.3 Time stepping.

MPM is a hybrid Lagrangian/Eulerian scheme. Accordingly, we actively conduct synergistic projects in advancing both pure Eulerian and pure Lagrangian simulation techniques. A first example is the recent BiMocq$^2$ [14] (n levels of Bi-directional mapping of convective quantities), an unconditionally stable, pure Eulerian-based advection scheme to efficiently preserve the advection accuracy of all physical quantities for long-term fluid simulations (Figure 4). Built upon the method of characteristic mapping, BiMocq$^2$ outperforms almost all state-of-art Eulerian and hybrid schemes at the cost of only two semi-Lagrangian advects. The resulting pipeline leads to a new standard in conservative fluid simulation with high order accuracy.

On the Lagrangian side, we worked with colleagues from Adobe Research in the invention of DOT [13], the Decomposed Optimization Time Integrator, a new domain-decomposed optimization method for solving the per time step, nonlinear problems of implicit numerical time integration. DOT is especially suitable for large time step simulations of deformable bodies with nonlinear materials and high-speed dynamics (Figure 5). It is efficient, automated, and robust at large, fixed-size time steps, thus ensuring stable, continued progress of high-quality simulation output.

Going a step beyond DOT, we realized that the idea of optimization-based time integrator could enable a wide range of new strategies for designing smarter and faster schemes. Correspondingly we invented HOT [9], the Hierarchical Optimization Time Integration for efficient implicit time-stepping of MPM irrespective of simulated materials and conditions. As an implicit MPM timestepper accelerated by a custom-designed Galerkin multigrid wrapped in a quasi-Newton solver, HOT provides convergent simulations out of the box across widely varying materials and computational resolutions without parameter tuning (Figure 6). It also outperforms existing state-of-the-art, heavily optimized implicit MPM codes with an up to 10× performance speedup.

1.4 Frictional contact.

Contacts weave through every aspect of our physical world, from daily household chores to acts of nature. Modeling and predictive computation of these phenomena for solid mechanics is important to every discipline concerned with the motion of mechanical systems, including engineering and animation. Nevertheless, efficiently time-stepping accurate and consistent simulations of real-world contacting elastica remains an outstanding computational challenge. To model the complex interaction of deforming solids in contact, in collaboration with Adobe and NYU, we invented Incremental Potential Contact (IPC) [5] – a new model and algorithm for variationally solving implicitly time-stepped nonlinear elastodynamics based on novel innovations in barrier energy based constrained optimization. IPC maintains an intersection- and inversion-free trajectory regardless of material parameters, time step sizes, impact velocities, severity of deformation, or boundary conditions enforced (Figure 7). The resulting time stepper solves contact problems that are intersection free (and thus robust), inversion-free, efficient (at speeds comparable to or faster than available methods that lack both convergence and feasibility), and accurate (solved to user-specified accuracies). To our knowledge this is the first implicit time-stepping method, across both the engineering and graphics literature that can consistently enforce these guarantees as we vary simulation parameters.

With IPC, numbers of extensions and applications beyond computer graphics await. Due to its unprecedented robustness and accuracy, a significant amount of challenging frictional contact tasks involving materials ranging from macroscale biomechanical tissues to mesoscale granular media can now be numerically simulated with almost no more difficulty than just inserting real physical parameters. Along this direction, I’m working on multiple projects in robotics and mechanical engineering to fully explore IPC’s power, in addition to several numerical modeling projects that focus on further improving its efficiency and flexibility.

1.5 Continuing research and directions on geometrical flexibility.

Leveraging experiences in numerical algorithmic development mentioned above, I observe a major challenge with regard to algorithmic foundations for computational solids: the lack of efficient and robust treatments for frictional contacts and topology changes in existing solvers for intricate geometric shapes largely impairs the result. These behaviors typically require highly domain-specific treatments or impractical modifications to existing algorithms (e.g., detecting and enforcing frictional contact of fibers in woven fabrics). In order to tackle problems involving difficulty in (re)meshing, constraints in the function space discretization of codimensional...
structures, and laboriousness in capturing contact and strong coupling, a primary focus within the next few years will be modeling topologically changing submanifold geometry with meshless representations. Optimal simulation of thin structures such as shells and rods depends on discretization schemes yielding convergence to their continuum foundations under refinement. Traditional mesh-based representations require high regularity of the discretization function space and impose great difficulty on the treatment of collisions and split-and-merge behaviors. I look forward to constructing fully meshless representations of the codimensional structures through a novel updated Lagrangian view of tracking the deformation. Exploiting my extensive experience with treating continuous contact and strong discontinuity, proper constitutive models will be developed to accurately capture the governing physics of the geometrically flexible particle models. More details of this research direction has been proposed in my awarded NSF CAREER grant.

2 Accommodating Versatile Physical Behaviors

Multi-physics simulation involves multiple physical fields and dominant processes described by different Differential Equations with extra constitutive relations incorporated to model their interaction. Finite Element and Finite Difference based software such as ANSYS [60] and COMSOL [61] provide existing solutions to many multi-physics problems such as thermal stress analysis, chemical interactions, and acousto-mechanical design. However, these solutions suffer from the aforementioned disadvantages in treating thin structures, adapting to unified representations, and fitting into modern computational platforms. Two more challenging aspects are (1) disagreement of underlying continuous versus discrete representations, and (2) modeling intricacy in the strong coupling of multiple materials and governing physics. More concretely, the questions we seek to answer are (1) how to discretely capture complex constitutive relations; (2) how to effectively discretize complex governing equations; (3) how to design more efficient underlying discretization schemes that adapt to special material cases.

In response to these challenges and in pursuing the objective of delivering a unified system that not only handles robust simulation of versatile geometries, but also accurately captures common multi-material multi-physics behaviors, my group focuses on several significant existing and new threads of investigation.

2.1 Fracture and damage mechanics.

Continuum fracture can be modeled non-locally through tracking of damage variables evolving over time, known as phase-field fracture (PFF) [62, 63]. Non-local damage mechanics predicts the crack tip and has the benefit of capturing all modes of fracture while producing mesh invariant crack propagation. To this end, my group has chosen to pair the non-local PFF with MPM for how naturally a scalar field evolution technique aligns with the grid-based computations of MPM.

---

Figure 5: DOT [13] offers stable, efficient, and robust Lagrangian elastodynamics simulation with high accuracy.

Figure 6: HOT [9] enables unified, consistent, and predictive simulations of metals with real-world mechanical parameters.
The resulting algorithm was named Continuum Damage Material Point Method \textbf{(CD-MPM)} \cite{11} (Figure 8) in our recent publication. CD-MPM builds the necessary foundations for addressing necessary steps for extending the continuous description to include flexible geometries and anisotropic damage rules. Our motivation is that in daily objects (especially food), modeling those with underlying structures that dictate preferred fracture directions is notoriously difficult. Thus we further developed \textbf{AnisoMPM} \cite{4}: a robust and general approach for animating the dynamic fracture of isotropic, transversely isotropic, and orthotropic materials. AnisoMPM has three core components: a technique for anisotropic damage evolution, methods for anisotropic elastic response, and a coupling approach. For anisotropic damage, we adopt a non-local continuum damage mechanics (CDM) geometric approach to crack modeling and augment this with structural tensors to encode material anisotropy. Furthermore, we discretize our damage evolution with explicit and implicit integration, giving a high degree of computational efficiency and flexibility. We also utilize a QR-decomposition based anisotropic constitutive model that is inversion safe, more efficient than SVD models, easy to implement, robust to extreme deformations, and that of anisotropy.

Our elasto-damage coupling is enforced through an additive decomposition of our hyperelasticity into a tensile and compressive component in which damage is used to degrade the tensile contribution to allow for material separation. For extremely stiff fibered materials, we further introduce a novel Galerkin weak form discretization that enables embedded directional inextensibility. We present this as a hard-constrained grid velocity solve that poses an alternative to our anisotropic elasticity that is locking-free and can model very stiff materials. AnisoMPM clearly stands as a robust and general approach to animating dynamic fracture of not only isotropic materials, but also the significantly more complex transversely isotropic and orthotropic materials. Furthermore, AnisoMPM offers both the efficient and easy to implement explicit damage and the stable and robust implicit damage, giving flexibility based on the desired fracture behavior and simulation constraints. Finally, AnisoMPM also poses numerous exciting directions for future study, ranging from an intuitive interface for artistic control to the exploration of other types of fracture effects.

2.2 Multi-phase and poromechanical mixtures.

Using particle quadrature for approximating level set represented cut-cell information, we developed a novel scheme for simulating two-way coupled interactions between nonlinear elastic solids and incompressible fluids \textbf{IQ-MPM} in \cite{6}. It leads to a stable and efficient method handling large time steps under the CFL limit while using a single monolithic solve for the coupled pressure fields, even in the case with highly nonlinear elastic solids. On
the other hand, phase-field as a continuum state also began as a popular approach for simulating multi-phase effects like the formation of ice crystals [64], and the mixing of fluids [65]. Similarly, [66] achieved fast multiphase fluid simulation using the Cahn-Hilliard equation [67]. Most recently, [68] attained complex multi-material simulation, expanding on this work by introducing a phase proportion variable evolved over time with the Allen-Cahn equation [69], which is quite similar in formulation to the continuum damage evolution governing equation used in [11]. On the other hand, my group has investigated poromechanical multi-species mixtures (described with continuous mixture theory) with a MPM discretization for Locally Averaged Navier-Stokes equations (MPM-LANS) [22] (Figure 10). Combining the phase-field model with the mixture theory shows great promise for enabling a relatively complete model of multi-material solid-fluid systems in both miscible and immiscible cases. A noteworthy application and experimental validation of cracking multi-phase systems is predictive avalanche using a novel Cohesive Cam Clay model\(^3\) (Figure 13) [40, 36, 39], ice sheet and glacier calving with coupled ocean dynamics, through my on-going collaboration with snow scientist Johan Gaume from EPFL.

### 2.3 Chemical-mechanical and thermal-mechanical systems.

MPM is also naturally suited for simulating continuous chemical/mechanical coupling, as demonstrated in the recent work of [70] for simulating lithiation and delithiation of a silicon battery anode. A major form of transport for chemical concentration can be modeled with diffusion in the concentration flux. As future research endeavors we will investigate proper modeling of arbitrary chemical potentials to complement the diffusion process. Thermal-mechanical systems range from heat-induced thermal-visco-elastoplasticity, thermal expansion, and frictional heating. Continuum theories based on the Gibbs free energy or the Helmholtz free energy as the thermodynamic potential enable the formulation of nonlinear thermal-mechanically coupled PDEs [71]. We will develop effective numerical formulations for the discretization of such systems with MPM. My preliminary research on this includes temperature dependent elasticity for MPM Thermomechanics [18] (Figure 15). In [3], we studied Non-Fourier Diffusion for diffusive phenomena such as hydrogel swelling, limited diffusive domain for smoke flow, and snowflake and dendrite formation.

Extending our inviscid euler solver discretized on a sphere [17], we further developed a chemomechanical simulation framework rooted in lubrication theory to enable the simulation of Soap Film Dynamics [7] on spherical bubbles both in free flow as well as under body forces such as gravity or external air flow (Figure 11). This flow is characterized by a complex coupling between surfactant concentration and Marangoni surface tension.

A particular challenge arises since thermo-mechanical systems often involve physically or chemically phase-changing materials (e.g., melting snow during frictional heating, burning polymers into gas). We will address algorithms that incorporate the effects of phase-field and concentration fields in the variational models and derivatives of the free energy. For small-scale fluid systems, we are also exploring Continuum Surface Stress (CSS)-based approaches [72] in modeling thermo-capillary Bénard-Marangoni convection in multi-phase fluids, which also finds many applications in my collaborative efforts on exascale additive metal manufacturing simulation [73] with the US Depart of Energy (DOE).

---

\(^3\)This publication is among the 50 (out of 5000+) most read Nature Communication articles on earth and planetary sciences (2018).
2.4 Algorithmic guarantees on efficient strong coupling, convergence, and stability.

We recently demonstrated in ADMM-MPM [12] (Figure 12) that variational formulations in elasticity, inelasticity, and damage mechanics can enable optimization-based time integrators that are efficient, unconditionally stable, and convergence-guaranteeing. Continuum formulations can be made to contain similar variational forms such that advances from nonlinear optimization can be adopted. To achieve strong coupling between various physical quantities, monolithic solvers based on the Newton-Raphson method are no longer an efficient option. It is thus reasonable to perform variable splitting for different physical variables, and explore the usage of the Alternating Direction Method of Multipliers (ADMM) [74, 12] for decoupling the monolithic nonlinear system into smaller and simpler sub-problems.

3 High-Performance Scalable Simulations

I strongly believe in the value of augmenting algorithms for use with high-performance platforms, especially multi-GPU architecture, for scalable simulations. Large-scale simulations are becoming more demanding and popular than ever. There are many efforts invested in algorithms that scale with supercomputers or cloud servers. The biggest advantage of HPC over cloud computing is that supercomputer nodes are usually more tightly coupled through faster and more stable connections (e.g. infiniband). For physics-based simulations, the frequent communications among its subdomains heavily relies on these interconnections. On the other hand, there has been a trend towards the deployment of general-purpose GPU (GPGPU) as an accelerator in modern high-end HPC systems [75]. Current HPC frameworks either leave most responsibilities to the programmer [76] or impose many rigid restrictions for specific usage scenarios. Although plenty of parallel physics-based simulation libraries and softwares already exist [77, 78, 79], none provide a multi-GPU-tailored MPM solution. Due to the unique hybrid particle-grid nature of MPM, most existing dynamic load balancing strategies [80, 81, 82] cannot be readily imported into MPM without significant algorithmic and software-engineering challenges. As such, most open-sourced MPM code does not adapt to modern heterogeneous HPC environments. Moreover, no existing MPM frameworks support the flexible geometry or versatile multi-physics we focus on. In next few years, leveraging prior experience in CPU/GPU
optimization of MPM and access to computing resources in national labs, my group is well prepared to design and benchmark innovative multi-node multi-GPU tailored large-scale simulation algorithms.

My group targets at developing new multi-GPU infrastructure to achieve high performance and scalability. The first step is to start with building intra-node parallel multi-GPU MPM. My work on single GPU MPM (Figure 15) based on hierarchical reduction provides a solid foundation for extension. This pipeline already approaches optimal efficiency, yet the requirement of small time steps hinders its use in time-constrained applications. Besides exploring novel algorithms, making good use of multi-GPU architecture requires a dedicated domain decomposition (DD) and workload distribution scheme. The task in each timestep is to distribute particles across discrete GPUs, perform time integration on the grid, and update particle states. Many commercial VFX softwares adopt a static partitioning strategy in their DD scheme, which is simple but far from optimal. On the other hand, the benefits of dynamic partitioning and scheduling can be offset by unexpected scenarios. In intra-node parallel computations, a non-uniform memory access (NUMA) system with centralized scheduling is assumed. It is interesting to take advantage of this infrastructure and conduct experiments on various scheduling algorithms.

We recently created a massively parallel and scalable framework called Multi-GPU MPM [8] which has enabled the simulation of hundreds of millions of particles on a single workstation computer with 4-8 GPUs. Such a groundbreaking result enables the possibility for large-scale simulation-based engineering tasks such as aircraft design, computer-aided machine manufacturing and human-robot interaction to achieve astonishing numerical accuracy (and thus real-world predictivity) within a time duration that is many orders of magnitudes shorter than any existing approaches (Figure 14).

Additionally, my group also has an emphasis on designing inter-node parallel aspects, including investigations on resilience, latency hiding, and HPC tailored programming models. Since 2019 I am involved in the The Exascale Computing Project (ECP), working with the US Department of Energy (DOE) on developing a new class of high-performance computing systems whose power will be measured in exaflops. ECP is a multi-billion dollar project by the US government for developing the next-generation super computers and using them to target the most challenging scientific and national security mission needs. We target at running simulations on the Summit supercomputer for solving important challenges in scientific discovery, energy assurance, economic competitiveness, and security, as part of the Exascale Additive Manufacturing Project (ExaAM). This involvement helps improve quality, reliability, and application breadth of additive manufacturing and helps accelerate innovation in clean energy manufacturing institutes.

4 Artificial Intelligence, Machine Learning, Vision, and Robotics

4.1 Improving Deep Learning with Numerical Optimization

Training a Neural Network is a highly nonconvex and nonlinear optimization task, as is the simulation of high-resolution elastodynamics under extreme loads, big time steps, or the presence of complex frictional contacts. Both tasks apparently show great potential of significantly benefiting from the development of advanced techniques in numerical optimization. For simulations of solids with large deformation, we have demonstrated significant advantages through the development of Newton-Barrier methods and Quasi-Newton methods augmented with Domain Decomposition [13] and Galerkin Multigrid [9]. It is thus intriguing to explore aspects from

4https://www.exascaleproject.org/
such methods in the training of Neural Networks. A second order stochastic Newton-Inexact-CG NN training scheme [46] efficiently augmented with complex-step derivative approximation [83] presents our first step in this direction. In deep learning, first-order methods prevail and remain the dominant solution for modern deep architectures. However, we tackle the practical obstacle of Hessian calculation by using the complex-step finite difference (CSFD) – a numerical procedure adding an imaginary perturbation to the function for derivative computation. For example, to approximate gradients, CSFD uses $f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + hi) - f(x_0)}{ih} + O(h^2) \approx im(f(x_0 + hi))$. CSFD is highly robust, efficient, and accurate (as accurate as the analytic result, no floating point errors commonly seen in finiteifference). This method allows us to literally apply any known second order optimization methods for deep learning training. Based on it, we design an effective Newton Krylov procedure. The key mechanism is to terminate the stochastic Krylov iteration as soon as a disturbing direction is found so that unnecessary computation can be avoided. During the optimization, we monitor the approximation error in the Taylor expansion to adjust the step size. This strategy combines advantages of line search and trust region methods making our method preserves good local and global convergence at the same time. We have tested our methods in various deep learning tasks. The experiments show that our method outperforms exiting methods, and it often converges one-order faster. We believe our method will inspire a wide-range of new algorithms for deep learning and numerical optimization.

**Next steps:** Combining with the barrier formulation for handling inequality constraints [5], our next step is investigating the high order training of Neural Networks under geometrical and physical constraints of the input, the weights, and the network structures.

In the mean while, I am also interested in exploring next-generation hardware opportunities for machine learning. Collaborating with colleagues from UPenn and UCSD who are experts on photonics hardware, we have started a big NSF project to perform investigations that will leverage the state-of-the-art integrated photonics technology to develop an innovative programmable photonic computation accelerators (PPCA) for large scale machine learning, accelerating the computation speed and reducing the cost and energy consumption to sustain long term performance requirements in ML. I expect that developing the innovative programmable photonic computation accelerators could be applied in the domains which demand extreme speed, energy efficiency, parallelism, significant complexity, and high scalability on an ultra-compact footprint, and full programmability.

### 4.2 Simulations for Intelligent Systems.

Simulations can be used to extract features, provide guidance, and synthesize data for many computer vision, scene understanding, and robotics tasks. **Smart robots.** Towards this end, we developed physics-based simulation that is applied to emulate various robot actions and mirrors the actions that are functionally equivalent to the human’s in the sense of causing the same state changes by exerting similar forces [49]. Through this approach, a robot reasons about which forces to exert and what goals to achieve to generate actions (i.e., mirroring), rather than strictly mimicking demonstration (i.e., overimitation) a bioluminescence cell-based assay that reflects Snail1 protein stability. We demonstrated this approach by teaching a real Baxter robot with a complex manipulation task involving haptic feedback—opening medicine bottles. In another project [48], we use fluids simulations to train robots on pouring tasks. We autonomously pour from unknown symmetric containers found in a typical wet laboratory for the development of a robot-assisted, rapid experiment preparation system. The robot estimates the pouring container symmetric geometry, then leverages simulated pours as priors for a given fluid to pour precisely and quickly in a single attempt. This system is implemented on the Rethink Robotics Sawyer and KUKA LBR iiwa manipulators.

**Scene understanding.** We also perform simulations or rendering for Computer vision tasks, such as in photorealistic indoor scene synthesis [50, 51] and physical scene understanding [55]. In such applications, simulations help generate high quality training data and invisible physics-based features in machine learning algorithms for vision tasks.

### 4.3 Inverse Problems through Combining ML with Differentiable Simulation/Computing

Revisiting the study of Material Point Methods over years, and the recent development of our Incremental Potential Contact (IPC) [5] method for Lagrangian mechanics frictional contact simulated with FEM, we realize

---

that both methods present an extremely appealing property: differentiability. Combining such simulations with automatic differentiation techniques (AutoDiff) and discrete adjoint methods for sensitivity analysis, we could use these high-fidelity simulations for many modern inverse problems in computational design, computation fabrication, AI, and robotics.

Conducting research using differentiable simulation and differentiable scientific computing becomes an extremely promising direction. Based on such ideas we recently started to work on inverse problems through differentiable elasticity simulations. Our first study along this direction is differentiable topology sensitivity. In [38], a multi-resolution hybrid Lagrangian-Eulerian topology optimization method is proposed where the elastic force equilibrium is solved by MPM. The method transfers density information from freely movable Lagrangian carrier particles to a fixed set of Eulerian quadrature points. The transfer is based on a smooth radial kernel involved in the compliance objective to avoid the artificial checkerboard pattern. The quadrature points act as MPM particles embedded in a lower-resolution grid and enable a sub-cell multi-density resolution of intricate structures with a reduced computational cost. A quadrature-level connectivity graph based method is adopted to avoid the artificial QR patterns commonly existing in multi-resolution topology optimization method. Numerical experiments are provided to demonstrate the efficacy of the proposed approach.

Immediately following the optimization of compliance mechanism, we performed another research on soft robotics, specifically soft fabricated aerial vehicles [47]. Unmanned aerial vehicles have been demonstrated successfully in a variety of tasks, including surveying and sampling tasks over large areas. These vehicles can take many forms. Quadrotors’ agility and ability to hover makes them well suited for navigating potentially tight spaces, while fixed wing aircraft are capable of efficient flight over long distances. Hybrid aerial vehicles (HAVs) attempt to achieve both of these benefits by exhibiting multiple modes; however, morphing HAVs typically require extra actuators which add mass, reducing both agility and efficiency. In this project we developed a bistable morphing hybrid aerial vehicle with folding wings that exhibits both a quadrotor and a fixed wing mode without requiring any extra actuation. This is achieved by leveraging the motion of a bistable mechanism at the center of the aircraft to drive folding of the wing using only the existing motors and the inertia of the system. Our prototype successfully transitions between both modes and our experiments demonstrate that the behavior of the fabricated prototype is consistent with that of the simulation. The next step is to combine the differentiable dynamic simulation (through discrete adjoint sensitivity analysis) with deep Neural Networks for (1) designing controllers of the vehicle, and (2) better models for automatic design space explorations.

**More thoughts:** In the mean time, we are also exploring combining physics-aware Neural Networks with differentiable simulators for efficient design of numerical schemes. Examples include designing new Galerkin kernels / function spaces for spatial-temporal discretization with differentiable frictional contact and fracture. We can also utilize differentiable numerical linear algebra and using NNs to optimize operators such as those in multigrid methods for particular types of linear and nonlinear physical problems. Furthermore, Neural networks provide a way of representing functions with smooth surrogates. We are thus interested in exploring automatic NN representations of dissipative physics (friction, plasticity, viscoplasticity, fracture). For example, if we could find a NN to locally approximate elastoplasticity with hyperelasticity (trained over different deformation regions), then we would be able to develop efficient variational integrators for simulating plastic materials.

Finally, combining NNs and differentiable simulations with computer vision opens an opportunity of creating physically functional digital doubles from video captured objects. Starting from video footage, we could combine NN constitutive models, differentiable forward simulations, and CNN-based visual learning networks, to discover new material models and even new terms in differential equations from data.
References Cited


