Mechanochemical regeneration of NaBH4

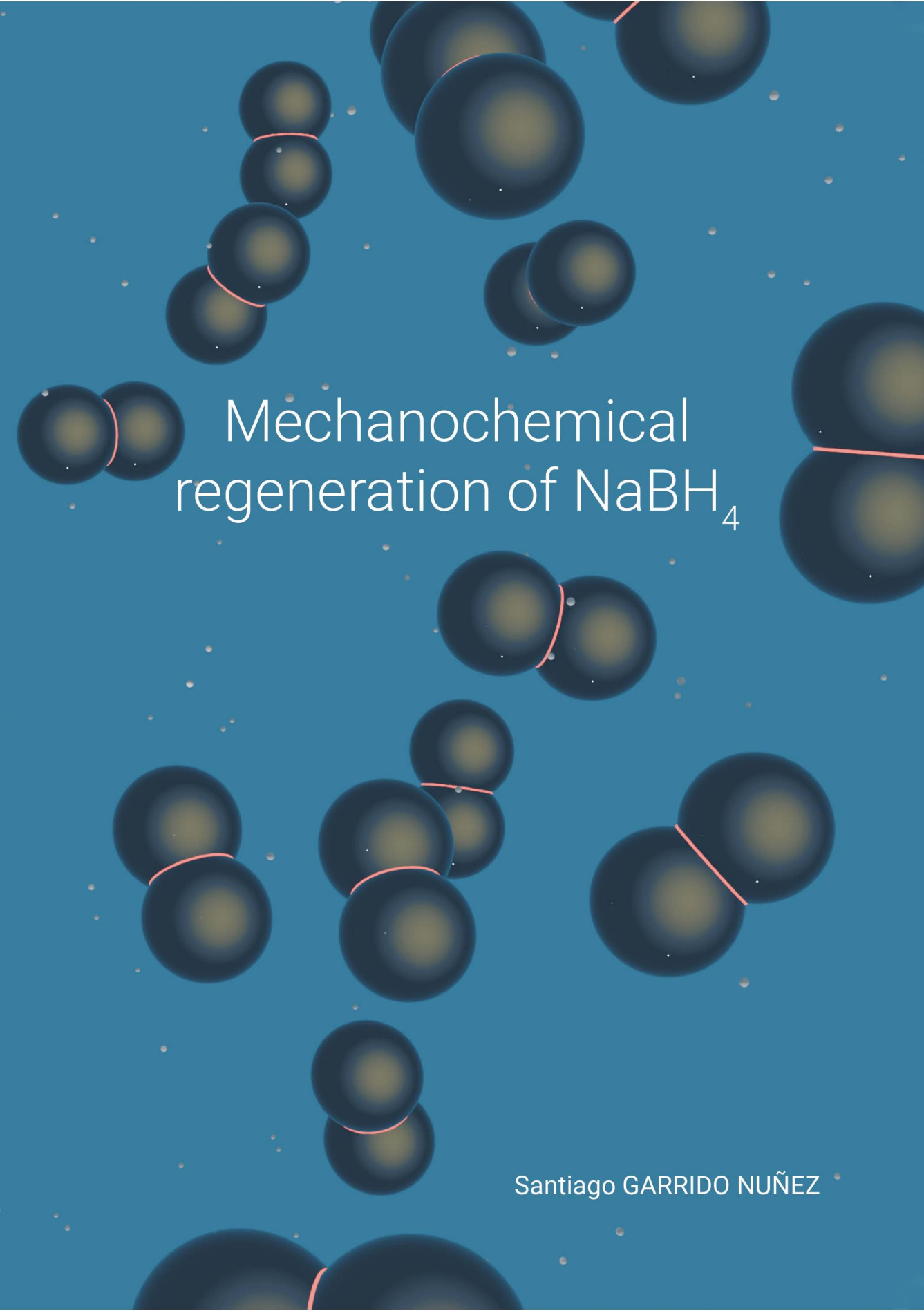
Garrido Nuñez, S.

DOI
10.4233/uuid:75ffbd05-1789-4df7-ac8a-cd0260a04e44

Publication date
2026

Document Version
Final published version

Citation (APA)
Garrido Nuñez, S. (2026). *Mechanochemical regeneration of NaBH4*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:75ffbd05-1789-4df7-ac8a-cd0260a04e44



Mechanochemical regeneration of $NaBH_4$

Santiago GARRIDO NUÑEZ

Mechanochemical regeneration of $NaBH_4$

Mechanochemical regeneration of $NaBH_4$

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, Prof. dr. ir. H. Bijl,
chair of the Board for Doctorates
to be defended publicly on
Monday the 2nd of February 2026 at 12:30

by

Santiago GARRIDO NUÑEZ

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. ir. J.T. Padding,	Delft University of Technology, *promotor*
Prof. dr. ir. D.L Schott,	Delft University of Technology, *promotor*

Independent members:

Prof. dr. ir. W. de Jong,	Delft University of Technology
Prof. dr. F.M. Mulder,	Delft University of Technology
Prof. dr. C. Weidenthaler,	Max Planck Institute for Coal Research, Germany
Dr. I. Vollmer,	Utrecht University
Ir. K. Visser,	Delft University of Technology

Reserve member:

Prof. dr. ir. C. Poelma,	Delft University of Technology

This work was supported by the project SH2IPDRIVE: Sustainable Hydrogen Integrated Propulsion Drives, funded by the RVO under grant MOB21013.

Keywords: Mechanochemistry, Sodium Borohydride, Discrete Element Method, Process Optimization, Machine Learning, Surrogate Model

Printed by: proefschriftenprinten.nl

ISBN 978-94-6518-229-2

An electronic copy of this dissertation is available at
https://repository.tudelft.nl/.

La novela perfecta rechazaría al lector.
(The perfect novel would reject the reader.)

Carlos Fuentes

Contents

Summary

Sodium borohydride ($NaBH_4$) combines an exceptionally high gravimetric hydrogen capacity (≈ 10.9 wt %) with ambient condition stability, making it an appealing solid-state hydrogen carrier for heavy-duty and maritime applications. Its practical adoption, however, is limited by the energy-intensive and costly regeneration of the spent product, sodium metaborate ($NaBO_2$). Mechanochemical routes, where high-energy ball milling drives solid-state synthesis, offer a solvent-free, room-temperature alternative, but published data remain fragmented, difficult to reproduce, and largely confined to laboratory scale. To tackle these bottlenecks, this dissertation develops a robust regeneration pathway for $NaBH_4$ and, in parallel, advances a generally applicable methodology for scalable, data-rich mechanochemistry.

The work begins by identifying the divergence in $NaBH_4$ mechanochemical yields reported for similar reactant systems. These differences are traced to un-reported or poorly controlled milling parameters such as rotational speed, ball-to-powder ratio, jar geometry, and the working principle of the milling machine. A comprehensive fractional design of experiments (DoE) is therefore established to quantify both individual and interaction effects of four key operational variables on chemical yield: milling time, molar ratio, ball-to-powder ratio (BPR) and rotational speed. Leveraging these insights, the highest literature yield (90 %) is reproduced while operating the mill 20 % more slowly, cutting specific energy demand and metal wear. The same procedure simultaneously delivers a "ready-to-use" aqueous $NaBH_4$ solution, eliminating hazardous separation steps.

To move beyond empirical screening, the thesis introduces a physics-based description of the milling process. Discrete Element Method (DEM) simulations reveal that normal and tangential energy dissipation per collision, together with specific collision frequency, constitute a set of mechanical invariants that uniquely characterize ball-mill operation. When experimental conditions are reproduced in terms of these dimensionless groups, different datasets collapse onto universal master curves, providing a transferable "mechanical fingerprint" that links equipment of different sizes and designs to comparable chemical performance.

Building on this mechanistic insight, the influence of the ratio between tangential and normal stressing events is systematically explored under constant-power and constant-rotational-speed conditions. A low fill ratio of 6 %, which maximizes tangential dissipation, increases the specific yield to 0.28 yield W^{-1} and achieves a record conversion to 94 %, while shifting the fill ratio to 17 % drives the system towards normal impacts and reduces the yield by 40-50 percentage points. The resulting guidelines enable targeted adjustment of dynamics and filling conditions to favor shear over less productive compression, provid-

ing a practical route to maximize regeneration efficiency while minimizing energy demand.

Data-driven models extend the framework. A two-stage Gaussian-Process-Regression ensemble, trained on combined chemical parameters and DEM-based mechanical descriptors, predicts out-of-sample $NaBH_4$ yields with $R^2 = 0.83$, allowing unpromising parameter combinations to be discarded computationally before a single gram of powder is milled. Complementarily, a graph neural-network surrogate offers a novel alternative to traditional DEM simulations by predicting particle trajectories with a mean-squared error of 2×10^{-4} m^2 with a time-step more than 100 times larger than DEM. Additionally, the model can dynamically predict the energy dissipation of the system, showcasing how these surrogates can be used beyond a simple kinematic fit.

Together, these elements form a closed feedback loop: statistically designed experiments feed physics-based simulations, then these simulations yield scale-independent descriptors. These descriptors drive machine learning models, significantly cutting time and energy demands. Lastly, a proof-of-concept approach with the surrogate offers a low-entry-barrier tool for the mechanochemistry community, intending to standardize reporting protocols when the knowledge of physical simulations is limited. The dissertation thus delivers a validated, energy-efficient route for mechanochemical $NaBH_4$ regeneration, and a transferable template for reproducible, scalable mechanochemistry. By combining empirical, mechanistic and data-centric approaches, the study advances $NaBH_4$ towards a truly circular hydrogen carrier and lays the groundwork for broader adoption of green mechanochemical synthesis across the chemical industry.

Samenvatting

Natriumboorhydride ($NaBH_4$) combineert een uitzonderlijk hoge gravimetrische waterstofcapaciteit (≈ 10,9 wt %) met stabiliteit onder omgevingscondities, wat het tot een aantrekkelijke vastestofwaterstofdrager maakt voor zware en maritieme toepassingen. De praktische toepassing wordt echter beperkt door de energie-intensieve en kostbare regeneratie van het uitgeputte product, natriummetaboraat ($NaBO_2$). Mechanochemische routes, waarbij hoogenergetische kogelmaling de vastestofsynthese aandrijft, bieden een oplosmiddelvrij alternatief bij kamertemperatuur, maar gepubliceerde gegevens blijven gefragmenteerd, moeilijk reproduceerbaar en grotendeels beperkt tot laboratoriumschaal. Om deze knelpunten aan te pakken ontwikkelt dit proefschrift een robuust regeneratiepad voor $NaBH_4$ en, parallel daaraan, een algemeen toepasbare methodologie voor schaalbare, data-rijke mechanochemie.

Het onderzoek begint met het identificeren van de variatie in mechanochemische opbrengsten van $NaBH_4$ die voor vergelijkbare reagentia-systemen zijn gerapporteerd. Deze verschillen worden herleid tot niet-gerapporteerde of slecht gecontroleerde maalparameters zoals rotatiesnelheid, kogel-tot-poeder-verhouding, potgeometrie en het werkingsprincipe van de maalinstallatie. Daarom wordt een uitgebreid fractioneel experimenteel ontwerp (DoE) opgesteld om zowel de individuele als interactie-effecten van vier belangrijke operationele variabelen op de chemische opbrengst te kwantificeren: maaltijd, molverhouding, kogel-tot-poeder-verhouding (BPR) en rotatiesnelheid. Met deze inzichten wordt de hoogste literatuuropbrengst (90 %) gereproduceerd terwijl de molen 20 % langzamer draait, waardoor de specifieke energiebehoefte en metaal-slijtage afnemen. Dezelfde procedure levert tegelijkertijd een “gebruiksklare” waterige $NaBH_4$ oplossing, waardoor gevaarlijke scheidingsstappen vervallen.

Om verder te gaan dan empirische screening introduceert het proefschrift een natuurkundige beschrijving van het maalproces. Discrete Element Method (DEM) simulaties tonen aan dat normale en tangentiële energiedissipatie per botsing, samen met de specifieke botsfrequentie, een set mechanische invarianten vormen die het functioneren van de kogelmolen uniek karakteriseren. Wanneer experimentele condities worden gereproduceerd in termen van deze dimensieloze groepen, vallen verschillende datasets samen op universele mastercurves, wat een overdraagbare “mechanische vingerafdruk” oplevert die apparatuur van verschillende grootte en ontwerp koppelt aan vergelijkbare chemische prestaties.

Voortbouwend op dit mechanistisch inzicht wordt de invloed van de verhouding tussen tangentiële en normale belastingsevents systematisch onderzocht onder zowel constante vermogens- als constante rotatiesnelheidscondities. Een lage vulgraad van

6 %, die de tangentiële dissipatie maximaliseert, verhoogt de specifieke opbrengst tot 0,28 yield W^{-1} en bereikt een recordomzetting van 94 %, terwijl het verhogen van de vulgraad tot 17 % het systeem naar normale impacten verschuift en de opbrengst met 40–50 procentpunten verlaagt. De resulterende richtlijnen maken een gerichte afstemming van dynamiek en vullingscondities mogelijk om schuifspanningen te bevorderen boven minder productieve compressie, wat een praktische route biedt om de regeneratie-efficiëntie te maximaliseren en tegelijk het energieverbruik te minimaliseren.

Datagedreven modellen breiden het kader uit. Een tweestaps Gaussian Process Regression ensemble, getraind op gecombineerde chemische parameters en DEM-gebaseerde mechanische descriptoren, voorspelt *out-of-sample* $NaBH_4$ opbrengsten met $R^2 = 0.83$, waardoor weinigbelovende parametercombinaties computationeel kunnen worden uitgesloten voordat er één gram poeder wordt gemalen. Aanvullend biedt een grafeneuraalnetwerksurrogaat een nieuw alternatief voor traditionele DEM simulaties door deeltjesbanen te voorspellen met een gemiddelde kwadratische fout van 2×10^{-4} m^2 bij een tijdstap die meer dan 100 keer groter is dan die van DEM. Bovendien kan het model de energiedissipatie van het systeem dynamisch voorspellen, wat laat zien hoe deze surrogaten verder kunnen gaan dan een eenvoudige kinematische fit.

Samen vormen deze elementen een gesloten feedbacklus: statistisch ontworpen experimenten voeden natuurkundige simulaties, die vervolgens schaalonafhankelijke descriptoren opleveren. Deze descriptoren sturen machine learning modellen aan, waardoor tijd en energiebehoefte aanzienlijk worden verlaagd. Ten slotte biedt een proof-of-concept-benadering met het surrogaat een hulpmiddel met lage instapdrempel voor de mechanochemiegemeenschap, bedoeld om rapportageprotocollen te standaardiseren wanneer de kennis van fysische simulaties beperkt is. Het proefschrift levert daarmee een gevalideerde, energie-efficiënte route voor mechanochemische regeneratie van $NaBH_4$, en een overdraagbare blauwdruk voor reproduceerbare, schaalbare mechanochemie. Door empirische, mechanistische en data-gerichte benaderingen te combineren, brengt de studie $NaBH_4$ dichter bij een werkelijk circulaire waterstofdrager en legt zij de basis voor bredere toepassing van groene mechanochemische synthese in de chemische industrie.

Nomenclature

Latin capitals

symbol	units	description
A	m	Oscillation amplitude (jar radius)
BPR	–	Ball-to-powder mass ratio
C	–	Regularisation / kernel scale (SVR, GPR)
$\bar{E}_n$	J	Average normal energy dissipation
$\bar{E}_t$	J	Average tangential energy dissipation
$\tilde{P}_{\mathrm{spec},n}$	–	Normalised specific normal power
$\tilde{P}_{\mathrm{spec},t}$	–	Normalised specific tangential power
F_c	N	Total contact force on particle
F_n	N	Normal contact force
F_t	N	Tangential contact force
G	GPa	Shear modulus of contact material
I	$\mathrm{kg\,m^2}$	Moment of inertia (particle)
K_n	$\mathrm{N\,m^{-1}}$	Hertz normal stiffness
K_t	$\mathrm{N\,m^{-1}}$	Mindlin tangential stiffness
M	kg / –	Mass; or number of estimators (RF/XGB)
P_n	W	Normal power dissipation
P_t	W	Tangential power dissipation
P_{tot}	W	Power drawn by mill
R	m	Particle / ball radius
R^2	–	Coefficient of determination (metric)
RMSE	–	Root-mean-squared-error
MAE	–	Mean-absolute-error
MAPE	%	Mean-absolute-percentage-error
S_n	$\mathrm{N\,m^{-1}}$	Normal contact stiffness scale
S_t	$\mathrm{N\,m^{-1}}$	Tangential contact stiffness scale
T	K	Absolute temperature
T_z	N m	Driving torque about z-axis
ΔE_{norm}	–	Normalised incremental energy

Latin lower case

symbol	units	description
a	$\mathrm{m\,s^{-2}}$	Acceleration
d_b	mm	Ball diameter
$d_{\max}$	–	Maximum tree depth (RF/XGB)
f_{col}	$\mathrm{s^{-1}}$	Collision frequency
$\tilde{f}_{\mathrm{coll}}$	–	Normalised collision frequency
f_r	–	Volumetric fill ratio of balls
g	$\mathrm{m\,s^{-2}}$	Gravitational acceleration
m	kg	Particle mass
n	$\mathrm{min^{-1}}$	Jar rotational speed (rpm)
$n_{\mathrm{estimators}}$	–	Number of trees / estimators (RF/XGB)
n_{split}	–	Min. samples to split a node (RF)
n_{leaf}	–	Min. samples in a leaf (RF)
p	Pa	Pressure
r_c	m	Connectivity radius (graph)
t	s	Time variable
u	$\mathrm{m\,s^{-1}}$	Speed magnitude
v_n, v_t	$\mathrm{m\,s^{-1}}$	Relative normal / tangential velocity
x	–	Hydration degree in $\mathrm{NaBO_2 \cdot} x\mathrm{H_2O}$
w_i	–	Weight in weighted-MSE loss
ℓ	–	Kernel length-scale (GPR)

Greek upper case

symbol	units	description
Ω	–	Regularisation term in XGBoost objective
Θ	$\mathrm{m^2\,s^{-2}}$	Granular “temperature” (velocity variance)

Greek lower case

symbol	units	description
α	–	Lagrange multiplier (SVR)
β	–	Logarithmic damping factor
γ_n	$\mathrm{kg\,s^{-1}}$	Normal damping coefficient
γ_t	$\mathrm{kg\,s^{-1}}$	Tangential damping coefficient
γ	–	RBF kernel width / XGBoost pruning strength
δ_n	m	Normal overlap
δ_t	m	Tangential overlap
ϵ	–	ϵ-insensitive tube half-width (SVR)
η	–	Learning rate (XGBoost)
θ	–	Trainable network parameters (SGN)

symbol	units	description
λ	–	L_2 regularisation weight (XGBoost)
μ_f	–	Static friction coefficient (DEM)
μ_r	–	Rolling friction coefficient (DEM)
ν	–	Poisson ratio
ϕ_{SDF}	m	Signed-distance-field value
ρ	$\mathrm{kg\,m^{-3}}$	Density
σ_n^2	–	Observation-noise variance (GPR)
τ	–	History-window length (GNN)

Sub-/superscripts and decorators

symbol	units	description
*	–	Effective (reduced) quantity (e.g. Y^*)
n	–	Normal component / direction
t	–	Tangential component / direction
norm	–	Indicates a normalised variable

1

Introduction

1.1. Solid hydrogen carriers

Hydrogen is widely regarded as a key clean energy carrier for a sustainable future, but a major challenge remains in storing hydrogen safely and efficiently [1, 2]. Traditional storage of hydrogen as a compressed gas or cryogenic liquid poses safety risks and offers low volumetric energy density [3]. Solid hydrogen carriers, which chemically bind hydrogen in a solid matrix, are being actively explored to overcome these issues. By chemically binding hydrogen within a solid, these materials enable safe, compact, and high-density hydrogen storage that could meet the demands of a future "hydrogen economy," in which hydrogen is produced and used as a widespread fuel [1]. Indeed, solid-state hydrogen storage in certain hydride materials has shown the potential for greater hydrogen density and safer handling than conventional tanks [2]. However, no single material yet satisfies all practical requirements: current candidates face different trade-offs in storage capacity, release temperature, kinetics, and cost [1, 4]. This section introduces the main classes of solid hydrogen carriers, including metal hydrides, complex hydrides, and other chemical hydrides like ammonia borane, discussing their potential and the challenges that motivate continued research.

1.1.1. Types of solid hydrogen carriers

Solid hydrogen carriers are materials that store hydrogen via chemical bonds or adsorptive interactions in a solid matrix, releasing H_2 upon suitable triggers, such as heat or water. This category encompasses a range of materials. In general, they can be grouped into three families: metal hydrides, which are hydrides of elemental metals or intermetallic compounds, such as MgH_2 and $LaNi_5H_6$; complex hydrides, ionic solids in which hydrogen resides in polyatomic anions, such as borohydrides ($NaBH_4$) and alanates ($NaAlH_4$); and chemical molecular hydrides, neutral compounds that store hydrogen in covalent bonds and are solid under ambient conditions, such as ammonia borane (NH_3BH_3). All of these materials offer material-based storage as opposed to physical compression, and often feature a higher volumetric hydrogen density and inherent safety since they don't rely on high pressures [5].

Metal hydrides

Metal hydrides are compounds formed by the reaction of hydrogen gas with metals or alloys, often yielding solid materials where hydrogen atoms occupy interstitial sites in a metal lattice or form metal–hydrogen bonds [6]. Classic examples include alloys like $LaNi_5H_6$, and simple metal hydrides like magnesium hydride (MgH_2). Metal hydrides are attractive because they can reversibly absorb and release hydrogen numerous times by changing temperature or pressure, effectively acting as a rechargeable hydrogen "battery". They also tend to offer very high volumetric hydrogen densities, since the hydrogen is stored in a compact solid matrix. Moreover, metal hydrides are safe materials for hydrogen storage under mild conditions and can have higher hydrogen densities than liquid hydrogen [6]. A notable case is alane (AlH_3), which has a volumetric hydrogen density about twice that of liquid H_2 [7].

Despite these advantages, metal hydrides often suffer from excess weight

and thermal constraints. Many intermetallic hydrides contain heavy elements, so their gravimetric hydrogen capacity is modest. $LaNi_5H_6$, for example, stores only about 1.4 wt% hydrogen [2], although it operates near room temperature with excellent kinetics [8]. Lighter metal hydrides like MgH_2 and lithium hydride LiH have much higher hydrogen content: MgH_2 contains 7.6 wt% H_2, and LiH contains 12.7 wt% H_2 [9]. These light metal hydrides exemplify the gravimetric potential of metal hydrides; however, they are too thermodynamically stable, requiring high temperatures to release hydrogen. For instance, bulk MgH_2 only begins to desorb H_2 at roughly 300 ºC and above, with very slow kinetics unless improved by catalysts or nanostructuring [10, 11]. LiH is even more extreme as it is stable at nearly 700 °C, and thus is impractical for on-demand H_2 release by heating. Although LiH can release hydrogen readily if reacted with water, it is highly reactive and corrosive, posing handling challenges [9]. Another illustrative example is alane (AlH_3). It contains approximately 10 wt% hydrogen and, as mentioned, has an exceptional volumetric hydrogen density, but it is a metastable compound. Rehydriding aluminum metal to AlH_3 requires enormous hydrogen pressures of up to 7000 bar in bulk [7], making it very difficult to regenerate once decomposed. In general, while metal hydrides are often fully reversible in theory, many have unfavorable thermodynamics that demand either high temperatures or high pressures to cycle, and their weight efficiency is a concern [1]. Research continues to improve metal hydrides by alloying and catalysis, but as of now, all metal hydrides fall short in meeting the necessary requirements for a viable hydrogen economy. The key challenges can be summarized as insufficient storage capacity, slow reaction rates, and hydrogen release occurring only at impractical temperature levels [1].

Complex hydrides

Complex hydrides are a broad category of hydrogen-rich compounds containing anionic complexes such as borohydride $[BH_4]^-$ or aluminohydride $[AlH_4]^-$, often paired with lightweight metal cations (Li^+, Na^+, K^+, Mg^{2+}). These materials have attracted great interest because they typically involve only light elements (H, B, Al, Li, etc.), giving them very high gravimetric hydrogen densities which in many cases surpass those of simple metal hydrides [12]. For example, borohydrides are among the most hydrogen-rich solids known: lithium borohydride ($LiBH_4$) has a theoretical hydrogen content of about 18.5 wt% [13], and sodium borohydride ($NaBH_4$) contains 10.6 wt% H_2 [14]. Such high hydrogen capacities exceed the U.S. DOE targets of 5–7 wt% usable hydrogen, and are very appealing for lightweight storage [15]. Alanates have intermediate capacity and, when doped with Ti-based catalysts, can reversibly hold approximately 5 wt%, but are still much lighter than transition-metal hydrides. Complex hydrides also tend to be stable solids under ambient conditions, making them nonvolatile, which is good for safety and storage [12].

The main challenge with complex hydrides is achieving controlled hydrogen release and uptake. These materials usually have strong covalent bonds in their anionic groups (B – H, Al – H, etc.), which means high thermal stability and slow dehydrogenation kinetics. As an illustration, $NaBH_4$ has a decomposition enthalpy of ≈110 kJ per mole H_2

and does not release significant hydrogen until heated above about (500ºC) [14]. In fact, if one simply heats solid $NaBH_4$, it may melt and partially decompose to sodium metal and boron, but liberates much less H_2 than expected because of side reactions like phase separation and evaporation of sodium [14]. While not easily dehydrogenated by heat alone, it is well known as a chemical hydrogen generator via hydrolysis to produce H_2 at room temperature, making it useful for on-demand hydrogen via a fuel cell. The trade-off is that this hydrolysis yields sodium metaborate ($NaBO_2$) as a spent product, and regenerating $NaBH_4$ from $NaBO_2$ is notoriously difficult with conventional methods [14]. Similarly, $LiBH_4$ has a high decomposition temperature (380–400°C) [16] and releases hydrogen in multiple steps, often forming stable boron compounds that are hard to rehydrogenate. These high temperatures are undesirable for practical systems. On the positive side, some complex hydrides can be induced to release H_2 at lower temperatures through clever modifications: adding catalysts, making composite mixtures, or nanoconfinement [17, 18]. To give an example, doping sodium alanate ($NaAlH_4$) with a small amount of Ti-based catalyst enables it to reversibly release and reabsorb about 4–5 wt% hydrogen at 120–180°C, far lower than its undoped decomposition temperature [19]. Nanoscale engineering is another approach: confining hydrides in porous scaffolds can reduce particle size and alter thermodynamics, aiding hydrogen release. Despite such advances, most complex hydrides remain irreversible or only partially reversible under practical conditions. In summary, complex hydrides offer some of the highest hydrogen storage densities and the appeal of solid-state safety, but they face serious challenges in hydrogen release kinetics and especially in reversibility. Many borohydrides and high-capacity complex hydrides either decompose to inert solids that won't take up H_2 again under mild conditions, or require very high pressures and temperatures to re-hydrogenate [20].

Molecular Chemical Hydrides

Beyond the metal and complex hydrides, there are molecular chemical hydrides that store hydrogen in covalent bonds within discrete molecules or polymeric frameworks. The prime example is ammonia borane ($NH_3 . BH_3$), a solid at room temperature that has been intensely studied as a hydrogen storage material. Ammonia borane (AB) contains 19.6 wt% hydrogen, making it one of the highest gravimetric hydrogen contents of any stable compound, and a volumetric hydrogen density of about 146 g H_2 per liter of solid [21]. It is also non-toxic and stable under ambient conditions, which makes it very appealing on paper. AB releases hydrogen upon heating through a series of thermolysis reactions: around 110 °C it begins to release H_2, and with further heating up to 150 °C it can liberate roughly two equivalents of H_2 (about 13–14 wt% of its hydrogen) in an exothermic process [22, 23]. This hydrogen release is relatively low in temperature, and researchers have developed catalysts to lower the release temperature or to achieve more complete dehydrogenation [23]. However, ammonia borane exemplifies the key challenges of chemical hydrides as the decomposition does not cleanly yield H_2 and a simple spent fuel. Instead, it produces boron–nitrogen byproducts (e.g. polyborazylene and borazine among others) that are chemically complex. Once AB has released its hydrogen, the remaining spent material is a mixture of solid residues that are difficult to reprocess back into AB. Regeneration of ammonia borane from these B–N spent

products is a complex task requiring multiple chemical steps and substantial energy input, which undermines its practicality [21]. Additionally, the hydrogen release from AB can be accompanied by trace amounts of ammonia or borazine, impurities that can damage fuel cells if not managed. Similar issues affect other high-hydrogen chemical hydrides like hydrazine borane or metal amine-boranes, which have been explored but often suffer from stability or handling issues [24, 25]. In general, the advantages of chemical hydrides like AB are their extraordinarily high hydrogen content and the possibility of rapid hydrogen generation at low temperature, while their disadvantages include stability issues, potential side-product formation, and the lack of a simple regeneration cycle [21].

In summary, solid hydrogen carriers present an innovative solution to store hydrogen in a dense and safer form, aligning with the needs of a future zero-emission energy landscape [1]. They offer clear benefits in theory (high hydrogen density, no need for extreme pressurization, and potential for reversibility), but each class has its own set of challenges in practice, from thermodynamic limitations to recharging difficulties. The current state of research reflects a balancing act: improving the performance of these materials (via catalysts, nanostructuring, new compositions) while also developing processes to efficiently regenerate and reuse the carriers. Achieving a practical closed-loop hydrogen carrier fuel cycle, paired with efficient, low-waste regeneration, would be disruptive for the hydrogen economy. Figure 1.1 illustrates this compromise explicitly, mapping the practical release temperature of representative metal, complex, and molecular chemical hydrides against their achievable gravimetric hydrogen capacities and regeneration capability.

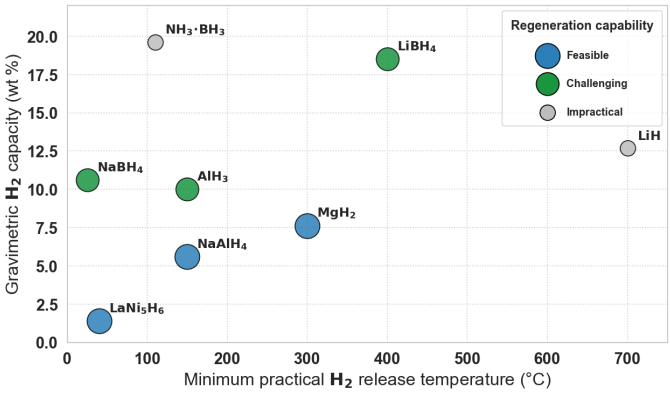


Figure 1.1: Trade-off between practical hydrogen release temperature, gravimetric capacity, and regeneration capability for representative solid carriers.

1.1.2. Towards the regeneration of sodium borohydride

Among the different solid hydrogen carrier alternatives, sodium borohydride ($NaBH_4$) stands out because it combines a good gravimetric capacity (10.92 wt% H_2) with

powder-form stability, benign handling, and an existing industrial supply chain [26]. These qualities make it one of the few materials realistic for on-board hydrogen generation [12, 14]. It can also deliver hydrogen on demand through simple, room-temperature hydrolysis (see Equation 1.1), yet it is non-volatile, non-pyrophoric, and transportable without high-pressure cylinders or cryogenic insulation. This unmatched safety-energy-density balance makes it particularly attractive for vehicles and portable power systems [27, 28].

$$NaBH_4 + (2 + x)H_2O \rightarrow NaBO_2 \cdot xH_2O + 4\,H_2 \tag{1.1}$$

where x is the level of hydration.

The critical limitation of $NaBH_4$ is the difficulty of regenerating it after use. When $NaBH_4$ releases hydrogen via hydrolysis, it is converted into sodium metaborate ($NaBO_2$) or related borates. Converting $NaBO_2$ back into $NaBH_4$ requires the input of considerable energy and hydrogen, essentially reversing the original reaction. Conventional chemical methods for $NaBH_4$ production, such as the Brown-Schlesinger process, are complex and expensive, which have so far prevented $NaBH_4$ from being a truly sustainable hydrogen carrier in a closed cycle [29]. This challenge has motivated researchers to seek new regeneration methods that are more energy-efficient and practical. One promising approach is mechanochemical regeneration in which a high-energy ball mill is used to drive the chemical reduction of $NaBO_2$ back to $NaBH_4$. Recent studies have shown that by ball milling sodium metaborate with reductive agents, such as MgH_2, it is possible to synthesize $NaBH_4$ directly from its spent form at room temperature, without the need for high-temperature furnaces or high-pressure hydrogen gas [17, 30]. This mechanochemical route essentially uses mechanical action in the form of impact and shear forces to overcome kinetic barriers and facilitate the reaction of the metaborate with H_2-bearing solids, forming fresh $NaBH_4$. While still in early stages, such approaches hint at a feasible recycling loop for $NaBH_4$ that could dramatically improve its practicality as a hydrogen carrier [31] (Figure 1.2).

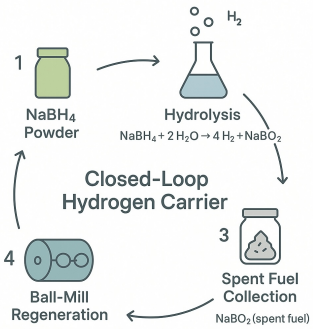


Figure 1.2: Basic $NaBH_4$ closed-loop

1.2. Mechanochemistry

Mechanochemistry is the branch of chemistry concerned with chemical reactions induced by mechanical energy, typically by grinding or milling solid reactants [32]. Specifically in high-energy ball milling, mechanical forces from collisions and friction drive reactions without the need for bulk solvents or high temperatures (see Figure 1.3) [17, 33]. This solvent-free, room-temperature approach is attractive for sustainable synthesis, as it avoids hazardous reagents and can enable transformations not attainable by conventional methods [34]. Mechanochemistry has been applied across materials science, from inorganic alloy formation to organic synthesis [33, 35, 36].

1.2.1. Reproducibility gap and standardization challenges

The outcomes of mechanochemical reactions are highly sensitive to milling conditions. Important parameters include the ball-to-powder ratio (BPR), which represents the mass ratio of milling media to reactant powders, the milling time, rotational or oscillation speed, and other factors like ball size, fill volume of the jar, and even the type of mill [37, 38]. Despite this clear cause-and-effect cycle, mechanochemical reactions have a notorious reputation for poor reproducibility since the literature underreports or overlooks some of these parameters. For instance, significantly different product yields are often reported for the same starting materials, owing to variations in milling conditions and equipment. Factors like the exact jar geometry, material of the milling media, and even subtle differences in how a procedure is executed (pre-treatment of reagents, temperature, etc.) can also dramatically affect outcomes [39, 40].

In the context of solid hydrogen carriers, Kuziora et al. [41] kept a nominal BPR constant but systematically varied jar volume, ball size, and powder mass while testing the hydrogen-storage performance of MgH_2. The resulting capacities diverged widely, proving that a single BPR value cannot guarantee identical milling conditions or outcomes. In other applications, Julien et al. [38] used in-situ Raman spectroscopy inside a mixer mill to follow a Knoevenagel condensation and showed that varying the milling frequency led to different regimes of reaction kinetics. Likewise, an inter-laboratory benchmarking study concluded that dependencies on mill model and jar size remain “globally overlooked” in published procedures, frustrating theory-to-practice transfer and reproducibility [42]. Collectively, these findings illustrate the common pattern

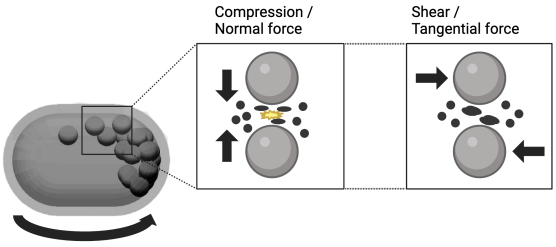


Figure 1.3: Mechanical action inside milling machine

of reporting only a fraction of relevant parameters, even though neglected variables can affect yields or selectivity by orders of magnitude. Addressing that blind spot, for example, by full disclosure of milling geometry, fill level, media composition, and liquid additives, is imperative for genuine reproducibility and for building transferable kinetic models. As expected, the same limitations are encountered for the mechanochemical regeneration of $NaBH_4$ [17]. Figure 1.4 visualizes this issue by plotting the best-reported regeneration yields of different studies against rotational speeds. Even within the narrow 1000-1450 rpm window, the yields can scatter by more than 20 percentage points, and other studies cannot be placed in the abscissa at all because their mill speed was never disclosed. The plot, therefore, makes the reproducibility gap tangible and underscores the need for a complete mechanical descriptor set.

1.2.2. Real-Time monitoring and mechanistic understanding

Another fundamental limitation in mechanochemistry has been the difficulty of observing and controlling reactions in real time. Ball milling typically occurs in sealed, opaque jars undergoing rapid motion, which limits direct visual or spectroscopic monitoring of the reaction progress [43]. This lack of real-time insight means that the mechanisms of energy transfer, from mechanical to chemical, remain only partially understood. Pioneering developments in the last decade have introduced in situ monitoring techniques, such as time-resolved X-ray diffraction and Raman spectroscopy performed on modified milling setups [38, 43], offering glimpses into phase transformations during milling. However, such methods are still in their infancy and typically require specialized, modified equipment. As a result, mechanochemical reactions are often treated as “black boxes,” where only the initial and final states are known. An additional

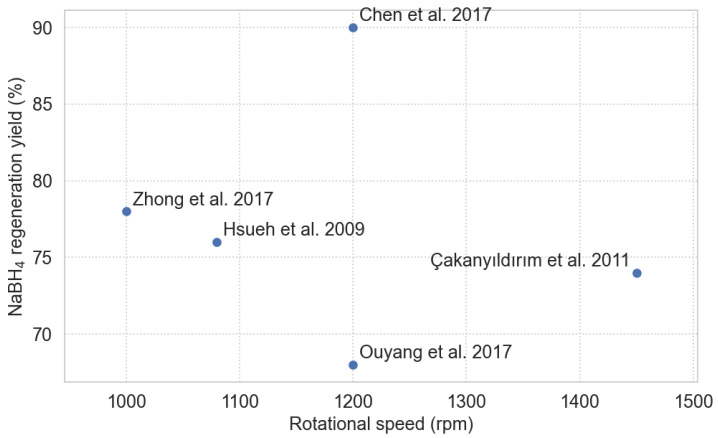


Figure 1.4: Scatter of best-reported yields for the mechanochemical regeneration of $NaBH_4$ versus the rotational speed of the ball mill. The large yield spread highlights process variability and the impact of incompletely specified milling conditions.

limitation of this is that, while lab-scale equipment may be modifiable to include these inspection technologies, the translation to large-scale equipment cannot be taken for granted. This is critical since it has already been established that differences in mills can lead to different mechanochemical outputs. Recent state-of-the-art reviews emphasize that advanced computational simulations can bridge this gap, providing a powerful complementary route for understanding and optimizing mechanochemical transformations [44].

1.2.3. Scale-up limitations

Despite promising lab-scale results, translating mechanochemistry to industrial production remains a major hurdle. Most successful reactions have been demonstrated in small laboratory mills, and scaling up the process is non-trivial. The technology developed in academic labs faces significant scale-up challenges before it can reach industry [45–48].

One of the biggest obstacles is the lack of proven strategies for continuous or large-volume mechanochemical processing. Very few demonstrations exist of mechanochemical syntheses on the kilogram or ton scale, and conventional ball mills operate in batch mode, which can become inefficient for large throughputs [45, 47]. There is ongoing research into adapting continuous milling techniques, such as twin-screw extrusion, to perform mechanochemical reactions at scale [46]. For instance, in an extruder, rotating screws can mash reactants together continuously, offering more controlled conditions and easier scalability than batch ball mills [48]. Nonetheless, significant innovation and investment are required to bridge the gap between lab and industry, and to ensure that mechanochemical processes remain efficient and controllable at scale.

1.3. Research objective & thesis outline

This thesis confronts a two-fold challenge. On the one hand, mechanochemical $NaBH_4$ regeneration is a promising route towards a closed borohydride fuel cycle, yet the published data remains fragmented and hard to reproduce. On the other hand, mechanochemistry still lacks unified protocols, representative diagnostics, and scale-up strategies. The goal of this work is therefore to close both gaps by delivering a robust, transferable route to sustainable $NaBH_4$ regeneration, and a general template for reliable, scalable mechanochemistry. The following main research question is therefore defined:

Which mechanochemical conditions, descriptors, and reporting standards enable reproducible, scalable, and energy-efficient regeneration of $NaBH_4$?

To unpack this main question, the following sub-research questions are defined:

- What are the key operational variables and interactions in mechanochemical $NaBH_4$ regeneration, and do multiple operating windows produce comparable yields?
- How can the internal dynamics of a mechanochemical process be linked to machine-level operating conditions to facilitate reproducibility and transferability across mills and scales?
- Does mechanochemical yield depend more on total mechanical input or its partitioning between normal and tangential stressing, and how can we test this experimentally?
- To what extent can the mechanochemical yield be predicted by coupling mechanical descriptors with chemical stoichiometry via machine learning models?
- How can physics-based simulations, specifically discrete element modeling, be accelerated through surrogate or reduced-order approaches to reduce computational cost?

These questions outline the path from mechanistic understanding to scalable technology, situating the work within the broader landscape of green chemistry innovation. The thesis, therefore, unfolds in three distinct parts:

I (chapter 2) establishes an efficient experimental methodology that maps the vast operational space of a typical mechanochemical process with a reduced set of trials. This allows operational parameters to be ranked against yield targets and the identification of statistical significance. Also, a water-based yield quantification protocol replaces conventional handling of ethylenediamine and simultaneously produces a "ready-to-use" solution of $NaBH_4$.

II (chapter 3 and chapter 4) turns to physics-based modeling. In chapter 3, Discrete Element Method (DEM) simulations are used to define a set of mechanically invariant descriptors that collapse multiple milling conditions onto so-called master curves. Framed in this way, any mill, at any scale, can be compared or tuned using a single 'mechanical blueprint', allowing fair comparisons between equipment and forward predictions of conditions that favor a desired mechanochemical outcome. Then, in chapter 4, the master curve framework is exploited to fairly assess the influence of the distribution between tangential and normal stressing events in the system, thereby providing the first unbiased benchmark for bulk-scale $NaBH_4$ mechanochemical reactivity.

III (chapter 5 and chapter 6) combines data-driven tools to extend both reach and practicality. In chapter 5, machine learning models fuse experimental conditions with the DEM-derived descriptors to predict regeneration across the entire design space, steering experiments toward high-yield regions while cutting trial-and-error costs. Finally, chapter 6 introduces a graph neural network surrogate that reproduces

DEM physics at a fraction of the computational cost, lowering the entry barrier for rapid mill design and scale-up studies.

Collectively, the three parts progress from systematic experimentation, through mechanistic insight, to predictive and accelerated design of mechanochemical processes. In doing so, the thesis offers both a concrete pathway for circular hydrogen carriers based on $NaBH_4$ and a broadly applicable framework for next-generation industrial mechanochemistry.

References

[1] J. Abe, A. Popoola, E. Ajenifuja and O. Popoola. 'Hydrogen energy, economy and storage: Review and recommendation'. In: *International Journal of Hydrogen Energy* 44.29 (2019), pp. 15072–15086. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2019.04.068. URL: https://www.sciencedirect.com/science/article/pii/S036031991931465X.

[2] B. Sakintuna, F. Lamari-Darkrim and M. Hirscher. 'Metal hydride materials for solid hydrogen storage: A review'. In: *International Journal of Hydrogen Energy* 32.9 (2007), pp. 1121–1140. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2006.11.022. URL: https://www.sciencedirect.com/science/article/pii/S0360319906005866.

[3] A. S. Mekonnin, K. Wacławiak, M. Humayun, S. Zhang and H. Ullah. 'Hydrogen Storage Technology, and Its Challenges: A Review'. In: *Catalysts* 15.3 (2025). ISSN: 2073-4344. DOI: 10.3390/catal15030260. URL: https://www.mdpi.com/2073-4344/15/3/260.

[4] Z. Chen, Z. Ma, J. Zheng, X. Li, E. Akiba and H. Y. Li. 'Perspectives and challenges of hydrogen storage in solid-state hydrides'. In: *Chinese Journal of Chemical Engineering* 29 (Jan. 2021), pp. 1–12. DOI: 10.1016/j.cjche.2020.08.024. URL: https://doi.org/10.1016/j.cjche.2020.08.024.

[5] S.-i. Orimo, Y. Nakamori, J. R. Eliseo, A. Züttel and C. M. Jensen. 'Complex Hydrides for Hydrogen Storage'. In: *Chemical Reviews* 107.10 (2007). PMID: 17848101, pp. 4111–4132. DOI: 10.1021/cr0501846. eprint: https://doi.org/10.1021/cr0501846. URL: https://doi.org/10.1021/cr0501846.

[6] T. G. Manivasagam, K. Kiraz and P. H. L. Notten. 'Electrochemical and Optical Properties of Magnesium-Alloy Hydrides Reviewed'. In: *Crystals* 2.4 (2012), pp. 1410–1433. ISSN: 2073-4352. URL: https://www.mdpi.com/2073-4352/2/4/1410.

[7] V. Stavila, S. Li, C. Dun, M. A. T. Marple, H. E. Mason, J. L. Snider, J. E. Reynolds III, F. El Gabaly, J. D. Sugar, C. D. Spataru, X. Zhou, B. Dizdar, E. H. Majzoub, R. Chatterjee, J. Yano, H. Schlomberg, B. V. Lotsch, J. J. Urban, B. C. Wood and M. D. Allendorf. 'Defying Thermodynamics: Stabilization of Alane Within Covalent Triazine Frameworks for Reversible Hydrogen Storage'. In: *Angewandte Chemie International Edition* 60.49 (2021), pp. 25815–25824. DOI: https://doi.org/10.1002/anie.202107507. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.202107507. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.202107507.

[8] M. Pęska, J. Dworecka-Wójcik, T. Płociński and M. Polański. 'The Influence of Cerium on the Hydrogen Storage Properties of La1-xCexNi5 Alloys'. In: *Energies* 13.6 (2020). ISSN: 1996-1073. DOI: 10.3390/en13061437. URL: https://www.mdpi.com/1996-1073/13/6/1437.

[9] P. T. Aakko-Saksa, C. Cook, J. Kiviaho and T. Repo. ‘Liquid organic hydrogen carriers for transportation and storing of renewable energy – Review and discussion’. In: *Journal of Power Sources* 396 (2018), pp. 803–823. ISSN: 0378-7753. DOI: https://doi.org/10.1016/j.jpowsour.2018.04.011. URL: https://www.sciencedirect.com/science/article/pii/S0378775318303483.

[10] X. Zhang, Y. Liu, R. Zhuanghe, X. Zhang, J. Hu, Z. Huang, Y. Lu, M. Gao and H. Pan. ‘Realizing 6.7 wt% reversible storage of hydrogen at ambient temperature with non-confined ultrafine magnesium hydride’. In: *Energy & Environmental Science* 14 (Dec. 2020). DOI: 10.1039/D0EE03160G.

[11] J.-C. Crivello, B. Dam, R. Denys, M. Dornheim, D. Grant, J. Huot, T. Jensen, P. Jongh, M. Latroche, C. Milanese, D. Milčius, G. Walker, J. Webb, C. Zlotea and V. Yartys. ‘Review of magnesium hydride-based materials: Development and optimisation’. In: *Applied Physics A* 122 (Jan. 2016), 122:97. DOI: 10.1007/s00339-016-9602-0.

[12] E. S. V. Rheenen, J. T. Padding, J. C. Slootweg and K. V. and. ‘Hydrogen carriers for zero-emission ship propulsion using PEM fuel cells: an evaluation’. In: *Journal of Marine Engineering & Technology* 23.3 (2024), pp. 166–183. DOI: 10.1080/20464177.2023.2282691. eprint: https://doi.org/10.1080/20464177.2023.2282691. URL: https://doi.org/10.1080/20464177.2023.2282691.

[13] W. Zhang, X. Zhang, Z. Huang, H.-W. Li, M. Gao, H. Pan and Y. Liu. ‘Recent Development of Lithium Borohydride-Based Materials for Hydrogen Storage’. In: *Advanced Energy and Sustainability Research* 2.10 (2021), p. 2100073. DOI: https://doi.org/10.1002/aesr.202100073. eprint: https://advanced.onlinelibrary.wiley.com/doi/pdf/10.1002/aesr.202100073. URL: https://advanced.onlinelibrary.wiley.com/doi/abs/10.1002/aesr.202100073.

[14] Z. Lu and L. Zhang. ‘Recent advances in sodium borohydride for hydrogen storage’. In: *E3S Web of Conferences* 385 (May 2023). DOI: 10.1051/e3sconf/202338504025.

[15] Q. He, D. Zhu, X. Wu, D. Dong, X. Jiang and M. Xu. ‘The Dehydrogenation Mechanism and Reversibility of LiBH4 Doped by Active Al Derived from AlH3’. In: *Metals* 9.5 (2019). ISSN: 2075-4701. DOI: 10.3390/met9050559. URL: https://www.mdpi.com/2075-4701/9/5/559.

[16] Y. Xu, Y. Zhou, Y. Li, M. Ashuri and Z. Ding. ‘Engineering LiBH4-Based Materials for Advanced Hydrogen Storage: A Critical Review of Catalysis, Nanoconfinement, and Composite Design’. In: *Molecules* 29.23 (2024). ISSN: 1420-3049. DOI: 10.3390/molecules29235774. URL: https://www.mdpi.com/1420-3049/29/23/5774.

[17] S. Garrido Nuñez, D. L. Schott and J. T. Padding. 'Optimization of operational parameters in the mechanochemical regeneration of sodium borohydride (NaBH4)'. In: *International Journal of Hydrogen Energy* 97 (2025), pp. 640–648. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2024.11.360. URL: https://www.sciencedirect.com/science/article/pii/S0360319924050511.

[18] S. Munshi, G. S. Walker, K. Manickam, T. Hansen, M. Dornheim and D. M. Grant. 'Understanding the reaction pathway of lithium borohydride-hydroxide-based multi-component systems for enhanced hydrogen storage'. In: *J. Mater. Chem. A* 12 (41 2024), pp. 28326–28336. DOI: 10.1039/D4TA05368K. URL: http://dx.doi.org/10.1039/D4TA05368K.

[19] Q. Lai, Y. Sun, T. Wang, P. Modi, C. Cazorla, U. B. Demirci, J. R. Ares Fernandez, F. Leardini and K.-F. Aguey-Zinsou. 'How to Design Hydrogen Storage Materials? Fundamentals, Synthesis, and Storage Tanks'. In: *Advanced Sustainable Systems* 3.9 (2019), p. 1900043. DOI: https://doi.org/10.1002/adsu.201900043. eprint: https://advanced.onlinelibrary.wiley.com/doi/pdf/10.1002/adsu.201900043. URL: https://advanced.onlinelibrary.wiley.com/doi/abs/10.1002/adsu.201900043.

[20] A. G. Gebretatios, F. Banat and C. K. Cheng. 'A critical review of hydrogen storage: toward the nanoconfinement of complex hydrides from the synthesis and characterization perspectives'. In: *Sustainable Energy Fuels* 8 (22 2024), pp. 5091–5130. DOI: 10.1039/D4SE00353E. URL: http://dx.doi.org/10.1039/D4SE00353E.

[21] Q. Liu, W. Ran, W. Bao and Y. Li. 'A Review on Catalytic Hydrolysis of Ammonia Borane for Hydrogen Production'. In: *Energies* 18.5 (2025). ISSN: 1996-1073. DOI: 10.3390/en18051105. URL: https://www.mdpi.com/1996-1073/18/5/1105.

[22] C. W. Hamilton, M. M. Montemore and M. R. Zachariah. 'Pyrolytic decomposition of ammonia borane to boron nitride'. In: *Inorganic Chemistry* 50 (3 2011), pp. 881–889. DOI: 10.1021/ic101020k.

[23] X. Zhang, L. Kam and T. J. Williams. 'Dehydrogenation of ammonia borane through the third equivalent of hydrogen'. In: *Dalton Trans.* 45 (18 2016), pp. 7672–7677. DOI: 10.1039/C6DT00604C. URL: http://dx.doi.org/10.1039/C6DT00604C.

[24] B. Roy, A. Hajari, J. Manna and P. Sharma. 'Supported ammonia borane decomposition through enhanced homopolar B–B coupling'. In: *Dalton Trans.* 47 (18 2018), pp. 6570–6579. DOI: 10.1039/C8DT00789F. URL: http://dx.doi.org/10.1039/C8DT00789F.

[25] T. A. Seminary, A. M. Stowe, T. Autrey, B. M. Schmid, W. J. Shaw, J. C. Linehan and J. J. Warren. 'Spectroscopic studies of dehydrogenation of ammonia borane in carbon cryogel'. In: *The Journal of Physical Chemistry C* 112 (1 2008), pp. 242–247. DOI: 10.1021/jp077057t.

[26] H. X. Nunes, D. L. Silva, C. M. Rangel and A. M. F. R. Pinto. 'Rehydrogenation of Sodium Borates to Close the NaBH4-H2 Cycle: A Review'. In: *Energies* 14.12 (2021). ISSN: 1996-1073. DOI: 10.3390/en14123567. URL: https://www.mdpi.com/1996-1073/14/12/3567.

[27] M. Dragan. 'Hydrogen Storage in Complex Metal Hydrides NaBH4: Hydrolysis Reaction and Experimental Strategies'. In: *Catalysts* 12.4 (2022). ISSN: 2073-4344. DOI: 10.3390/catal12040356. URL: https://www.mdpi.com/2073-4344/12/4/356.

[28] S. S. Muir and X. Yao. 'Progress in sodium borohydride as a hydrogen storage material: Development of hydrolysis catalysts and reaction systems'. In: *International Journal of Hydrogen Energy* 36.10 (2011), pp. 5983–5997. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2011.02.032. URL: https://www.sciencedirect.com/science/article/pii/S0360319911003491.

[29] C. Lang, Y. Jia, J. Liu, H. Wang, L. Ouyang, M. Zhu and X. Yao. 'NaBH 4 regeneration from NaBO 2 by high-energy ball milling and its plausible mechanism'. In: *International Journal of Hydrogen Energy* 42 (Apr. 2017). DOI: 10.1016/j.ijhydene.2017.04.014.

[30] M. Huang, H. Zhong, L. Ouyang, C. Peng, X. Zhu, W. Zhu, F. Fang and M. Zhu. 'Efficient regeneration of sodium borohydride via ball milling dihydrate sodium metaborate with magnesium and magnesium silicide'. In: *Journal of Alloys and Compounds* 729 (Sept. 2017). DOI: 10.1016/j.jallcom.2017.09.262.

[31] L. Kong, C. Xinyu, H. Jin, J. Wu, H. Du and T. Xiong. 'Mechanochemical Synthesis of Sodium Borohydride by Recycling Sodium Metaborate'. In: *Energy & Fuels* 23 (Sept. 2009). DOI: 10.1021/ef900619y.

[32] T. H. El-Sayed, A. Aboelnaga, M. A. El-Atawy and M. Hagar. 'Ball Milling Promoted N-Heterocycles Synthesis'. In: *Molecules* 23.6 (2018). ISSN: 1420-3049. DOI: 10.3390/molecules23061348. URL: https://www.mdpi.com/1420-3049/23/6/1348.

[33] J.-L. Do and T. Friščić. 'Mechanochemistry: A Force of Synthesis'. In: *ACS Central Science* 3.1 (2017). PMID: 28149948, pp. 13–19. DOI: 10.1021/acscentsci.6b00277. eprint: https://doi.org/10.1021/acscentsci.6b00277. URL: https://doi.org/10.1021/acscentsci.6b00277.

[34] V. Štrukil. 'Mechanochemical Organic Synthesis: The Art of Making Chemistry Green'. In: *Synlett* 29 (Jan. 2018). DOI: 10.1055/s-0036-1591868.

[35] L. M. Martínez, J. Cruz-Angeles, M. Vázquez-Dávila, E. Martínez, P. Cabada, C. Navarrete-Bernal and F. Cortez. 'Mechanical Activation by Ball Milling as a Strategy to Prepare Highly Soluble Pharmaceutical Formulations in the Form of Co-Amorphous, Co-Crystals, or Polymorphs'. In: *Pharmaceutics* 14.10 (2022). ISSN: 1999-4923. DOI: 10.3390/pharmaceutics14102003. URL: https://www.mdpi.com/1999-4923/14/10/2003.

[36] V. A. Rodriguez, L. Ribas, A. Kwade and L. M. Tavares. 'Mechanistic modeling and simulation of a wet planetary ball mill'. In: *Powder Technology* 429 (2023), p. 118901. ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2023.118901. URL: https://www.sciencedirect.com/science/article/pii/S003259102300685X.

[37] E. Colacino, M. Carta, G. Pia, A. Porcheddu, P. C. Ricci and F. Delogu. 'Processing and Investigation Methods in Mechanochemical Kinetics'. In: *ACS Omega* 3.8 (2018). PMID: 31459054, pp. 9196–9209. DOI: 10.1021/acsomega.8b01431. eprint: https://doi.org/10.1021/acsomega.8b01431. URL: https://doi.org/10.1021/acsomega.8b01431.

[38] P. A. Julien, I. Malvestiti and T. Friščić. 'The effect of milling frequency on a mechanochemical organic reaction monitored by in situ Raman spectroscopy'. In: *Beilstein Journal of Organic Chemistry* 13 (2017), pp. 2160–2168. ISSN: 1860-5397. DOI: 10.3762/bjoc.13.216. URL: https://doi.org/10.3762/bjoc.13.216.

[39] S. Garrido Nuñez, D. L. Schott and J. T. Padding. 'Predictive models for energy dissipation in mechanochemical ball milling'. In: *Powder Technology* 457, 120919 (2025). ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2025.120919. URL: https://www.sciencedirect.com/science/article/pii/S0032591025003146.

[40] A. M. Belenguer, G. I. Lampronti and J. K. M. Sanders. 'Reliable Mechanochemistry: Protocols for Reproducible Outcomes of Neat and Liquid Assisted Ball-mill Grinding Experiments'. In: *Journal of Visualized Experiments* 131 (Jan. 2018), p. 56824. DOI: 10.3791/56824.

[41] P. Kuziora, M. Wyszyńska, M. Polanski and J. Bystrzycki. 'Why the ball to powder ratio (BPR) is insufficient for describing the mechanical ball milling process'. In: *International Journal of Hydrogen Energy* 39 (June 2014), pp. 9883–9887. DOI: 10.1016/j.ijhydene.2014.03.009.

[42] O. Jafter, S. Lee, J. Park, C. Cabanetos and D. Lungerich. 'Navigating Ball Mill Specifications for Theory-to-Practice Reproducibility in Mechanochemistry'. In: *Angewandte Chemie International Edition* 63 (Oct. 2024), e202409731. DOI: 10.1002/anie.202409731.

[43] S. Lukin, L. S. Germann, T. Friščić and I. Halasz. 'Toward Mechanistic Understanding of Mechanochemical Reactions Using Real-Time *In Situ* Monitoring'. In: *Accounts of Chemical Research* 55.9 (May 2022). PMID: 35446551 Epub 2022 Apr 21, pp. 1262–1277. DOI: 10.1021/acs.accounts.2c00062.

[44] L. Dong, L. Li, H. Chen, Y. Cao and H. Lei. 'Mechanochemistry: Fundamental Principles and Applications'. In: *Advanced Science* 12.24 (2025), e2403949. DOI: 10.1002/advs.202403949. URL: https://doi.org/10.1002/advs.202403949.

[45] E. Colacino, V. Isoni, D. Crawford and F. Garcia. 'Upscaling Mechanochemistry: Challenges and Opportunities for Sustainable Industry'. In: *Trends in Chemistry* 3 (Mar. 2021). DOI: 10.1016/j.trechm.2021.02.008.

[46] M. Baláž, V. Balema, J. D. Batteas, R. G. Blair, C. Bolm, L. Borchardt, A. B. Braunschweig, S. L. Craig, F. Emmerling, M. Ferguson, T. Friščić, S. James, J. Leitch, J. Mack, S. Mohamed, K. Nagapudi, F. Puccetti and M. E. Rivas. 'Scale up and industrial implementation: general discussion'. In: *Faraday Discuss.* 241 (0 2023), pp. 387–393. DOI: 10.1039/D2FD90083A. URL: http://dx.doi.org/10.1039/D2FD90083A.

[47] J. F. Reynes, V. Isoni and F. García. 'Tinkering with Mechanochemical Tools for Scale Up'. In: *Angewandte Chemie International Edition* 62.44 (2023), e202300819. DOI: https://doi.org/10.1002/anie.202300819. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.202300819. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.202300819.

[48] D. E. Crawford, C. K. G. Miskimmin, A. B. Albadarin, G. Walker and S. L. James. 'Organic synthesis by Twin Screw Extrusion (TSE): continuous, scalable and solvent-free'. In: *Green Chem.* 19 (6 2017), pp. 1507–1518. DOI: 10.1039/C6GC03413F. URL: http://dx.doi.org/10.1039/C6GC03413F.

2

Optimization of operational variables and interactions for $NaBH_4$ regeneration

In this chapter, we investigate mechanochemical $NaBH_4$ regeneration from $NaBO_2 \cdot 4\,H_2O$ and MgH_2 on a Retsch Emax ball mill. A screening design of experiments identifies the key operational variables and their interactions. We map operating windows in the space of rotational speed, milling time, BPR (ball-to-powder ratio), and molar ratio, showing when different combinations reach comparable yields with different trade-offs. A water-based method facilitates conversion quantification while avoiding toxic solvents and produces a ready-to-use $NaBH_4$ solution.

Parts of this chapter are adapted from: Garrido Nuñez, S., Schott, D. L. and Padding, J. T. 'Optimization of operational parameters in the mechanochemical regeneration of sodium borohydride ($NaBH_4$)'. In: International Journal of Hydrogen Energy 97, 640-648 (2025).

2.1. Introduction

The ongoing transition to low-carbon and no-carbon energy systems has incentivized the creation of many green power generation solutions that have the potential to sustain our energetic needs. This transition faces a great challenge in finding an effective and efficient energy carrier that can keep up with the demands of high-energy industries and applications such as maritime transport [1, 2]. One of the alternatives to mitigate fossil fuel dependency is hydrogen, a clean energy carrier with zero emissions. However, hydrogen storage and transportation in pure gaseous or liquid form is challenging due to high-pressure or low-temperature working conditions [3, 4].
A potential solution involves using solid hydrogen carriers, which enable the storage of hydrogen at ambient temperature and pressure conditions. Sodium borohydride ($NaBH_4$) has a high theoretical gravimetric hydrogen storage capacity (10.92 wt%) and thus, is a promising solid hydrogen carrier [5]. Hydrogen can be released via the hydrolysis of $NaBH_4$ (Equation 4.1) with dry or hydrated sodium metaborate ($NaBO_2 \cdot xH_2O$) as byproduct, typically referred to as spent fuel. Therefore, the regeneration of $NaBH_4$ from the spent fuel is critical for considering it a viable contributor to the energy transition since it would allow its usage in a circular, cheap, and sustainable manner [6, 7].

$$NaBH_4 + (2 + x)H_2O \rightarrow NaBO_2 \cdot xH_2O + 4\,H_2 \tag{2.1}$$

where x is the level of hydration.

Mechanochemical, electrochemical, and thermochemical methods have been reported for regenerating $NaBH_4$ [6, 8, 9]. Thermochemical processes require high-pressure and temperature conditions, while electrochemical methods tend to be inefficient and can produce toxic or harmful byproducts that have a negative impact on the environment. Mechanochemical methods, in contrast, are appealing and environmentally friendly options because they function without the need for an electrolyte solution and can be performed at room temperature and pressure conditions [6].

Within a mechanochemical process, the mechanical action of the system induces the breakage of primary bonds, giving rise to surface reconstruction and chemical reactions among the surrounding media as milling balls collide [10]. For the chemical reaction to take place, a sufficiently large contact area and contact time must be involved in the process. Otherwise, the reactants may not have sufficient interplay.
In the specific case of $NaBH_4$, its mechanochemical regeneration has been reported to be viable in high-energy mills [11–17]. A summary of the highest yield obtained in different studies is presented in Table 4.1, which are based on the following five distinct chemical processes.

$$NaBO_2 + 2\,MgH_2 \rightarrow NaBH_4 + 2\,MgO \tag{2.2}$$

$$NaBO_2 \cdot 2\,H_2O + 4\,Mg \rightarrow NaBH_4 + 4\,MgO \tag{2.3}$$

$$NaBO_2 \cdot 2\,H_2O + 4\,MgH_2 \rightarrow NaBH_4 + 4\,MgO + 4\,H_2 \tag{2.4}$$

$$NaBO_2 \cdot 4\,H_2O + 6\,MgH_2 \rightarrow NaBH_4 + 6\,MgO + 8\,H_2 \tag{2.5}$$

$$NaBO_2 \cdot 2\,H_2O + 2\,Mg_2Si \rightarrow NaBH_4 + 4\,MgO + 2\,Si \tag{2.6}$$

It can be noticed that significantly different yields have been achieved using the same reactants. This variation is due to differences in the working conditions of the ball mills used (e.g., rotational speed, fill ratio, ball size, jar shape, ball mill motion, or ball-to-powder ratio (BPR)). These parameters fundamentally affect the mechanical action within the milling jar. It is important to emphasize that mechanical conditions depend on the specific ball mill and the collisions occurring inside the jar. Therefore, simply replicating operational conditions does not guarantee the same chemical yield unless the same mechanical action inside the jar is ensured. The process parameters reported in the studies shown in Table 4.1 are presented in Table 2.2.
The results presented in Table 4.1 and Table 2.2 highlight the significant potential of mechanochemistry in facilitating the regeneration of $NaBH_4$, which serves as a motivation for this study. However, it is worth noting that many authors have omitted crucial details for the complete characterization of their experimental setup, thereby hindering reproducibility.

Moreover, most authors typically investigate the impact of only one or two operational variables at a time while holding the remaining variables constant. For instance, Hsueh et al. [11] simultaneously varied milling time and molar ratio while keeping rotational speed and BPR constant. In contrast, Lang et al. [14], Çakanyildirim et al. [13], and Chen et al. [16] independently varied milling time and molar ratio while maintaining the remaining variables fixed. Kong et al. [12] independently varied milling time, molar ratio, and ball-to-powder ratio (BPR) while keeping the rest of the variables constant. Notably, the work of Ouyang et al. [15] stands out as they studied the simultaneous effects of varying milling time, molar ratio, and BPR. However, their maximum yield fell short compared to other studies. Finally, Zhong et al. [17] also analyzed the combined influence of milling time and molar ratio while keeping the other variables constant.
It is understandable why authors have decided to follow this approach. The number of

Table 2.1: Reported yields for the mechanochemical regeneration of $NaBH_4$.

Metaborate	2nd reactant	Yield (%)	Year	Ref.
$NaBO_2$	MgH_2	76	2009	[11]
$NaBO_2$	MgH_2	71	2009	[12]
$NaBO_2$	MgH_2	74	2011	[13]
$NaBO_2$	MgH_2	89	2017	[14]
$NaBO_2 \cdot 2\,H_2O$	Mg	68	2017	[15]
$NaBO_2 \cdot 2\,H_2O$	MgH_2	90	2017	[16]
$NaBO_2 \cdot 4\,H_2O$	MgH_2	88	2017	[16]
$NaBO_2 \cdot 2\,H_2O$	Mg_2Si	78	2017	[17]

Table 2.2: Process parameters reported for the mechanochemical regeneration of $NaBH_4$. Parameters that were *not* specified in the original papers are indicated with an en-dash (–).

Reference	[11]	[12]	[13]	[14]
Metaborate	$NaBO_2$	$NaBO_2$	$NaBO_2$	$NaBO_2$
2nd reactant	MgH_2 (98 %)	MgH_2 (79.3 %)	MgH_2 (99 %)	MgH_2 (95 %)
Mass metaborate (g)	0.66	–	0.86	–
Mass 2nd reactant (g)	0.26	–	0.89	–
Total mass (g)	0.92	–	1.75	–
Molar ratio	2.8	2.0	2.6	2.7
Rotational speed (rpm)	1080	–	1450	–
Milling time (h)	6	2	12	12
Jar volume (mL)	65	–	45	–
Ball diameter (mm)	13	–	10/4	–
Number of balls	4	–	2/3	–
Ball volume in jar (%)	7	–	2.5	–
Ball-to-powder ratio (kg:kg)	39	50	10	50
Mill type	SPEX 8000	QM-3A	CertiPrep 8000M	QM-3A
Milling vessel material	Steel	–	–	–
Yield (%)	76	71	70	89

Reference	[15]	[16]	[17]
Metaborate	$NaBO_2 \cdot 2\,H_2O$	$NaBO_2 \cdot 4\,H_2O$	$NaBO_2 \cdot 2\,H_2O$
2nd reactant	Mg (99.8 %)	MgH_2 (99 %)	Mg_2Si (99.5 %)
Mass metaborate (g)	0.43	0.44	0.31
Mass 2nd reactant (g)	0.57	0.56	0.69
Total mass (g)	1.00	1.00	1.00
Molar ratio	5.5	5.0	3.0
Rotational speed (rpm)	1200	1200	1000
Milling time (h)	10	15	20
Jar volume (mL)	80	80	–
Ball diameter (mm)	–	–	–
Number of balls	–	–	–
Ball volume in jar (%)	8	8	–
Ball-to-powder ratio (kg:kg)	50	50	50
Mill type	QM-3C	QM-3C	QM-3C
Milling vessel material	Steel	Steel	–
Yield (%)	68	90	78

experiments increases exponentially as more parameters are varied simultaneously, and the effect can be even more dramatic depending on the number of levels or values assigned to these variables. However, it is crucial to recognize that all these operational variables impact the chemical yield. To provide a simple example, while it may be intuitive to fix the rotational speed of the milling jar at a constant value that can deliver sufficient kinetic energy to the milling balls, one needs to realize that it also directly influences the collision frequency. This affects the number of events (effective collisions) that ultimately enable the chemical reaction to take place. If one then decreases the BPR to include more powder, the chemical yield will inevitably be reduced as the collisions available to treat the total amount of powder have effectively changed. In this situation, one could arguably compensate by increasing the rotational speed or the milling time. Needless to say, the molar ratio also plays a key role as it impacts the amount of effective collisions where both reactants are crushed together. This chain of events can be visualized in Figure 2.1.

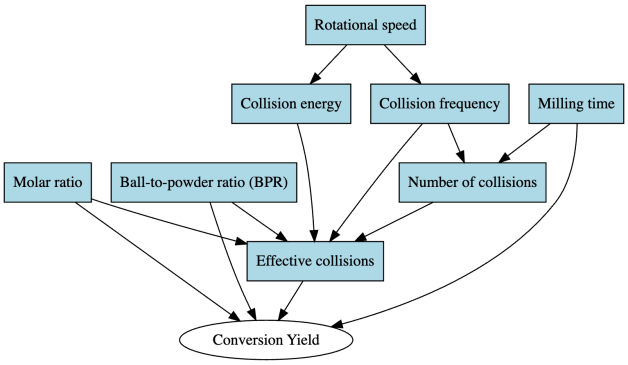


Figure 2.1: Graphical model of variable interplay

As a result of the previous observations and the insights gained by the studies shown in Table 2.2, this study aims to optimize and enhance the understanding of the mechanochemical regeneration of $NaBH_4$ by investigating the individual and joint effects of milling time, BPR, molar ratio, and rotational speed on the chemical yield. Consistent with our prior investigations, we maintain a fill ratio of 10%, achieved with twenty-four 10 mm milling balls. In our milling machine, this fill ratio achieves an optimal balance between normal and tangential dissipation during collisions, while the size of milling balls primarily affects the distribution of energy dissipation per rotation cycle [18]. Furthermore, we opt for the chemical pathway consisting of $NaBO_2 \cdot 4\,H_2O$ and MgH_2 (Equation 3.27) as this reaction has shown to enable high conversion yields and it eliminates the need to dedicate additional energy drying the sodium metaborate or to artificially create a hydrogen atmosphere inside the milling jar [16].

2.2. Materials and Methodology

2.2.1. High-Energy ball milling

The Emax high-energy ball mill is a device produced and distributed by the German company Retsch. It offers a novel approach to ball milling by combining high friction and impact results with a temperature control system, allowing for controlled grinding. The system was set up to allow a maximum temperature of 50 ˚C. The machine can allocate proprietary grinding jars with 125 ml of volume that follow a circular motion with a rotational speed n up to 2000 revolutions per minute (corresponding to an angular frequency of $\omega = 2\pi n/60 = 209$ rad/s) with an amplitude (radius) A of 1.7 cm, see Figure 6.1.

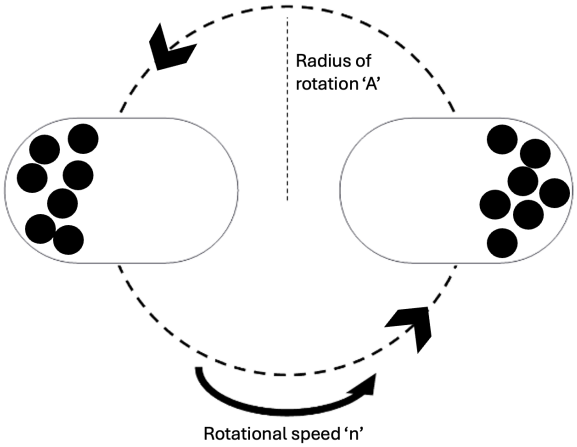


Figure 2.2: Schematic of jar movement

Stainless steel milling balls with a diameter of 10 mm (1.4034 G100 DIN 5401), purchased from Kugel Pompel, were utilized in all experiments.

2.2.2. Chemicals

Hydrated sodium metaborate ($NaBO_2 \cdot 4\,H_2O$) (≥ 99%) was purchased from Sigma-Aldrich. Magnesium hydride (MgH_2) (≥ 99.9%, ≤ 50μm) was purchased from Nanoshel. All chemicals were used as received. Additionally, all samples for ball milling were prepared in a glove box under an Argon environment where concentrations of oxygen and water were below 0.1 ppm.

2.2.3. Equipment cleaning

To preserve similar conditions for all our experimental cases, after every 3 experiments, the jars were cleaned, and the milling balls were replaced with new ones. We have found that cleaning the jar by milling 1 g of silica sand and adding 10 ml of isopropyl

alcohol provides excellent results in removing any leftover contamination. The typical duration for the cleaning process was 6 minutes. Upon finishing, the jar is then rinsed thoroughly with pure water, air-blasted with compressed air, and left to dry open to the ambient.

2.2.4. Quantification of conversion yield

In the studies presented in Table 4.1, the standard approach for quantifying the regenerated $NaBH_4$ via mechanochemistry involves utilizing ethylenediamine (EDA). EDA possesses a distinct advantage as it selectively dissolves $NaBH_4$ while leaving the remaining reactants intact, facilitating an efficient separation process through subsequent filtering and sublimation. However, EDA poses risks such as corrosion, toxicity, and health hazards. Hence, we propose a cheaper, safer, and more environmentally friendly method to assess the conversion yield, leveraging any unconverted MgH_2 in the process.
After the ball milling process is completed, the jar is opened to the atmosphere, and pure water is added. This promptly initiates the self-hydrolysis of unconverted MgH_2, elevating the solution's pH above 11 within seconds [19]. Hydrogen is released from this hydrolysis until dense passivation layers of magnesium hydroxide form over unconverted magnesium hydride [20–22]. While hydrogen released in this step is carefully disposed of, it could also be utilized to enhance overall hydrogen release in the system. The half-life of this solution, representing the time for half of the $NaBH_4$ solution to decompose, can be calculated as 10.2 hours at pH = 11 and temperature of 25 ˚C, based on the work by Kreevoy and Jacobson (Equation 2.7) [23]. Thus, a significant advantage of the rapid pH increase is the immediate inhibition of $NaBH_4$ self-hydrolysis, minimizing hydrogen losses for quantification purposes [24, 25].

$$\log_{10}(t_{1/2}) = \text{pH} - (0.034T - 1.92) \tag{2.7}$$

where $t_{1/2}$ represents the half-time in minutes and T is the temperature in Kelvin.

After the stabilization of MgH_2 by self-hydrolysis, the jar is sealed again with a lid featuring a unidirectional flow valve and connected to a gas-collection over water system. Subsequently, Ru-based catalysts are introduced into the solution within the jar, and the temperature is raised to 80˚C. This catalyst specifically assists the hydrolysis of $NaBH_4$, as the hydrolysis of MgH_2 remains relatively inactive at temperatures below 200˚C [26–30] for non-catalized mixtures and 146˚C for catalyzed mixtures [31, 32]. Moreover, Ru-based catalysts are favored for $NaBH_4$ hydrolysis due to their high hydrogen generation rates, durability, and efficient catalytic activity, particularly in comparison to other metals like Pt or Pd. These catalysts are well-suited for hydrogen-on-demand applications because of their rapid reaction rates and stability under various conditions [33–35]. Independent experiments using commercially available $NaBH_4$ demonstrated that the catalyst can hydrolyze more than 99% of these solutions under the conditions of our experimental setup.

When the hydrogen release ceases, we assess the actual volume of hydro-

gen released against the theoretical volume, derived from the ideal gas law, that would have been obtained if all initial $NaBO_2 \cdot 4\,H_2O$ had converted to $NaBH_4$ (refer to Equation 3.27). This method enables the quantification of the chemical conversion yield without the need for additional separation steps, chemicals, or equipment. Although the accuracy may not reach the same level as separating with EDA, it offers a cleaner, cheaper, and simpler process sufficient to estimate the influence of the investigated operational variables. Furthermore, it facilitates the production of a "ready-to-use" solution, the half-life of which can be extended by further pH or temperature adjustments. Such a solution offers a new approach to storing regenerated $NaBH_4$ and streamlines parallel operations. Moreover, the low solubility of the remaining solids in the jar, namely MgO and $Mg(OH)_2$, allows for a simple and straightforward separation process via filtration.

2.2.5. Experimental cases definition: Fractional design of experiments

In this paper, we investigate the simultaneous effects that varying molar ratio, BPR, milling time, and rotational speed have on the mechanochemical conversion yield. In principle, the number of experimental cases n_{cases} is defined by Equation 2.8.

$$n_{cases} = m^x \tag{2.8}$$

Where m is the number of possible values (or levels) for each variable and x is the number of variables (or factors).
By defining three general levels for each of the four variables (high(2), medium(1), and low(0)), we can account for potential non-linear behavior, and the total number of cases is 81. The number of experiments is prohibitive due to time, cost, and equipment availability constraints. As a result, we decided to employ a fractional design of experiments technique, namely screening, to reduce the number of cases needed to identify pertinent information about the main effects and two-factor interactions on the conversion yield [36].

As expected, this approach has some limitations. The most relevant is the risk of confounding, which binds together the effects of multiple factors, potentially leading to inaccurate conclusions about the true relationships between them. To minimize this risk, we select a resolution of the fourth (IV) degree, reducing the total number of experiments to 27. This configuration has no main effects confounded with interactions, but at least one pair of two-variable interactions is confounded together [37]. To define the values for each level, it is important to define a wide and reasonable range [38]. For this, we use as a reference both the operational limitations of our milling machine and the process parameters shown in Table 2.2. The levels for BPR (A) are defined as 10, 30, and 50. In the case of molar ratio (B), we use a molar ratio of 33%, 66%, and 100% compared to the stoichiometric value for MgH_2. The milling time (C) is varied between 5, 12.5, and 20 hours. Finally, the rotational speed (D) is set to 600, 800, and 1000 rpm. To ensure robustness, we applied randomization in the selection order of cases throughout the experimental process. Key cases with high conversion

results were repeated twice to confirm reproducibility. The repeatability tests showed a maximum variation of 2.5%, which is significantly smaller than the variation observed between different parameter settings. Therefore, we do not include error bars in the interaction plots. The final screening design is shown in Table 2.3, which was created with Altair's HyperStudy v2022.1.

2.3. Results and Discussion

In this section, the results of the experimental cases are presented, and we leverage the screening design of experiments to assess the relevance of each of the studied

Table 2.3: Screening design of experiments: experimental factors and levels.

Case	BPR (a)	Molar ratio (b)	Time (h) (c)	Speed (rpm) (d)	Treatment code
1	10	8	5	600	0000
2	10	8	12.5	800	0011
3	10	8	20	1000	0022
4	10	10	5	600	0100
5	10	10	12.5	1000	0112
6	10	10	20	600	0120
7	10	12	5	1000	0202
8	10	12	12.5	600	0210
9	10	12	20	800	0221
10	30	8	5	800	1001
11	30	8	12.5	1000	1012
12	30	8	20	600	1020
13	30	10	5	1000	1102
14	30	10	12.5	600	1110
15	30	10	20	800	1121
16	30	12	5	600	1200
17	30	12	12.5	800	1211
18	30	12	20	1000	1222
19	50	8	5	1000	2002
20	50	8	12.5	600	2010
21	50	8	20	800	2021
22	50	10	5	600	2100
23	50	10	12.5	800	2111
24	50	10	20	1000	2122
25	50	12	5	800	2201
26	50	12	12.5	1000	2212
27	50	12	20	600	2220

operational variables: BPR (A), molar ratio (B), milling time (C), and rotational speed (D). Additionally, we evaluate the linear (L) and quadratic (Q) dependency of each operational variable to identify trends in yield performance. Then we assess specific cases of interest that are worth discussing in more detail. The conversion results, along with the linear and quadratic mapping, are presented in Table 2.4.

Table 2.4: Conversion results for the screening experiments. Linear (L) and quadratic (Q) orthogonal contrasts are shown for the coded factors A–D: A = BPR, B = molar ratio, C = milling time, D = rotational speed.

Case	Treatment	Yield (%)	AL	AQ	BL	BQ	CL	CQ	DL	DQ
1	0000	12	-1	1	-1	1	-1	1	-1	1
2	0011	22	-1	1	-1	1	0	-2	0	-2
3	0022	30	-1	1	-1	1	1	1	1	1
4	0100	28	-1	1	0	-2	-1	1	-1	1
5	0112	39	-1	1	0	-2	0	-2	1	1
6	0120	45	-1	1	0	-2	1	1	-1	1
7	0202	40	-1	1	1	1	-1	1	1	1
8	0210	61	-1	1	1	1	0	-2	-1	1
9	0221	73	-1	1	1	1	1	1	0	-2
10	1001	26	0	-2	-1	1	-1	1	0	-2
11	1012	37	0	-2	-1	1	0	-2	1	1
12	1020	42	0	-2	-1	1	1	1	-1	1
13	1102	50	0	-2	0	-2	-1	1	1	1
14	1110	71	0	-2	0	-2	0	-2	-1	1
15	1121	88	0	-2	0	-2	1	1	0	-2
16	1200	21	0	-2	1	1	-1	1	-1	1
17	1211	32	0	-2	1	1	0	-2	0	-2
18	1222	49	0	-2	1	1	1	1	1	1
19	2002	25	1	1	-1	1	-1	1	1	1
20	2010	62	1	1	-1	1	0	-2	-1	1
21	2021	74	1	1	-1	1	1	1	0	-2
22	2100	31	1	1	0	-2	-1	1	-1	1
23	2111	73	1	1	0	-2	0	-2	0	-2
24	2122	90	1	1	0	-2	1	1	1	1
25	2201	41	1	1	1	1	-1	1	0	-2
26	2212	62	1	1	1	1	0	-2	1	1
27	2220	57	1	1	1	1	1	1	-1	1

2.3.1. Analysis of variance (ANOVA)

The main principle behind an analysis of variance (ANOVA) is to partition the total variation in experimental data into different components, using sums of squares, to assess how well a statistical model fits the data. Including all relevant operational variables in the model enables the evaluation of individual statistical significance in influencing the outcome variable. The mapping presented in Table 2.4 was used to fit the models to assess the statistical significance of each variable, as well as to examine the dominance of linear versus quadratic effects. The linear mapping, denoted by 'L' for each factor (e.g., AL, BL), is derived directly from the treatment combination and follows this pattern: 0 is mapped to -1, 1 to 0, and 2 to 1. For instance, in case 9, where the treatment combination is defined as 0221, the mappings are as follows: AL is mapped to -1, BL and CL are mapped to 1, and DL is mapped to 0. The quadratic mapping, denoted by 'Q' (e.g., AQ, BQ), is then generated by squaring the linear values. However, this method results in identical values for zero, which is not ideal. To address this, we assign a value of -2 to the quadratic components corresponding to a linear value of 0 [39]. Returning to the example of case 9, AQ is calculated by squaring AL, resulting in 1, while BQ and CQ are also 1 after squaring their respective linear components. Finally, DQ is assigned a value of -2 because the corresponding linear component is 0.

Once the mapping is established, a full ANOVA model was fitted, incorporating all factors and their interactions to assess their effect on the response variable. This analysis allows for the identification of significant factors and interactions, setting the stage for a more detailed examination of the individual contributions of linear and quadratic effects.
Subsequently, a separate ANOVA model was used to specifically evaluate the linear and quadratic components of each factor. This approach helped to clarify whether the relationship between the factors and the response variable was predominantly linear or quadratic. By comparing the results from both models, the most influential factors were identified, providing a deeper understanding of their impact on the response variable. A p-value is then employed as a metric to balance the risk between making type 1 errors (false positives) and type 2 errors (false negatives). Here, we opt for a cutoff p-value of 0.05 to determine statistical significance. The selection of this value means that there is a 5% chance of observing the obtained data if the null hypothesis is true (i.e., that the relevance of an operational variable does not affect the chemical yield). The results of the ANOVA are presented in Table 2.5.

The results presented in Table 2.5 indicate that the BPR, molar ratio, and milling time are statistically significant in affecting the yield of the mechanochemical regeneration of $NaBH_4$. Additionally, the confounded two-factor interaction (AB, CD) between BPR/molar ratio and milling time/rotational speed also has statistical relevance. The ranking of these variables is effectively visualized with the Pareto plot in Figure 2.3. In this Pareto plot, the F-value is used to rank the significance of each factor and interaction. The F-value quantifies the ratio of variance explained by a factor relative to the variance not explained by the model. A higher F value indicates a greater impact on the response variable.

Table 2.5: Analysis of variance (ANOVA) for the screening design. Main factors A–D correspond to: A = BPR, B = molar ratio, C = milling time, D = rotational speed. "L" and "Q" denote orthogonal linear and quadratic contrasts, respectively.

Source of variation	Degrees of freedom.	Sum of squares	*p*-value
A (BPR)	2	0.153	0.013
AL	1	0.151	–
AQ	1	0.002	–
B (Molar ratio)	2	0.191	0.007
BL	1	0.062	–
BQ	1	0.129	–
C (Milling time)	2	0.434	0.0009
CL	1	0.417	–
CQ	1	0.017	–
D (Rot. speed)	2	0.023	0.293
DL	1	0.003	–
DQ	1	0.020	–
AB, CD	6	0.270	0.011
AC, BD	6	0.050	0.286
AD, BC	6	0.009	0.874

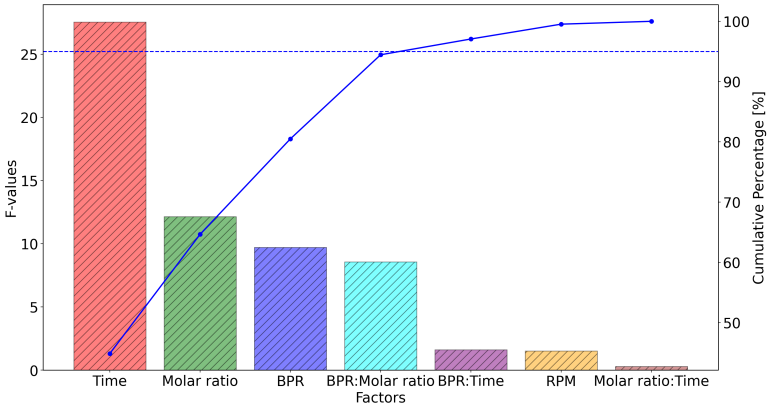


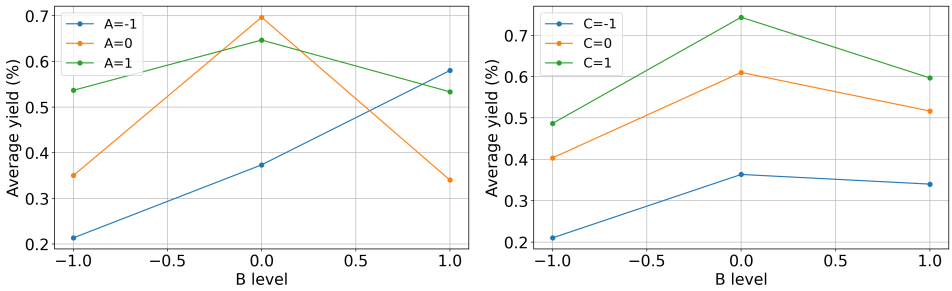
Figure 2.3: Pareto plot of operational variables

The success in the mechanochemical yield has the strongest correlation with the milling time of the process. This is supported by the findings of previous studies as presented in Table 2.2 where it can be seen that high conversion yields necessitate long milling times. Moreover, while a proper individual selection of molar ratio and BPR is also significant for the success of the process, it is worth noting that they are affected by each other. As a consequence of this, their influence on the chemical conversion is leveled. This interaction has the risk of confounding with the two-factor interaction between milling time and rotational speed. However, our results suggest that rotational speed is not statistically significant in this study, and thus, the risk of confounding is minimized. The lack of statistical significance of rotational speed is intriguing since previous studies consistently rely on high rotational speeds (> 1000 rpm) to achieve high yields. Our findings suggest that this high energetic input is unnecessary, as other operational variables will dominate the process. This result is compelling because it can decrease energy requirements and the wear experienced by the milling balls and jar during the process. Specifically, we observed that at higher rotational speeds (> 1000 rpm) in the Emax, the material loss from the milling balls can become so significant that it contaminates the sample. One potential explanation is that the Emax can supply sufficient energy to regenerate $NaBH_4$ even at comparatively low rotational speeds. However, quantifying this is challenging with current state-of-the-art methods, as ball milling machines are often treated as black boxes. This is an area we plan to investigate further in future studies.

The results presented in Table 2.5 also provide a means to quantify linear and quadratic trends among individual factors. This is achieved by comparing the sum of squares between linear and quadratic terms. Specifically, the evolution of ball-to-powder ratio (BPR) and milling time is linearly explained, whereas the molar ratio is explained quadratically. This implies that there exists an optimal point for the molar ratio beyond which the chemical yield will be impacted negatively. This result is supported by the findings of Chen et al. [16] and can be visualized in the interaction plots presented in Figure 2.4. These interaction plots are created by grouping the data based on two

selected factors and plotting the average yield for each combination of their levels. Variables not plotted are averaged out, meaning their effects are integrated into the overall means, allowing the focus to be on the interaction between the plotted factors.
Interestingly, the quadratic dependency of the molar ratio tends to become linear as the second interacting factor decreases in level (A, C = -1). This behavior is observed in both interaction plots, but the slope is considerably different. In the case of the low BPR, the increase in molar ratio results in higher yields, whereas in the case of low milling time, the increase in molar ratio leads to stagnation. The effect of milling time is straightforward to explain, as it is a critical variable in the process: longer milling times allow more opportunity for the quadratic dependency to exert influence, thus increasing yield. In contrast, the behavior with BPR presents a more intriguing phenomenon. The results suggest that the maximum yield shifts as BPR changes. Although the static nature of the selected levels prevents precise identification of these maxima, the trend indicates that a lower BPR requires a higher molar ratio to achieve maximum yield values. While this could potentially allow for the regeneration of more $NaBH_4$ in the same batch, it also implies an increased waste of material due to the need for a greater excess of MgH_2, which is undesirable.

The other statistically significant confounded interaction (C, D), milling time/ rotational speed, may also be analyzed using the interaction plot shown in Figure 2.5, even if rotational speed is not statistically significant as an individual factor.
As expected, longer milling times result in higher overall conversion yields, as shown by the green curve where C (milling time) = 1 (20 hours). However, the same plot illustrates why rotational speed is not a significant variable for improving yield in this study. At each level of milling time, the highest average yield occurs at different rotational speeds, indicating that other variables are more influential in the process. It is worth highlighting that when shorter milling times are used (C = -1), the average yield increases linearly with higher rotational speeds, likely due to the increased number of collisions. However,



(a) (b)

Figure 2.4: (a) Interaction plot of BPR (A) and Molar ratio (B). (b) Interaction plot of Milling time (C) and Molar ratio (B). BPR levels: -1 = 10, 0 = 30, 1 = 50; Molar ratio levels: -1 = 33%, 0 = 66%, 1 = 100%; Milling time levels: -1 = 5 h, 0 = 12.5 h, 1 = 20 h.

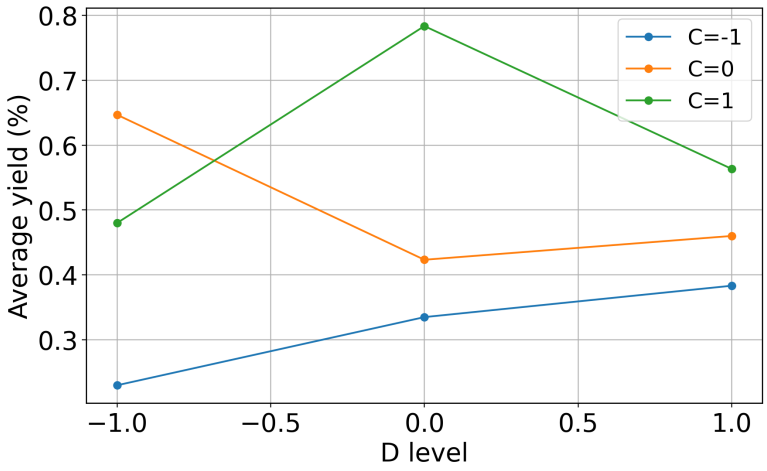


Figure 2.5: Interaction plot of Milling time (C) and Rotational speed (D). Milling time levels: -1 = 5 h, 0 = 12.5 h, 1 = 20 h; Rotational speed levels: -1 = 600 rpm, 0 = 800 rpm. 1 = 1000 rpm.

the average yields for short milling times never reach those achieved with longer milling times (C = 0, 1). For these longer milling times, the average yields intersect, indicating that with sufficient milling time, other variables in the process become more influential and the correlation with rotational speed is lost. It is also notable that high rotational speeds (D = 1) lead to convergence of the average yield regardless of the remaining variables. This does not necessarily mean that maximum performance is achieved under these conditions, but it does help to standardize the results, effectively brute-forcing the conversion. This may explain why many of the results shown in Table 2.2 report high rotational speeds with little care for the remaining operational variables.
Lastly, it is important to highlight that the statistical significance of each individual factor and the corresponding two-factor interactions are subject to the levels used in this study. We have explained the rationale behind the value selection for each level in subsection 2.2.5. Naturally, for the sake of the discussion, had we selected a broader range in rotational speed, its statistical significance in the process could have increased.

2.3.2. Specific case analysis

The results in Table 2.4 show that case 24 achieves the highest conversion yield (90%) in this study. This outcome is obtained with a BPR of 50, an excess molar ratio of 66%, a milling time of 20 hours, and a rotational speed of 1000 rpm. These conditions allow us to match the highest yield reported in the literature while reducing the rotational speed by 20% [16]. While this result is notable for achieving the same yield with lower energy requirements, it is also important to consider if other cases in this study, given the broad range of operational conditions, might be of interest.

The operating conditions of case 15 are particularly attractive as they achieve a regeneration yield of 88% while reducing the BPR by 40% and the rotational speed by an additional 20%. This supports our earlier observation that rotational speed does not statistically affect the conversion yield. Achieving nearly identical results with a reduced BPR demonstrates significant potential for optimizing the process conditions. This is also evident in Figure 2.4(a). While a high BPR consistently yields high conversion rates regardless of the molar ratio, reducing the BPR requires more precise fine-tuning to maintain high yields, thus narrowing the range of high performance. Similar findings are reported in Table 2.2, where Çakanyıldırım et al. [13] achieved a 70% yield with a low BPR of 10. In other words, lenient operating conditions (e.g., high BPR, long milling times, and high rotational speed) facilitate achieving high conversion yields but are generally more costly and less efficient. Conversely, more stringent conditions reduce the range within which the mechanochemical process can obtain high conversion yields, but they offer performance benefits.

The previous observations clearly demonstrate the critical role of milling time in achieving high conversion yields, as both cases 15 and 24 required 20 hours of operation. According to Figure 2.4(b), while there is a noticeable difference in yields between processing times of 12.5 hours and 20 hours, the difference between 12.5 hours and 5 hours is much more significant. This is highlighted by cases 23 and 14, which achieve yields of 73% and 71%, respectively, with only 12.5 hours of milling. Thus, reducing the milling time by 37.5%, from 20 to 12.5 hours, results in only a 17% decrease in yield. This finding is economically significant and suggests that a shorter milling time could be advantageous. A detailed techno-economic analysis could be of interest to further explore this potential benefit.

Finally, cases 1-9, which use a BPR of 10, generally perform poorly, yielding low conversion rates that are not appealing. This performance only improves with a significant excess of MgH_2, which is not technically or economically attractive since the resulting MgO needs to be managed for a circular fuel cycle. Instead of using lower BPRs to process more material in the same batch at the expense of wasted material, it would be more advantageous to develop a larger machine that can replicate the mechanical conditions of the Emax. We plan to quantify these mechanical conditions in future research.

2.3.3. Iron contamination analysis

Given the highly abrasive environment inside the milling jar, it is relevant to estimate the expected levels of iron contamination resulting from the wear experienced by the milling balls. A straightforward method for estimating contamination levels involves weighing the milling balls before the milling process and after cleaning them. For this, we examine the wear experienced by the balls at rotational speeds of 1000 rpm, and additionally, we implement 1200 rpm to serve as a comparison benchmark. These tests were conducted with a milling time of 20 hours and a BPR of 50, minimizing the amount of powder. Thus, this setup increases the potential percentage of impurities in the powder and reduces damping, resulting in more frequent and energetic collisions.

Therefore, this approach allows us to examine the worst-case scenario.

Table 2.6: Ball wear and resulting iron contamination measured after milling.

Ball size (mm)	Speed (rpm)	Unused mass (g)	Used mass (g)	Weight loss (%)	Wear (g)	Fe contam. (%)
10	1200	97.06	96.73	0.34	0.323	17
10	1000	97.06	96.94	0.12	0.117	6

As shown in Table 2.6, operating within the rotational speed range of 600-1000 rpm allows us to keep contamination levels below 6% even under the most abrasive conditions. If the milling time, BPR, or rotational speed are further reduced, contamination levels are expected to also decrease. This approximation does not account for contamination from the jar; however, we anticipate this to be minimal, as a layer of powder rapidly coats and protects the entire jar surface during the milling process.

2.4. Conclusions

We have conducted a comprehensive study on the importance of operational variables in the mechanochemical regeneration of $NaBH_4$, including ball-to-powder ratio, molar ratio, milling time, and rotational speed. Our results, covering a wide range of these variables, provide valuable insights for optimal selection and prioritization in future developments. Additionally, we introduced an inexpensive and straightforward method to quantify regeneration yield without additional chemicals, allowing the production of a 'ready-to-use' solution for on-demand hydrogen release.

Using a screening design of experiments, we investigated the influence of each operational variable on reaction yield. Our findings show that milling time is the most significant factor, followed by molar ratio and the interaction between molar ratio and ball-to-powder ratio. These results align with current state-of-the-art and offer insights into why previous studies selected specific conditions for high yields. We also found that rotational speed, often set at high values (i.e., >1000 rpm), does not significantly impact conversion yield compared to other variables. However, high rotational speeds homogenize yields, making them more consistent regardless of other conditions. While this doesn't ensure high yields, it results in yields converging toward an average value, which may explain their common use despite less optimization of other variables.

We reproduced the highest conversion yield reported using $NaBO_2 \cdot 4\,H_2O$ and MgH_2, with adjustments including a 20% reduction in rotational speed, offering energy savings. Given the wide range of conditions explored, we identified other attractive scenarios that, while not achieving the highest yields, offer economic advantages by processing more powder per batch or reducing milling time.

The ranking of operational variables holds within our experimental range. For instance, an extremely low rotational speed, like 10 rpm, would significantly reduce yield, despite being statistically insignificant in our analysis. This underscores the need for careful result interpretation and highlights substantial opportunities for optimization and future scale-up.

This work advances understanding of the complex interactions in the mechanochemical regeneration of $NaBH_4$. While we examined many variables, we believe the fill ratio, kept constant in this chapter, merits further exploration due to its potential impact on processing and yield. To address this, the next chapter uses discrete element simulations to characterize the mill's internal dynamics and map them to machine-level conditions (speed, fill ratio, ball size), establishing a mechanical basis for the operating windows identified here.

References

[1] J. M. Bergthorson. 'Recyclable metal fuels for clean and compact zero-carbon power'. In: *Progress in Energy and Combustion Science* 68 (2018), pp. 169–196. ISSN: 0360-1285. DOI: https://doi.org/10.1016/j.pecs.2018.05.001. URL: http://www.sciencedirect.com/science/article/pii/S0360128518300327.

[2] E. van Rheenen, E. Scheffers, J. Zwaginga and K. Visser. 'Hazard Identification of Hydrogen-Based Alternative Fuels Onboard Ships'. In: *Sustainability* 15.24 (2023). ISSN: 2071-1050. DOI: 10.3390/su152416818. URL: https://www.mdpi.com/2071-1050/15/24/16818.

[3] E. van Rheenen, J. Padding, C. Slootweg and K. Visser. 'A review of the potential of hydrogen carriers for zero emission, low signature ship propulsion systems'. In: *International Naval Engineering Conference and Exhibition*. Online: Conference Proceedings of INEC, 22nd Aug. 2022. DOI: 10.24868/10649. URL: https://doi.org/10.24868/10649.

[4] G. W. Crabtree, M. S. Dresselhaus and M. V. Buchanan. 'The Hydrogen Economy'. In: *Physics Today* 57.12 (Dec. 2004), pp. 39–44. ISSN: 0031-9228. DOI: 10.1063/1.1878333. eprint: https://pubs.aip.org/physicstoday/article-pdf/57/12/39/9876773/39_1_online.pdf. URL: https://doi.org/10.1063/1.1878333.

[5] E. Y. Marrero-Alfonso, A. M. Beaird, T. A. Davis and M. A. Matthews. 'Hydrogen Generation from Chemical Hydrides'. In: *Industrial & Engineering Chemistry Research* 48.8 (2009), pp. 3703–3712. DOI: 10.1021/ie8016225. eprint: https://doi.org/10.1021/ie8016225. URL: https://doi.org/10.1021/ie8016225.

[6] H. X. Nunes, D. L. Silva, C. M. Rangel and A. M. F. R. Pinto. 'Rehydrogenation of Sodium Borates to Close the NaBH4-H2 Cycle: A Review'. In: *Energies* 14.12 (2021). ISSN: 1996-1073. DOI: 10.3390/en14123567. URL: https://www.mdpi.com/1996-1073/14/12/3567.

[7] H. N. Abdelhamid. 'A review on hydrogen generation from the hydrolysis of sodium borohydride'. In: *International Journal of Hydrogen Energy* 46.1 (2021), pp. 726–765. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2020.09.186. URL: https://www.sciencedirect.com/science/article/pii/S0360319920336260.

[8] T. Tai, H. Cao, W. Feng, Z. Yin, H. Zhang and G. Zheng. 'High-efficient synthesis of NaBH4 by solid-phase electrolysis process on a core-shell-type cathode'. In: *International Journal of Hydrogen Energy* 51 (2024), pp. 172–183. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2023.11.043. URL: https://www.sciencedirect.com/science/article/pii/S0360319923057269.

[9] J. Urgnani, F. Torres, M. Palumbo and M. Baricco. 'Hydrogen release from solid state NaBH4'. In: *International Journal of Hydrogen Energy* 33.12 (2008). 2nd World Congress of Young Scientists on Hydrogen Energy Systems, pp. 3111–3115. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2008.03.031. URL: https://www.sciencedirect.com/science/article/pii/S0360319908003194.

[10] L. Takacs. 'What Is Unique About Mechanochemical Reactions?' In: *Acta Physica Polonica A* 126 (Oct. 2014), pp. 1040–1043. DOI: 10.12693/APhysPolA.126.1040.

[11] C.-L. Hsueh, C.-H. Liu, B.-H. Chen, C.-Y. Chen, Y.-C. Kuo, K.-J. Hwang and J.-R. Ku. 'Regeneration of spent-NaBH 4 back to NaBH 4 by using high-energy ball milling'. In: *International Journal of Hydrogen Energy - INT J HYDROGEN ENERG* 34 (Feb. 2009), pp. 1717–1725. DOI: 10.1016/j.ijhydene.2008.12.036.

[12] L. Kong, C. Xinyu, H. Jin, J. Wu, H. Du and T. Xiong. 'Mechanochemical Synthesis of Sodium Borohydride by Recycling Sodium Metaborate'. In: *Energy & Fuels* 23 (Sept. 2009). DOI: 10.1021/ef900619y.

[13] C. Cakanyildirim and M. Gürü. 'Processing of NaBH4 from NaBO2 with MgH2 by ball milling and usage as hydrogen carrier'. In: *Renewable Energy* 35 (Sept. 2010), pp. 1895–1899. DOI: 10.1016/j.renene.2010.01.001.

[14] C. Lang, Y. Jia, J. Liu, H. Wang, L. Ouyang, M. Zhu and X. Yao. 'NaBH 4 regeneration from NaBO 2 by high-energy ball milling and its plausible mechanism'. In: *International Journal of Hydrogen Energy* 42 (Apr. 2017). DOI: 10.1016/j.ijhydene.2017.04.014.

[15] L. Ouyang, W. Chen, J. Liu, M. Felderhoff, H. Wang and M. Zhu. 'Enhancing the Regeneration Process of Consumed NaBH 4 for Hydrogen Storage'. In: *Advanced Energy Materials* 7 (June 2017), p. 1700299. DOI: 10.1002/aenm.201700299.

[16] W. Chen, L. Ouyang, J. Liu, X. Yao, H. Wang, Z. Liu and M. Zhu. 'Hydrolysis and regeneration of sodium borohydride (NaBH 4) – A combination of hydrogen production and storage'. In: *Journal of Power Sources* 359 (Aug. 2017), pp. 400–407. DOI: 10.1016/j.jpowsour.2017.05.075.

[17] H. Zhong, L. Ouyang, J. Ye, J. Liu, H. Wang, X. Yao and M. Zhu. 'An one-step approach towards hydrogen production and storage'. In: *Energy Storage Materials* 7 (Mar. 2017). DOI: 10.1016/j.ensm.2017.03.001.

[18] S. G. Nuñez, D. L. Schott and J. T. Padding. 'Predicting the Energy Dissipation in a High-Energy Ball Mill with Discrete Element Modeling'. In: (Sept. 2024). DOI: 10.36227/techrxiv.172565505.54889130/v1. URL: http://dx.doi.org/10.36227/techrxiv.172565505.54889130/v1.

[19] M. Tegel, S. Schöne, B. Kieback and L. Röntzsch. 'An efficient hydrolysis of MgH2-based materials'. In: *International Journal of Hydrogen Energy* 42.4 (2017), pp. 2167–2176. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2016.09.084. URL: https://www.sciencedirect.com/science/article/pii/S0360319916327914.

[20] M.-H. Grosjean and L. Roué. 'Hydrolysis of Mg–salt and MgH2–salt mixtures prepared by ball milling for hydrogen production'. In: *Journal of Alloys and Compounds* 416.1 (2006), pp. 296–302. ISSN: 0925-8388. DOI: https://doi.org/10.1016/j.jallcom.2005.09.008. URL: https://www.sciencedirect.com/science/article/pii/S0925838805014118.

[21] Y. V. Verbovytskyy, V. V. Berezovets, A. R. Kytsya, I. Y. Zavaliy and V. A. Yartys. 'Hydrogen Generation by the Hydrolysis of MgH2'. In: *Materials Science* 56.1 (July 2020), pp. 1–14. ISSN: 1573-885X. DOI: 10.1007/s11003-020-00390-5. URL: https://doi.org/10.1007/s11003-020-00390-5.

[22] C. H. Chao and T. C. Jen. 'Reaction of Magnesium Hydride with Water to Produce Hydrogen'. In: *Applied Mechanics and Materials*. Vol. 302. Trans Tech Publications, Ltd. 2013, pp. 151–157. DOI: 10.4028/www.scientific.net/amm.302.151. URL: https://doi.org/10.4028/www.scientific.net/amm.302.151.

[23] M. M. Kreevoy and R. W. Jacobson. 'The rate of decomposition of NaBH4 in basic aqueous solutions'. In: *Ventron Alembic* 15 (1979), pp. 2–3.

[24] R. Retnamma, A. Q. Novais and C. Rangel. 'Kinetics of hydrolysis of sodium borohydride for hydrogen production in fuel cell applications: A review'. In: *International Journal of Hydrogen Energy* 36.16 (2011). European Fuel Cell 2009, pp. 9772–9790. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2011.04.223. URL: https://www.sciencedirect.com/science/article/pii/S0360319911011220.

[25] Q. Wang, L. F. Zhang and Z. G. Zhao. 'Hydrogen production by sodium borohydride in NaOH aqueous solution'. In: *IOP Conference Series: Materials Science and Engineering* 292.1 (Jan. 2018), p. 012031. DOI: 10.1088/1757-899X/292/1/012031. URL: https://dx.doi.org/10.1088/1757-899X/292/1/012031.

[26] M. S. Yahya and M. Ismail. 'Synergistic catalytic effect of SrTiO 3 and Ni on the hydrogen storage properties of MgH 2'. In: *International Journal of Hydrogen Energy* 43 (Mar. 2018). DOI: 10.1016/j.ijhydene.2018.02.028.

[27] H. Uesugi, T. Sugiyama, I. Nakatsugawa and T. Ito. 'Production of hydrogen storage material MgH2 and its application'. In: *Journal of Japan Institute of Light Metals* 60 (2010), pp. 615–618. URL: https://api.semanticscholar.org/CorpusID:13839396.

[28] M. Ismail, M. S. Yahya, N. Sazelee, N. Ali, F. Yap and N. Mustafa. 'The effect of K2SiF6 on the MgH2 hydrogen storage properties'. In: *Journal of Magnesium and Alloys* 8 (June 2020). DOI: 10.1016/j.jma.2020.04.002.

[29] T.-C. Jen, J. Adeniran, E. Akinlabi, C.-H. Chao, Y.-H. Ho and J. De Koker. 'Hydrogen Generation From Acetic Acid Catalyzed Magnesium Hydride Using an On-Demand Hydrogen Reactor'. In: Nov. 2016, V06AT08A034. DOI: 10.1115/IMECE2016-66459.

[30] J. Huot, G. Liang, S. Boily, A. Van Neste and R. Schulz. 'Structural study and hydrogen sorption kinetics of ball-milled magnesium hydride'. In: *Journal of Alloys and Compounds* 293-295 (1999), pp. 495–500. ISSN: 0925-8388. DOI: https://doi.org/10.1016/S0925-8388(99)00474-0. URL: https://www.sciencedirect.com/science/article/pii/S0925838899004740.

[31] T. Durojaiye, J. Hayes and A. Goudy. 'Rubidium Hydride: An Exceptional Dehydrogenation Catalyst for the Lithium Amide/Magnesium Hydride System'. In: *The Journal of Physical Chemistry C* 117.13 (2013), pp. 6554–6560. DOI: 10.1021/jp400961k. eprint: https://doi.org/10.1021/jp400961k. URL: https://doi.org/10.1021/jp400961k.

[32] D. Khan, J. Zou, S. Muhammad, N. A. Khan, S. Saud and S. Panda. 'The adaptable effect of Ru on hydrogen sorption characteristics of the MgH2 system'. In: *Materials Chemistry and Physics* 301 (2023), p. 127583. ISSN: 0254-0584. DOI: https://doi.org/10.1016/j.matchemphys.2023.127583. URL: https://www.sciencedirect.com/science/article/pii/S0254058423002912.

[33] Y. Zou, M. Nie, Y. Huang, J. Wang and H. Liu. 'Kinetics of NaBH4 hydrolysis on carbon-supported ruthenium catalysts'. In: *International Journal of Hydrogen Energy* 36.19 (2011), pp. 12343–12351. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2011.06.138. URL: https://www.sciencedirect.com/science/article/pii/S0360319911016107.

[34] C. Crisafulli, S. Scirè, M. Salanitri, R. Zito and S. Calamia. 'Hydrogen production through NaBH4 hydrolysis over supported Ru catalysts: An insight on the effect of the support and the ruthenium precursor'. In: *International Journal of Hydrogen Energy* 36.6 (2011). 3rd International Workshop in Hydrogen Energy, pp. 3817–3826. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2010.12.089. URL: https://www.sciencedirect.com/science/article/pii/S0360319910024316.

[35] Y. Zou, Y. Huang, X. Li and H. Liu. 'A durable ruthenium catalyst for the NaBH4 hydrolysis'. In: *International Journal of Hydrogen Energy* 36.7 (2011). Materials Technology: Materials in Clean Power System, pp. 4315–4322. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2011.01.027. URL: https://www.sciencedirect.com/science/article/pii/S0360319911000371.

[36] Y. Yan, R. Helmons, C. Wheeler and D. Schott. 'Optimization of a convex pattern surface for sliding wear reduction based on a definitive screening design and discrete element method'. In: *Powder Technology* 394 (2021), pp. 1094–1110. ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2021.09.041. URL: https://www.sciencedirect.com/science/article/pii/S0032591021008342.

[37] D. C. Woods and S. M. Lewis. ‘Design of Experiments for Screening’. In: *Handbook of Uncertainty Quantification*. Ed. by R. Ghanem, D. Higdon and H. Owhadi. Cham: Springer International Publishing, 2017, pp. 1143–1185. ISBN: 978-3-319-12385-1. DOI: 10.1007/978-3-319-12385-1_33. URL: https://doi.org/10.1007/978-3-319-12385-1_33.

[38] D. C. Montgomery. *Design and Analysis of Experiments*. 9th. John Wiley & Sons, 2017.

[39] A. Dean, D. Voss and D. Draguljić. *Design and Analysis of Experiments*. Springer International Publishing, 2017. DOI: 10.1007/978-3-319-52250-0. URL: http://dx.doi.org/10.1007/978-3-319-52250-0.

3

Linking internal dynamics to machine-level operating conditions.

Having mapped the experimental landscape in chapter 2, we move inside the milling jar to link internal dynamics to machine-level operating conditions via Discrete Element Method (DEM) simulations. We vary rotational speed, fill ratio, and ball size to characterize normal and tangential dissipation across all relevant collisions in the mill, and we define setup-independent mechanical descriptors (mean normal/tangential energy per collision and specific collision frequency). From these descriptors, we construct master curves that reveal distinct operational regimes and define transferable metrics for reproducibility and scale-up. Used in inverse mode, the curves can predict combinations of machine-level operational conditions that reproduce specified internal-dynamics configurations.

Parts of this chapter are adapted from: Garrido Nuñez, S., Schott, D. L. and Padding, J. T. 'Predictive models for energy dissipation in mechanochemical ball milling'. In: Powder Technology 457, 120919 (2025).

3.1. Introduction

High-energy ball milling is a versatile method that harnesses mechanical forces to drive physical and chemical material transformations. In recent years, it has emerged as an attractive technique that can support green chemistry, offering synthesis capabilities without reliance on organic solvents or extreme temperature-pressure conditions [1, 2]. Its application in various domains, such as sodium borohydride ($NaBH_4$) mechanochemical regeneration, showcases the growing preference for this novel mechanical method over traditional chemical routes [3]. It has also opened up the possibility of obtaining stable supramolecular and organic compounds that would otherwise be hard or impossible to obtain with traditional methods [2]. Additionally, it has been shown to facilitate ultra-fine grinding and the amorphization of crystalline materials below glass transition temperatures [4, 5]. Nonetheless, regardless of the application, an important challenge lies in comprehending and predicting the key mechanical dissipation interactions that influence the success of the process and that are fundamental for optimization and up-scaling [6].

Laboratory-scale milling processes are typically carried out in machines known as ball mills. The function and usage of these machines can vary according to the motion they induce, the size and density of the milling balls used to impact the processed material, and the shape of the container where they are placed. For example, shaker ball mills follow a linear left-to-right motion, promoting head-on impacts between the milling balls where normal energy transfer dominates, and are typically used with small samples. On the other hand, planetary ball mills use centrifugal forces by inducing a double-axis rotational motion. These centrifugal forces contribute towards tangential energy transfer and attempt to emulate the working mechanism behind industrial-sized roller mills in which potential energy is exploited, offering a direct path for scaling up [2]. However, these are only two general types of ball mills. As applications diversify and increase in complexity, specialized milling machines tailored to distinct motions and mechanical phenomena emerge. While the use of specialized machinery may offer benefits to individual use cases, it also highlights a pressing challenge: the lack of fundamental understanding regarding the underlying phenomenology makes reproducibility across different devices and scales problematic [1, 2].

Currently, some control on the outcome of a milling process can be offered with the definition of operational parameters, such as rotational speed or fill ratio. However, it has become apparent that when the process's complexity increases, or efficiency and scale-up become relevant, this is no longer sufficient [1]. We hypothesize that distinguishing between normal and tangential energetic contributions in ball milling becomes critical due to their distinct mechanical effects. Normal interactions predominantly induce compression forces, causing direct contact and facilitating material compaction or deformation. In contrast, tangential interactions induce shear forces, resulting in sliding or relative motion between surfaces. Understanding these distinctions is vital as they dictate energy transfer mechanisms, influencing the extent of particle deformation and the resultant effects, such as fragmentation, amorphization, or chemical conversion.

To tackle these limitations, some attempts have been made to characterize milling processes from the point of view of the fundamental mechanics inside the milling jar. Chen et al. [7] make use of DEM (Discrete Element Method) simulations to study the dynamics of a shaker ball mill and quantitatively describe the collision events that take place during a mechanical alloying process. The research focused on the energy dissipation that occurs as a result of ball-ball and ball-wall interactions by analyzing the changes in kinetic energy before and after collision. Although the study provides valuable information, it falls short in distinguishing between the roles of normal and tangential dissipation caused by impact and shearing collisions, which affect how the processed material is treated. Following a similar approach, Broseghini et al. [8] studied how varying the shape of the milling jar affected the efficiency of a high-energy planetary mill. To do this, an upper bound of the energy available for grinding is approximated using the same principles of kinetic energy transformation, with the main difference that the velocity vectors are decomposed into their normal and tangential components to distinguish the contribution of impacts and shearing. While this approach allows a more in-depth analysis of the different mechanical phenomena inside the jar, the approximations assume that the milling balls do not rotate and the contact model used is not capable of taking into account the non-linear elastic contact behavior of ball-ball and ball–wall interactions. Moreover, we hypothesize that assessing the interaction between the ball and the wall by measuring the change in kinetic energy will not accurately reflect the true extent of the impact. This is because the energy transferred from the wall to the ball could increase the ball's kinetic energy, making it challenging to precisely determine the amount of energy dissipated.

In the field of mechanochemistry, Burmeister et al. [9, 10] employed DEM simulations to examine the impact of various stressing conditions in planetary ball mills to obtain Knoevenagel synthesis. To achieve a more accurate representation of the ball-ball and ball-wall interactions, the Hertz-Mindlin model was used. However, their analysis is centered around the dissipation of energy in the normal direction due to head-on collisions, which may be insufficient in cases where shearing has the potential of being a critical component to determine the success of the process [11]. Moreover, the characterization parameters presented are limited to the reactants used and cannot be generalized any further.

Lastly, in the field of ultra-fine milling and particle breakage, Oliveira et al. [12] and Rodriguez et al. [5] have employed a state-of-the-art mechanistic and phenomenological model (UFRJ) to characterize the breakage mechanisms that occur in vertical stir mills and planetary ball mills, respectively. The model has only been validated for normal collisions, and as such, the tangential component is typically ignored. This is a limitation that can underestimate the resulting product of the process, as noted by Beinert et al. [13]. In an attempt to overcome this, Oliveira et al. account for the tangential contribution by matching the model predictions with experimental data via back-fitting. This method enables good agreement between experimental and simulation results, given that the proportion of tangential contribution is accurately

adjusted, but naturally, it is susceptible to over-fitting and possible misinterpretation of results. Moreover, this method is viable given that the process involves only a physical transformation (i.e., particle size reduction). In a process where a chemical transformation is expected, such as mechanochemistry, the back-fitting recursion method becomes inadequate. This limitation arises because the Discrete Element Method (DEM) is unable to simulate chemical processes, thus neglecting the possible role of tangential energy dissipation that could influence the rate of a chemical reaction. In other words, it becomes impossible to match experimental and simulation results because the simulations cannot predict the chemical component of the process.

3.2. Model and methods

3.2.1. Modelling approach

The Discrete Element Method (DEM) is used to simulate the interactions between the jar and the grinding media. In this study, Altair EDEM 2021.2 was used as the DEM solver, and Python 3.9.12 was used for data post-processing. EDEM follows a soft-sphere approach by calculating the contact forces for each particle interaction using Hertz and Mindlin's contact model. Then, Newton's laws of motion are used to calculate the instantaneous motion for each particle:

$$m_i \frac{d\mathbf{V}_i}{dt} = \mathbf{F}_{c,i} + m_i \mathbf{g} \tag{3.1}$$

$$I_i \frac{d\boldsymbol{\omega}_i}{dt} = \boldsymbol{\tau}_i \tag{3.2}$$

where m_i, I_i, $\mathbf{V}_i$, and $\boldsymbol{\omega}_i$ are the mass, moment of inertia, velocity, and angular velocity, respectively, of particle i. $\mathbf{F}_{c,i}$ and $\boldsymbol{\tau}_i$ represent the total contact force and total contact torque (relative to the particle's center of mass), respectively. The total force and torque are determined by summing over all neighbors in contact with particle i.

The Hertz–Mindlin model [14, 15] is selected to calculate the contact force on each pair because it is capable of capturing the non-linear behavior of particle–particle and particle–geometry interactions. The original model was modified to make the damping components accessible for data post-processing. Each discrete element has its own radius R, mass m, Young's modulus Y, shear modulus G, coefficient of restitution e, and Poisson ratio v. The contact force $\mathbf{F}_{c,ij}$ on a particle i due to its interaction with another particle j (or wall) is the vector sum of a normal $\mathbf{F}_{n,ij}$ and tangential $\mathbf{F}_{t,ij}$ force:

$$\mathbf{F}_{c,ij} = \mathbf{F}_{n,ij} + \mathbf{F}_{t,ij} = \left(K_n \boldsymbol{\delta}_{n,ij} - \gamma_n \mathbf{V}_{n,ij}\right) + \left(K_t \boldsymbol{\delta}_{t,ij} - \gamma_t \mathbf{V}_{t,ij}\right) \tag{3.3}$$

with:

$$K_n = \frac{4}{3} Y^* \sqrt{R^* \delta_n} \tag{3.4}$$

$$\gamma_n = -2\sqrt{\frac{5}{6}} \beta \sqrt{S_n m^*} \geq 0 \tag{3.5}$$

$$K_t = 8G^* \sqrt{R^* \delta_n} \tag{3.6}$$

$$\gamma_t = -2\sqrt{\frac{5}{6}} \beta \sqrt{S_t m^*} \geq 0 \tag{3.7}$$

$$S_n = 2Y^* \sqrt{R^* \delta_n} \tag{3.8}$$

$$S_t = 8G^* \sqrt{R^* \delta_n} \tag{3.9}$$

$$\beta = \frac{\ln(e)}{\sqrt{\ln^2(e) + \pi^2}} \tag{3.10}$$

$$\frac{1}{Y^*} = \frac{(1 - v_1^2)}{Y_1} + \frac{(1 - v_2^2)}{Y_2} \tag{3.11}$$

$$\frac{1}{G^*} = \frac{2(2 - v_1)(1 + v_1)}{Y_1} + \frac{2(2 - v_2)(1 + v_2)}{Y_2} \tag{3.12}$$

$$\frac{1}{R^*} = \frac{1}{R_1} + \frac{1}{R_2} \tag{3.13}$$

$$\frac{1}{m^*} = \frac{1}{m_1} + \frac{1}{m_2} \tag{3.14}$$

Here, $\mathbf{V}_{n,ij}$ and $\mathbf{V}_{t,ij}$ are the relative normal and tangential velocities between the two particles i and j at the point of contact. $\boldsymbol{\delta}_{n,ij}$ and $\boldsymbol{\delta}_{t,ij}$ represent the normal and tangential overlap vectors between the particles, the latter being found by integrating the relative tangential velocity with time and projecting it on the current tangential direction. K_n and K_t are the elastic coefficients for normal and tangential contact. Furthermore, γ_n and γ_t represent the viscoelastic damping coefficients for the normal and tangential contacts. On the right-hand side of Equation 6.3, the first term between parentheses is the normal force, and the second term is the tangential force. The normal force has two terms, a spring force and a normal damping force $\mathbf{F}_{n,d}$. The tangential force also has two terms, a shear force and a tangential damping force $\mathbf{F}_{t,d}$.

Additionally, the contact torque $\boldsymbol{\tau}_{ij}$ on particle i due to its interaction with particle (or wall element) j is calculated by cross-multiplying the vector $\mathbf{R}_{ij}$, which points from the center of mass of particle i to the contact point with particle j, by the tangential contact force $\mathrm{F}_{t,ij}$. Since the particles undergo constant rolling motion, especially against the wall, it is also necessary to account for slight non-sphericity with a rolling torque $\boldsymbol{\tau}_{r,ij}$ calculated by the coefficient of rolling friction μ_r, magnitude of the normal contact force $F_{n,ij}$, the distance from the center of mass to the contact point R_{ij} and the orientation of the angular velocity vector $\boldsymbol{\omega}_{rel}$ of the particle relative to the particle (or wall) it is in

contact with:

$$\boldsymbol{\tau}_{ij} = \mathbf{R}_{ij} \times \mathbf{F}_{t,ij} + \boldsymbol{\tau}_{r,ij} \tag{3.15}$$

$$\boldsymbol{\tau}_{r,ij} = -\mu_r F_{n,ij} R_{ij} \frac{\boldsymbol{\omega}_{rel}}{\omega_{rel}} \tag{3.16}$$

Lastly, the amount of dissipated energy in a time interval t_1 to t_2, attributed to the damping components γ_n and γ_t, can be calculated as follows:

$$E_n = \int_{t_1}^{t_2} \mathbf{F}_{n,d} \cdot \mathbf{V}_{n,ij} \mathrm{d}t = \int_{t_1}^{t_2} \gamma_n V_{n,ij}^2 \mathrm{d}t \tag{3.17}$$

$$E_t = \int_{t_1}^{t_2} \mathbf{F}_{t,d} \cdot \mathbf{V}_{t,ij} \mathrm{d}t = \int_{t_1}^{t_2} \gamma_t V_{t,ij}^2 \mathrm{d}t \tag{3.18}$$

3.2.2. Simulation setup and calibration

In this work, we model a commercially available ball mill. In particular, the Emax high-energy ball mill is a device produced and distributed by the German company Retsch. It offers a novel approach to ball milling by combining high friction and impact results with a temperature control system, allowing for controlled grinding. The machine can allocate proprietary grinding jars with 125 ml of volume that follow a circular motion with a rotational speed n up to 2000 revolutions per minute (corresponding to an angular frequency of $\omega = 2\pi n/60 = 209$ rad/s) with an amplitude (radius) A of 1.7 centimeters, see Fig. 6.1. The movement of the jar has been replicated in our simulations. To accurately represent the geometry of the milling jar, a CAD file was imported into EDEM and an automatic rigid body mesh was built using the settings in Table 3.1.
The system is initialized by generating the total amount of discrete media over a span of five seconds, allowing them to reach resting positions inside the jar before any movement is induced. The simulation is then run for an additional 15 seconds to observe the system's dynamics. To minimize the computational complexity of the model, the number of discrete elements has been limited to only include the milling balls. This approach is viable as the influence of the processed material can be represented by altering the friction and restitution coefficients [16–18].

As a result, two sets of coefficients are used. The first set corresponds to a clean, empty jar without powder, serving as a reference for the unaltered dynamics of

Table 3.1: Mesh-generation parameters used in the simulation.

Parameter	Value
Minimum mesh scaling factor	0.33
Maximum mesh scaling factor	4
Maximum deviation scaling factor	1
Maximum angle	0.261 rad

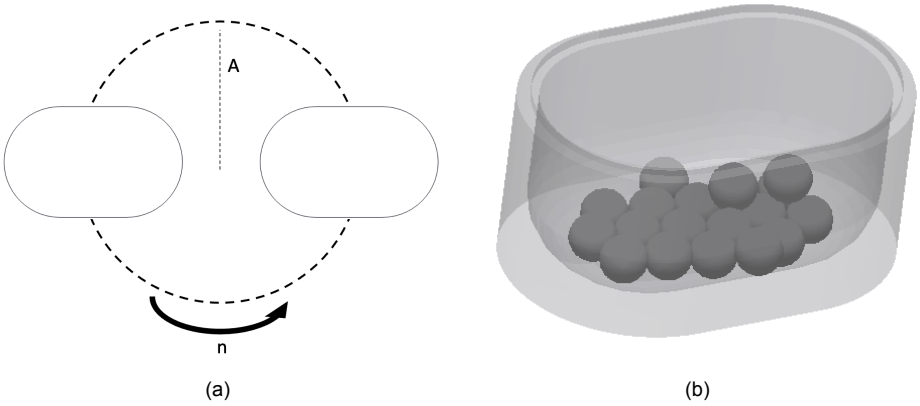


Figure 3.1: (a) Schematic of jar movement (b) 3D model of the jar.

the milling balls. The second set was obtained through calibration, considering the presence of hydrated sodium metaborate ($NaBO_2 \cdot 4\,H_2O$) and magnesium hydride (MgH_2). The calibration methodology combines the approaches of Dreizin et al. [17] and Burmeister et al. [10]. For an accurate calibration, it is essential to have a representative powder layer covering the jar and milling balls. To achieve this, we selected one of the experimental conditions from our previous work, where we explored the effects of various experimental parameters on the mechanochemical regeneration of $NaBH_4$ [19]. The selected conditions involve a ball-to-powder ratio of 30, a 66% molar ratio excess, a milling time of 12.5 hours, a fill ratio of 10%, and a rotational speed of 600 rpm, ultimately enabling a regeneration yield of 71%.
After conducting this experiment, the resulting powder layer inside the jar provides the most accurate representation of how the powder's presence affects the balls and jar, enabling a precise calibration process.

The methodology for determining the restitution coefficient is based on free-fall experiments. Milling balls were removed from the jar, and for each test, a ball was randomly selected and dropped onto the flat bottom surface of the milling jar. The restitution coefficient was then calculated as the ratio between the drop height and the bounce height, as shown in Equation 3.19. A total of 32 tests were conducted to ensure statistical significance, and the average value of 0.3 was chosen as the calibrated restitution coefficient.

$$e = \sqrt{\frac{h_{bounce}}{h_{drop}}} \tag{3.19}$$

The calibration of the friction coefficients involves the combination of a new experiment and its computational representation using the previously defined restitution coefficient

as a constant. In this experiment, the representative powder layer is retained, but instead of the flat surface at the bottom of the jar, the walls are used. The jar is positioned vertically, and a milling ball is placed at the midpoint of the jar's straight wall. The ball is then released, and a custom tracking code records its position as it oscillates along the walls until it comes to rest. An equivalent DEM model of this experiment is created, and the friction coefficients are adjusted until the simulated ball position aligns with the experimental curve. See Figure 3.2 and Figure 3.3 for the calibration results.

The comparison between the experimental and computational curves indicates that a static friction coefficient of 0.3 and a rolling friction coefficient of 0.045 provide a good fit. To further illustrate the sensitivity of the ball's motion to changes in these

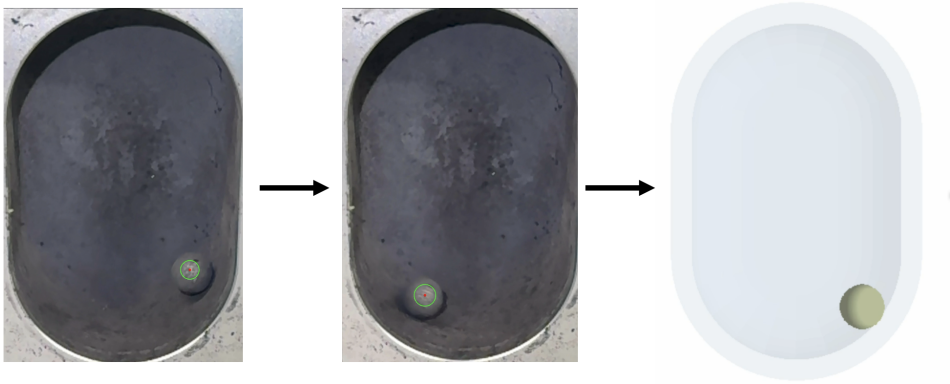


Figure 3.2: Frame of video and simulation used for the calibration of the friction coefficients.

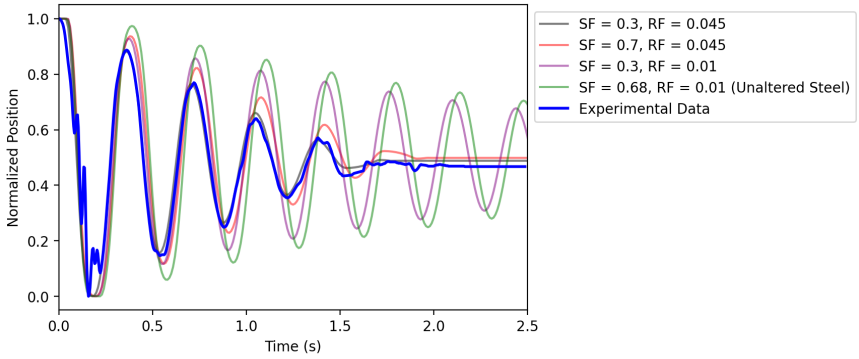


Figure 3.3: Normalized ball position as a function of time for calibration purposes. SF: static friction; RF: rolling friction. The bumps during the initial experimental oscillation of the ball are due to inaccuracies in ball detection during initial frames.

coefficients, we have included additional values in Figure 3.3. Notably, the rolling friction significantly affects the number of oscillations the ball undergoes before reaching its resting state, while the static friction primarily influences the height of each oscillation and the time it takes for the ball to reach its inflection points. Ultimately, both coefficients must be fine-tuned together. This finding aligns with the work of Dreizin et al. [17], but contrasts with Burmeister et al. [10], who suggest that the static friction coefficient has no significant impact. The final calibrated values and simulation settings used in the simulations can be found in Table 3.2 and Table 3.3.

Relevant dependent variables are derived from the normal and tangential damping forces shown in Equation 6.3. The normal interactions account for head-on collisions, while the tangential interactions account for glancing collisions. The benefit of extracting these forces directly from the Hertz-Mindlin model is that it becomes possible to access instantaneous values for both normal and tangential interactions at any time and not rely on averages derived from power calculations. This allows for a more precise representation of the evolution of the energy dissipation given by Equation 6.17 and Equation 6.18. Thus, it is possible to calculate the mean normal and tangential energy dissipation per collision by summing all individual dissipation events and dividing by the number of collisions in a given time.

Table 3.2: Material parameters for the milling balls and vial walls. "Standard" values correspond to clean X46Cr13 steel; the "calibrated" set accounts for the presence of $NaBO_2 \cdot 4\,H_2O$ and MgH_2. *Young's modulus was deliberately reduced to shorten simulation runtime (see Subsection 3.2.3).

Parameter	Standard value	Calibrated value
Restitution coefficient	0.68[20]	0.30
Static friction coefficient	0.70[21]	0.30
Rolling friction coefficient	0.01[20]	0.045
Density ($kg\,m^{-3}$)	7700[22]	
Young's modulus* (GPa)	205[22]	
Poisson's ratio	0.235[22]	

Table 3.3: Solver parameters used in the DEM simulations.

Parameter	Value
Simulation time step	9.5×10^{-7} s
Total simulation time	20 s
Time-integration method	Euler

$$\overline{E_n} = \frac{\sum E_n}{N_{collisions}} \tag{3.20}$$

$$\overline{E_t} = \frac{\sum E_t}{N_{collisions}} \tag{3.21}$$

Then, the normal and tangential dissipation power can be calculated by multiplying by the collision frequency f_{col} observed within the system:

$$P_n = f_{col}\overline{E_n} \tag{3.22}$$

$$P_t = f_{col}\overline{E_t} \tag{3.23}$$

By adding Equation 3.22 and Equation 3.23 together, the total dissipated power available to the processed material can be calculated. This value could then be compared to the total power of the machine to find an initial estimate of the efficiency of the process. Unfortunately, the Emax is not equipped with a torque sensor, so it needs to be estimated from the simulation. In EDEM, each geometry is modeled as a series of triangles, similar to a mesh. Therefore, the total torque on the geometry $\mathbf{T}$ around the center of rotation $\mathbf{r}_c$ can be calculated as follows:

$$\mathbf{T} = \sum_a \left(\mathbf{r}_a - \mathbf{r}_c\right) \times \mathbf{F}_a + \mathbf{T}_a \tag{3.24}$$

where $\mathbf{F}_a$ is the total force on triangle a, $\mathbf{r}_a$ the triangle's center of mass, and $\mathbf{T}_a$ the torque on triangle a. With this, it becomes possible to calculate the total power P_{total} to drive the system:

$$P_{total} = \omega\overline{T_z} \tag{3.25}$$

where $\omega = 2\pi n/60$ is the angular frequency with which the system is driven (n is the shaker frequency in rotations per minute), realizing the rotations of the jar only take place around the z-axis

The granular temperature (Θ) is an important parameter for the kinetic and hydrodynamic characterization of dynamic granular systems [23]. It is a quantity that measures the variance in the distribution of particle velocities, analogous to the fluctuations in molecular velocity distributions which are measured by the thermodynamic temperature [24]:

$$\Theta = \frac{1}{3}\left(\left\langle \mathbf{V}^2 \right\rangle - \langle \mathbf{V} \rangle^2\right) \tag{3.26}$$

where $\mathbf{V}$ represents the velocity vector of a particle, and pointy brackets $\langle \cdots \rangle$ indicate an average over all particles in a mesh cell dividing the geometry of the milling jar. In this paper, the granular temperature is used to identify high-energy collision zones within the milling jar. By doing this, it becomes possible to identify different collision regimes as the filling of the jar and the shaking speed change.

3.2.3. Evaluation of the effect of lowering Young's modulus

Using realistic values for the Young's modulus of steel necessitates using extremely small integration time steps, potentially making the simulations computationally very expensive. A way to tackle this limitation is to artificially lower Young's modulus as shown by Lommen et al. [25]. This is because reducing the stiffness weakens contact forces between particles and allows them to deform more, leading to larger overlaps under the same applied force. This enables particles to change their velocities more gently upon impact.
Lowering the Young's modulus should be done carefully to find a best-performance value where the accuracy of the simulation is preserved while allowing for shorter simulation times. Since the Young's modulus has a direct influence on the damping and elastic components of both tangential and normal forces, as shown in Equation 6.3, it influences the energy dissipation per collision. Thus, a sensitivity analysis with 19 discrete media with diameter d_b = 10mm, shown in Figure 3.4, was carried out. The total power exerted onto the system, based on Equation 4.5 has been measured for different values of Young's modulus, as this allows an analysis of the influence on both the elastic and damping components. By doing this, relevant particle properties, such as velocity and momentum, are also included in the decision-making. By decreasing the Young's modulus by a factor of 100 (from 205 GPa to 2.05 GPa), we can achieve a threefold reduction in computation time, while remaining within a range of two standard deviations from the mean power predicted for the realistic value. Thus, for all the following simulations, we set Young's modulus to 2.05 GPa.

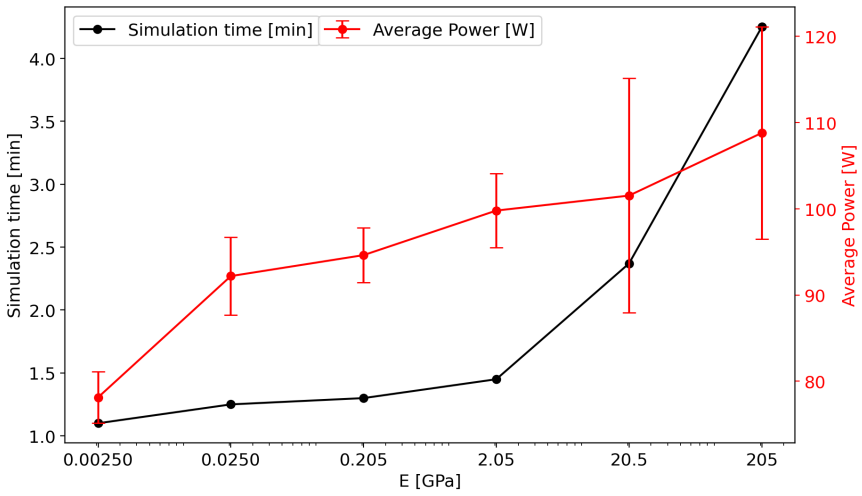


Figure 3.4: Young's modulus sensitivity analysis. Red line indicates the simulation's average power, and the black line indicates the simulation time. Whiskers show two standard deviations from the corresponding means. Ball size = 10mm, rotational speed = 1200 rpm.

3.2.4. Simulation variable parameters

The relevant operational parameters for this study are the rotational speed n, size of the milling balls d_b, and the fill ratio f_r, which is defined as the volumetric ratio between the total of all milling balls and that of the empty milling jar. Note that the fill ratio differs from the ratio of the stagnant layer thickness to the jar height because of the void space in a packed bed of monodisperse particles. The fill ratios used in the simulations are presented in Table 3.4. The rotational speed is varied from 600 to 1200 rpm.

3.2.5. Chemicals

Hydrated sodium metaborate ($NaBO_2 \cdot 4\,H_2O$) (≥ 99%) was obtained from Sigma-Aldrich, while magnesium hydride (MgH_2) (≥ 99.9%, ≤ 50μm) was sourced from Nanoshel. All reagents were used without further purification. The reaction between these reactants facilitates the following conversion:

$$NaBO_2 \cdot 4\,H_2O + 6\,MgH_2 \rightarrow NaBH_4 + 6\,MgO + 8\,H_2 \tag{3.27}$$

The sample preparation for ball milling was performed in a glove box under an argon atmosphere, with oxygen and water concentrations maintained below 0.1 ppm.

3.3. Results and Discussion

In this section, the performance of the Emax is evaluated in terms of energy and power dissipation in collision events. The motivation for focusing on these parameters is that they represent how kinetic energy is being transformed into usable energy for the mechanochemical reaction to take place [1]. First, the performance is assessed by varying the diameter of the milling balls from 5 mm to 10 mm, and altering the rotational speed from 600 rpm to 1200 rpm while maintaining a consistent fill ratio of 8%. Then, the fill ratio is varied following Table 3.4, and two rotational speeds are tested. With the

Table 3.4: Ball counts required to reach a given fill ratio f_r in a 125 mL jar. Two ball diameters are considered: d_b = 10 mm and d_b = 5 mm.

Fill ratio f_r	# balls (d_b = 10 mm)	# balls (d_b = 5 mm)
0.04	10	76
0.06	14	115
0.08	19	153
0.10	24	191
0.12	29	229
0.16	38	306
0.20	48	382
0.30	72	573
0.40	96	764

obtained results, master curves are built, which allow for generalization of the results in terms of the grinding media's density, ball diameter, rotational speed, amplitude (radius) of oscillation, collision frequency, number of balls, and powders used.

3.3.1. Ball size and rotational speed

We start with an analysis of the performance in energy dissipation as the ball size and rotational speed change while maintaining a constant fill ratio of 8%. Figure 3.5 presents the probability distributions of normal and tangential energy dissipation per cycle for different rotational speeds (600, 800, 1000, and 1200 rpm) and two milling ball diameters: 10 mm and 5 mm. The comparison includes unaltered steel conditions and calibrated conditions, where friction coefficients (static and rolling) and the restitution coefficient have been adjusted.

As expected, higher rotational speeds lead to greater energy dissipation due to the increased kinetic energy available during each cycle, which is reflected in the peak shift from lower values at 600 rpm to higher values at 1200 rpm. A key observation across the plots is that tangential energy dissipation increases at a faster rate compared to normal dissipation, as the rotational speed increases. This is easier to visualize in Figure 3.6 where the mean dissipation values are presented.

The difference between calibrated and non-calibrated values is reflected in a shift of peak dissipation values. In the case of 10 mm balls, this shift is guided towards smaller values, whereas in the case of 5 mm balls, the shift is much more subtle and is guided towards larger values. Specifically, this shift causes the mean tangential dissipation to no longer exhibit a significant difference between 5 mm and 10 mm balls, as it did by 13% with the non-calibrated conditions. In contrast, a difference of 20 % can now be identified for the mean normal dissipation.

This result indicates that the calibration process is much more critical as the size of the milling balls and the rotational speed increase. Additionally, the standard deviation and coefficient of variation decrease in both directions when smaller milling balls are used, as shown in Table 3.5. This suggests that the use of smaller milling balls can lead to a more consistent and predictable process, potentially increasing efficiency, if the required energy input for a given mechanochemical reaction is known.

3.3.2. Fill ratio

The analysis of the fill ratio is crucial as it determines the number of grinding elements involved in the mechanochemical process, impacting both the collision frequency and the milling balls' dynamic behavior. As a consequence, it defines operational ranges that enhance either normal or tangential dissipation. Figure 3.7 and Figure 3.8 show the mean energy dissipation in the normal and tangential directions, as the fill ratio is varied according to Table 3.4, for unaltered and calibrated conditions, respectively.
For both conditions, it is useful to define three operational zones. The first zone covers a fill ratio from 4% to 10%, the second extends from 10% to 20%, and the third goes

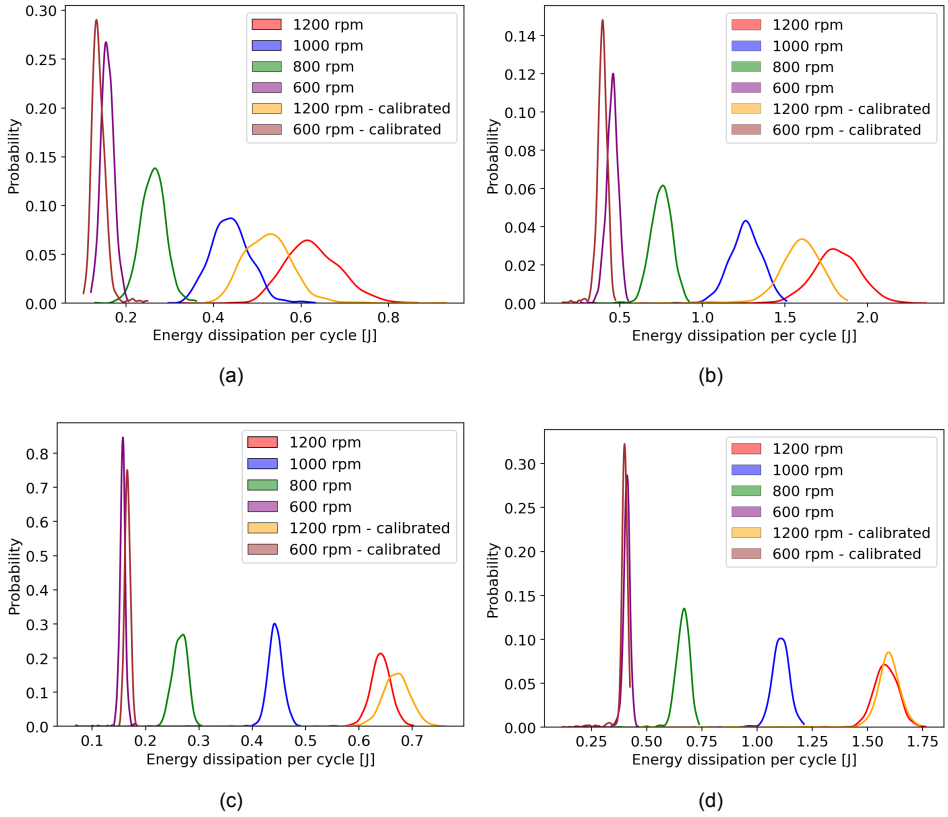


Figure 3.5: (a) Normal energy dissipation per cycle, d = 10 [mm] (b) Tangential energy dissipation per cycle, d = 10 [mm] (c) Normal energy dissipation per cycle, d = 5 [mm] (d) Tangential energy dissipation per cycle, d = 5 [mm]. Fill ratio = 8% for all cases.

Table 3.5: Descriptive statistics for energy dissipated in a single cycle. Two ball diameters are considered: d_b = 5 mm and d_b = 10 mm.

	Normal energy dissipation				Tangential energy dissipation			
	Std. dev.		Coeff. of variation		Std. dev.		Coeff. of variation	
Speed (rpm)	5 mm	10 mm	5 mm	10 mm	5 mm	10 mm	5 mm	10 mm
600	0.006	0.015	3.81%	9.27%	0.019	0.035	4.70%	7.76%
800	0.014	0.029	5.45%	10.95%	0.033	0.065	5.03%	8.64%
1000	0.015	0.046	3.50%	10.36%	0.047	0.098	4.22%	7.77%
1200	0.023	0.063	3.53%	10.11%	0.064	0.145	4.06%	8.02%

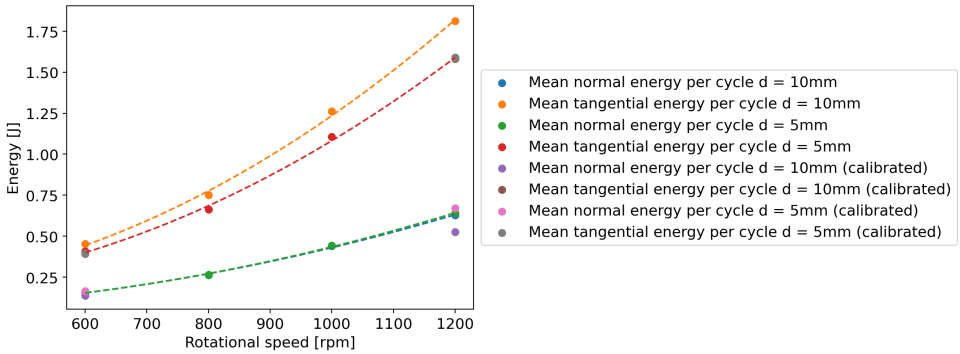


Figure 3.6: Mean energy dissipation per cycle. Note that the blue and green curves virtually overlap.

from 20% to 40%. In terms of normal dissipation, the optimal operating range falls within the second zone, where the majority of the maximum normal dissipation values are observed. However, clear differences emerge between unaltered and calibrated coefficients, as well as between small and large milling balls.

Regarding normal dissipation, the most noticeable difference occurs in the transition from zone 1 to zone 2. For large balls, an increase of up to 88% is estimated under calibrated conditions, while unaltered conditions only show an increase of 31%. Interestingly, when using small balls, unaltered conditions display a clear decrease of up to 67%, whereas calibrated conditions remain nearly stable, except at a fill ratio of 4%.

Tangential energy dissipation, on the other hand, exhibits a different behavior. The differences between large and small balls are subtler, with the primary change occurring in how dissipation evolves from zone 1 to zone 2. These results are consistent with the observations in Figure 3.6, where a shift from large to small balls led to differences in normal dissipation, but not in tangential dissipation. Under unaltered conditions, tangential dissipation decreases by as much as 150%, making zone 1 optimal for this type of dissipation. Conversely, calibrated conditions show a 42% increase, suggesting that zone 2 remains optimal for tangential dissipation as well.
Finally, when considering collision frequency, both unaltered and calibrated conditions exhibit growth as the number of balls increases, which is expected. However, the rate of growth is notably slower under calibrated conditions. This discrepancy can be explained by the differing behavior of the specific collision frequency in the two scenarios. Under calibrated conditions, the specific collision frequency decreases as the fill ratio increases. This is due to the increased damping, which enhances energy dissipation and reduces the relative velocities between balls. Consequently, collisions per ball occur less frequently as the system becomes denser. In contrast, under unaltered conditions, the lack of realistic damping allows for larger relative velocities, which increase the probability of collisions per ball as the jar

becomes more crowded. This upward trend in specific collision frequency amplifies the total collision frequency, resulting in a higher growth rate for unaltered conditions.

Regarding zone 3, it becomes evident that it does not provide any advantages over zones 1 or 2 in terms of energy dissipation. Moreover, the specific collision frequency plateaus in both unaltered and calibrated conditions, indicating the presence of choking behavior. Thus, zone 3 and beyond can be disregarded as zones of interest. The previously discussed results can be visualized by comparing the heat maps in Figure 3.9 and Figure 3.10. These maps depict the relative likelihood of finding particles in specific locations in the jar throughout the simulation, offering insight into how the system's dynamics evolve as different fill ratios give rise to the distinct operational zones. With a 4% fill ratio and calibrated coefficients, the combination of fewer balls and increased damping causes the balls to traverse the jar walls with minimal collisions among them, especially when compared to the uncalibrated scenario. This visually explains the significant increase in normal energy dissipation as more balls are added. Furthermore, it clarifies why the tangential dissipation remains almost constant for the calibrated values: most of the tangential dissipation arises from balls scraping along the jar walls. Since the balls maintain contact with the wall throughout the process, tangential dissipation reaches near-maximum values even at low fill ratios.
When examining the heat maps for the 10% fill ratio, the key difference is that with calibrated coefficients, collisions still do not occur in the center of the jar, unlike in the uncalibrated case. This visual representation also explains why the maximum normal dissipation for calibrated coefficients is achieved by further increasing the fill ratio to around 18%, as opposed to the 10% observed with uncalibrated coefficients. Since most of the normal dissipation results from head-on collisions between balls, allowing them to move slightly further toward the center of the jar leads to maximum dissipation. Lastly, the heat map for the 40% fill ratio shows, in both scenarios, why zone 3 is unfavorable. In both cases, there is a clogging effect, where the balls are unable to achieve high-energy collisions anywhere in the jar.

3.3.3. Master curves

The prior analysis provides a clear description of how the dynamics of the milling balls change as powder is introduced into the milling jar for a mechanochemical reaction. However, two limitations remain. First, while the dissipated energy in each collision is crucial for driving the reaction, it is also important to consider the dissipated power, which accounts for the rate of collisions that generate this energy. Ideally, maximizing power dissipation would combine a high frequency of collisions with the maximum amount of energy dissipated.
The second limitation relates to the fact that these results are, so far, specific to the operating conditions used in the simulations for this work. To address both issues simultaneously, we introduce master curves. These curves are designed to facilitate prediction capabilities when variables such as fill ratio, collision frequency, number of balls, ball density, rotational speed, ball diameter, and amplitude of rotation are varied. This approach offers significant flexibility for both characterization and iterative analysis. The master curves are constructed using the normalization formulas presented in

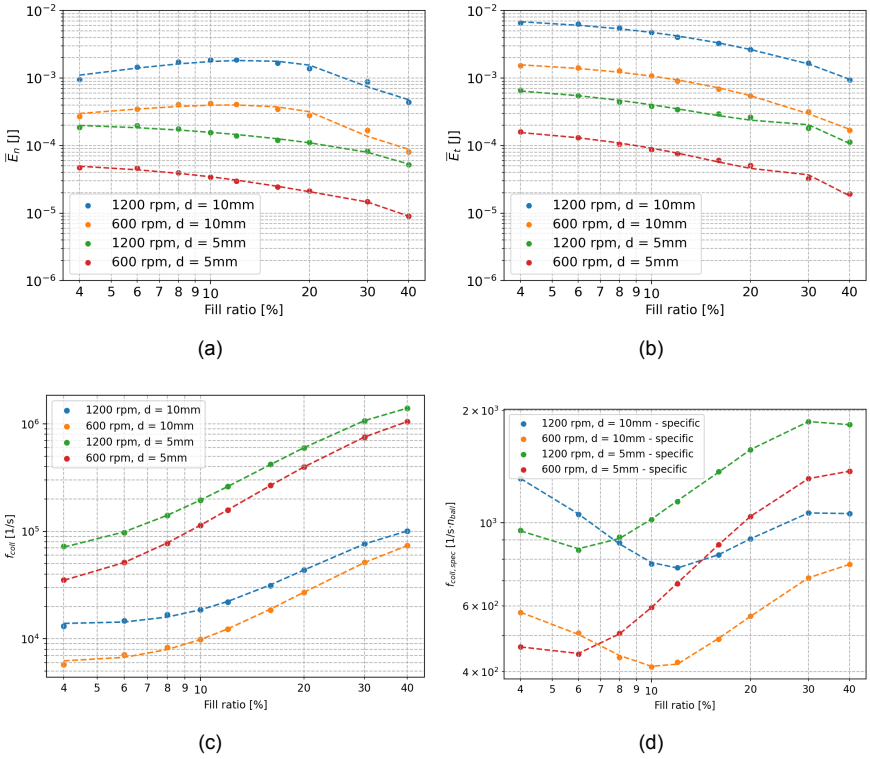


Figure 3.7: (a) Mean normal energy per collision $\overline{E}_n$ (b) Mean tangential energy per collision $\overline{E}_t$ (c) Collision frequency f_{coll} (d) Specific collision frequency $f_{coll,spec}$. Cases with fill ratio variation and unaltered steel coefficients.

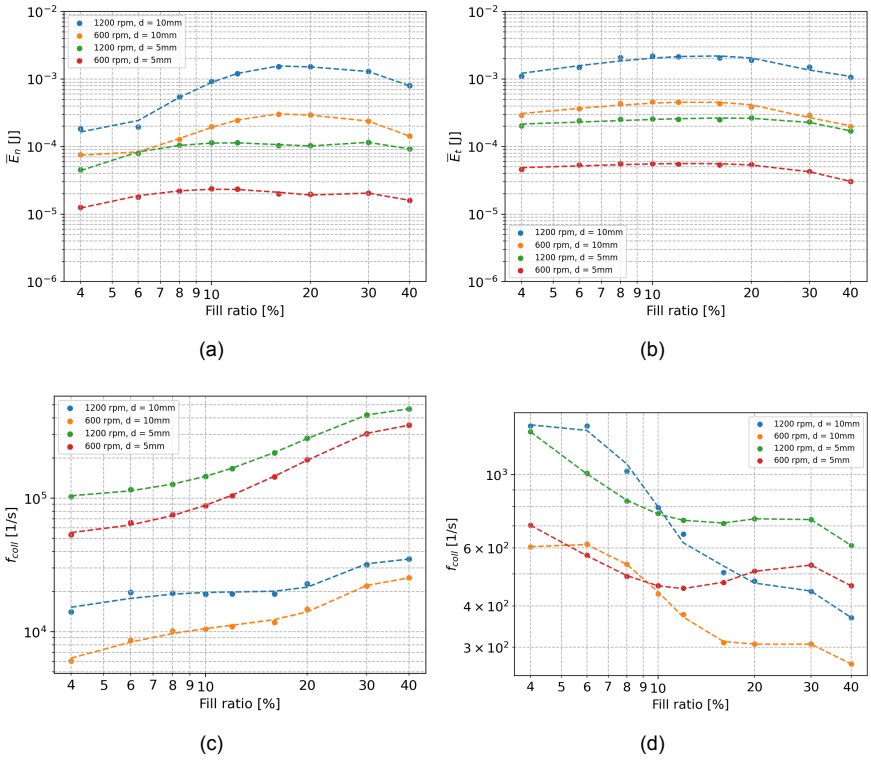


Figure 3.8: (a) Mean normal energy per collision $\overline{E}_n$ (b) Mean tangential energy per collision $\overline{E}_t$ (c) Collision frequency f_{coll} (d) Specific collision frequency $f_{coll,spec}$. Cases with fill ratio variation and calibrated coefficients (system comprising $NaBO_2 \cdot 4\,H_2O$ and MgH_2).

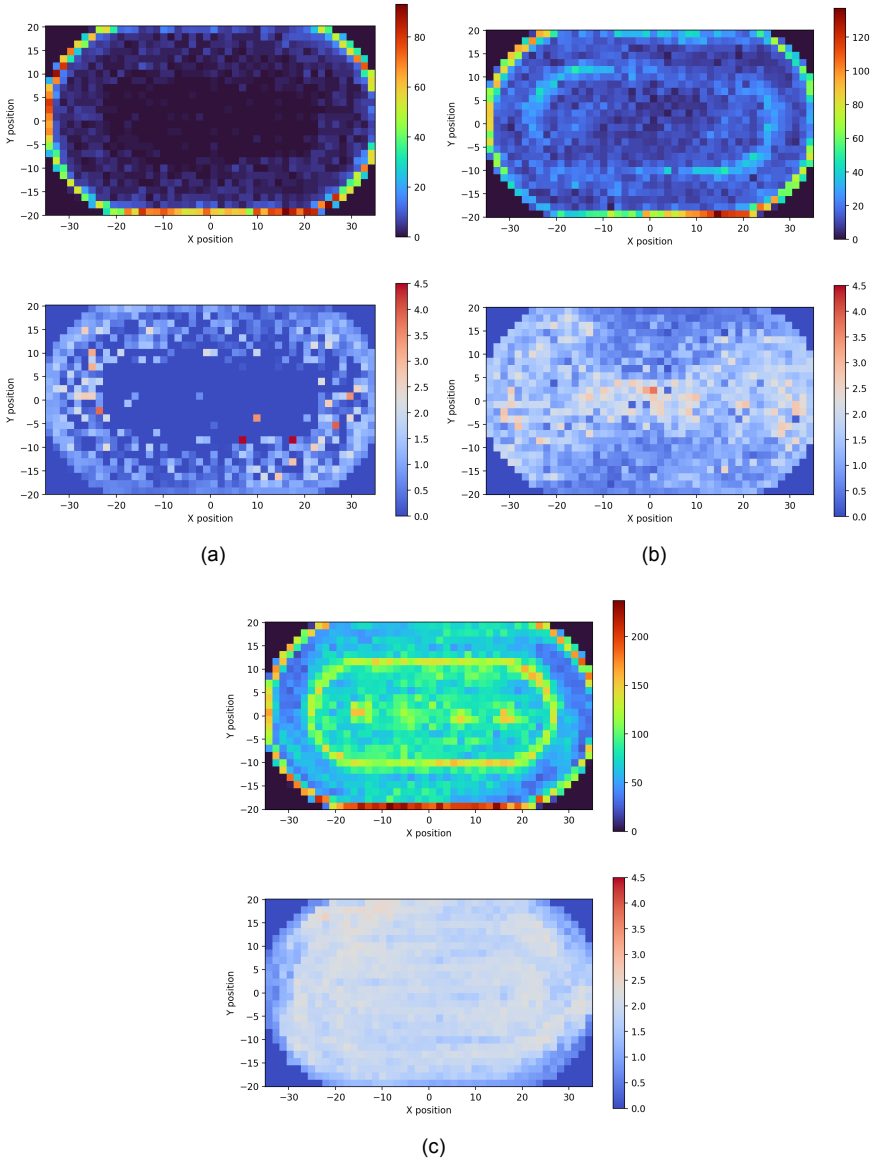


Figure 3.9: Particle heat map (blue-green-red) and granular temperature (blue-white-red). n = 1200 [rpm] and d = 10 [mm]. Fill ratio: (a) 4% (b) 10% (c) 40%. Unaltered steel values.

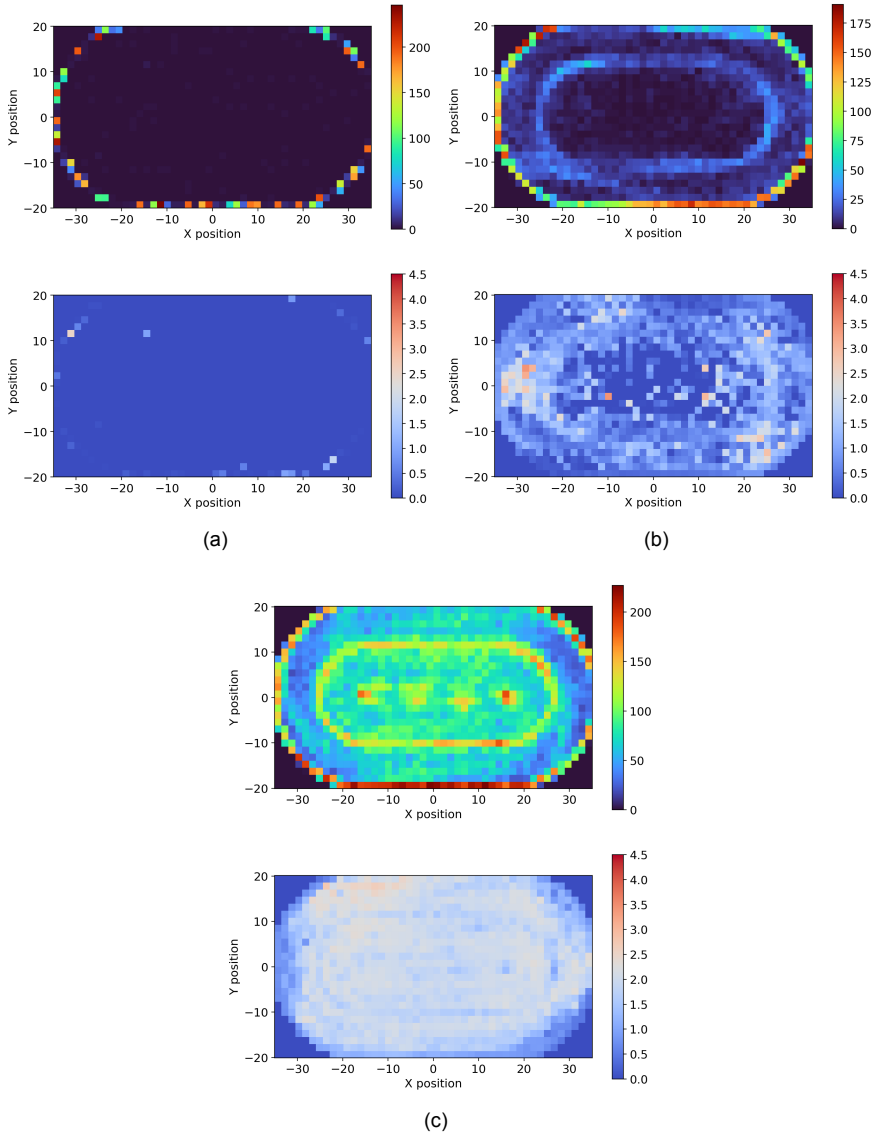


Figure 3.10: Particle heat map (blue-green-red) and granular temperature (blue-white-red). n = 1200 [rpm] and d = 10 [mm]. Fill ratio: (a) 4% (b) 10% (c) 40%. Calibrated values.

Equation 4.3 and Equation 4.4, which are based on basic mechanical principles. As shown in Figure 3.11 and Figure 3.12, these curves provide the mean specific power dissipation, along with the respective standard deviation, for both normal and tangential components. From there, mean energy dissipation values can be obtained by using the corresponding mean specific collision frequency curve.

$$\tilde{P}_{spec} = \frac{f_{col}\bar{E}}{\rho\omega^3 A^2 d_p^3 n_{ball}} \tag{3.28}$$

$$\tilde{f}_{coll} = \frac{f_{col}}{\omega n_{ball}} \tag{3.29}$$

Where f_{col} is the collision frequency, $\bar{E}$ is mean energy dissipation, ρ is the density of the grinding media, ω is the rotational speed of the mill (in Hz), A is the amplitude of oscillation, d_p is the diameter of the milling balls and n_{ball} is the number of milling balls in the system.

Additionally, Figure 3.11 (d) and Figure 3.12 (d) show the ratio between the mean specific tangential and normal dissipation powers for the unaltered and calibrated scenarios, respectively. In this plot, it becomes easier to distinguish the three operational zones previously discussed. The first zone goes from 4% to 10% and favors tangential dissipation. At this point, the elbow of the curve can be identified, meaning that this fill ratio provides the best balance between normal and tangential dissipation in this machine. Then, the second zone, which extends up to 20%, increases the relevance of normal dissipation. Lastly, the third zone, extending beyond 20%, provides no benefit as the ratio is no longer able to achieve values beyond those achievable at lower fill ratios. Furthermore, the ratio remains constant in this range since both types of dissipation begin to decay at the same rate. The identification of this limit is convenient as it establishes a clear upper boundary for the number of balls that should be used in the ball mill. Exceeding this value will result in inefficiency.

It is worth highlighting some key differences that arise in the calibrated master curves (see Figure 3.12), given that these represent the real system relevant for the mechanochemical regeneration of $NaBH_4$. Specifically, it can be observed that while the maximum zone for normal dissipation remains in the 10%-20% range, the actual fill ratio leading to maximum normal dissipation may fluctuate between 10% and 18% due to the standard deviation of the data. This is explained by the behavior identified in subsection 3.3.1, where it was noted that switching from small balls to large balls shifts the mean normal energy dissipation.
Similarly, while the maximum zone for tangential dissipation remains in the 4%-8% range, the standard deviation in the master curve allows for two interpretations: either a maximum is achieved at 4% or 8%. This variation arises due to the dynamic behavior depicted in Figure 3.10 (a), as previously discussed. At low fill ratios, the use of 10 mm balls causes energy dissipation to be dominated by collisions with the wall, effectively minimizing the contribution of ball-ball interactions. Smaller balls, however, do not exhibit this condition. Nevertheless, it is expected that if even lower fill ratios are used,

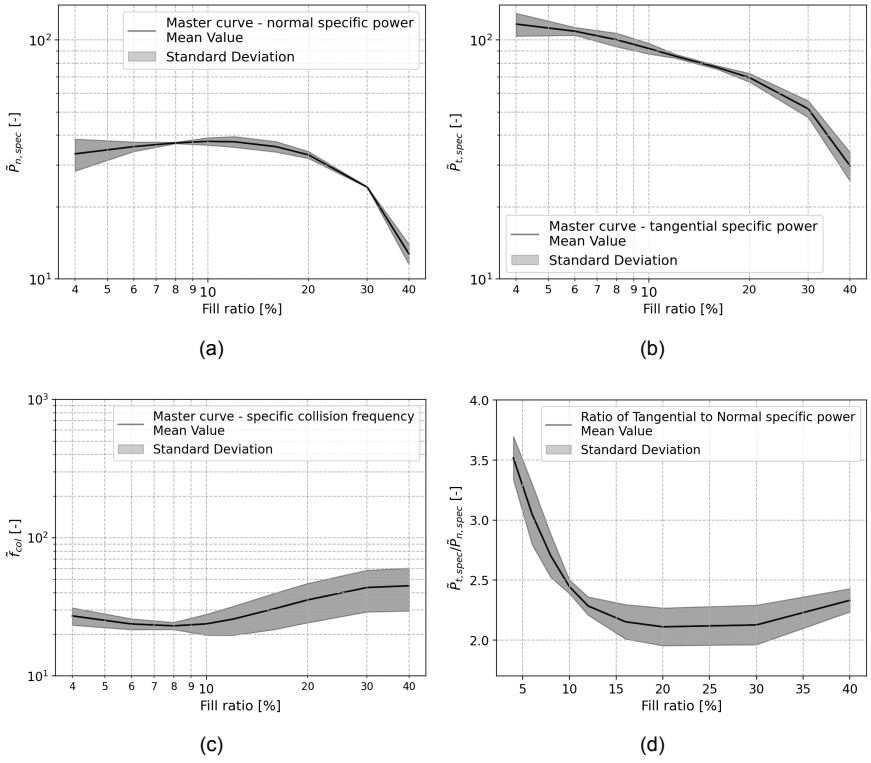


Figure 3.11: (a) Master curve - mean specific normal power dissipation per collision $\tilde{P}_{spec,n}$ (b) Master curve - mean specific tangential power dissipation per collision $\tilde{P}_{spec,t}$ (c) Master curve - mean specific collision frequency $\tilde{f}_{coll}$ (d) Ratio between mean specific tangential power dissipation per collision and mean specific normal power dissipation per collision $\frac{\tilde{P}_{spec,t}}{\tilde{P}_{spec,n}}$. Unaltered steel coefficients.

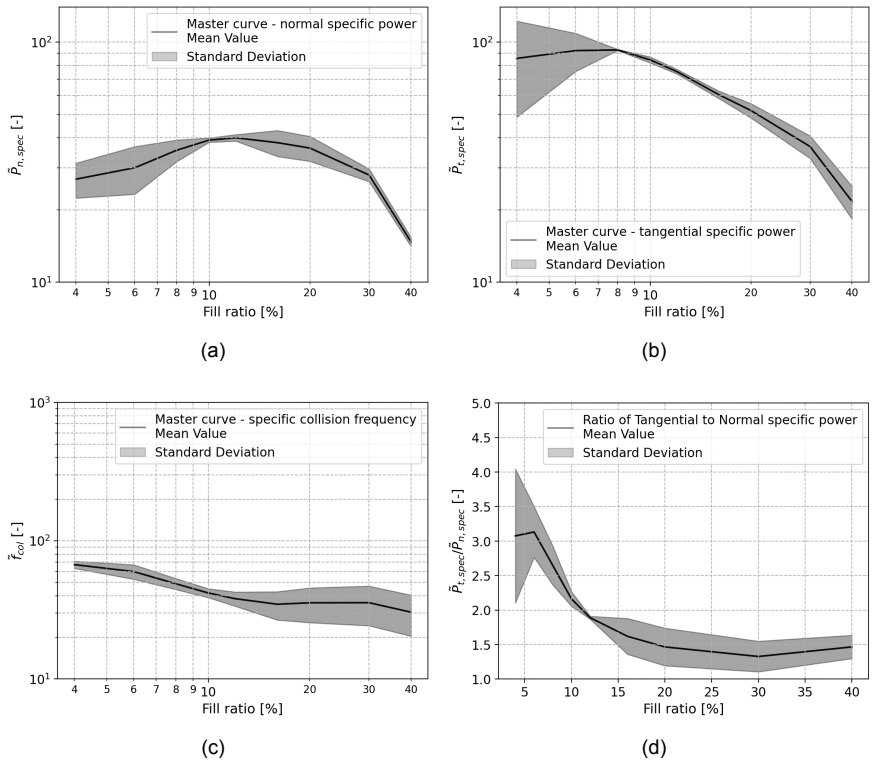


Figure 3.12: (a) Master curve - mean specific normal power dissipation per collision $\tilde{P}_{spec,n}$ (b) Master curve - mean specific tangential power dissipation per collision $\tilde{P}_{spec,t}$ (c) Master curve - mean specific collision frequency $\tilde{f}_{coll,spec}$ (d) Ratio between mean specific tangential power dissipation per collision and mean specific normal power dissipation per collision $\frac{\tilde{P}_{spec,t}}{\tilde{P}_{spec,n}}$. Calibrated coefficients (system comprising $NaBO_2 \cdot 4\,H_2O$ and MgH_2).

collisions between smaller balls will also eventually be minimized, as they will begin to traverse smoothly against the wall of the jar. In this operational range, tangential dissipation can be up to 4 times higher than normal dissipation. The elbow of the curve at 10% results in a ratio of 2.1, and with higher fill ratios, the ratio between tangential and normal dissipation can be minimized to 1.5.
To conclude, it is worth reiterating that these master curves encompass vastly different operational conditions. While the standard deviation may cause a slightly incorrect selection of the maxima, the master curves remain highly effective for facilitating a mechanical characterization of the regeneration of $NaBH_4$ without the need for additional simulations.

3.3.4. Testing the master curves

To test the master curves and their capability for characterization and prediction, it is essential to use an existing mechanochemical experiment along with its chemical yield. For this purpose, we utilize the same reference case used for calibration. This case involves a ball-to-powder ratio of 30, an excess molar ratio of 66%, a milling time of 12.5 h, a rotational speed of 600 rpm, a fill ratio of 10%, and a ball size of 10 mm. Under these conditions, a $NaBH_4$ regeneration yield of 71% was achieved [19]. By applying our master curves and Equation 3.31, we can propose a mechanical characterization of the mechanochemical process by defining three main characteristic values: the mean normal energy dissipation $\bar{E}_n$, calculated as 2.21×10^{-4} J; the mean tangential energy dissipation $\bar{E}_t$, calculated as 5.00×10^{-4} J; and a specific collision frequency $\frac{f_{col}}{n_{ball}}$ of 400 s^{-1}.

With these values defined, it is now possible to use the master curves to derive a new set of conditions that can replicate the characteristic mechanical values. We choose to preserve a fill ratio of 10%, as we wish to maintain the ratio between tangential and normal dissipation for this study. However, we opt to use 5 mm balls requiring a total of 191 milling balls, representing a 700% increase over the reference case. By changing the ball size and keeping the three characteristic mechanical values constant, Equation 3.31 allows us to solve for the density of the balls, rotational speed, or amplitude of rotation. We choose to solve for the rotational speed given the grinding media and equipment available. The new rotational speed is 1698 rpm, which represents a 183% increase over the reference case. A new experiment is conducted under these conditions, yielding a regeneration of 67%, just 4% below the reference case of 71%. This experiment was carried out twice to ensure reproducibility.

3.3.5. Universal master curves

In subsection 3.3.3, we presented two distinct sets of master curves, each independently characterizing the mechanical performance of the ball mill under different friction coefficients, which vary according to the presence of processed material. To develop a universal master curve that can characterize any processed material with calibrated coefficients within the ranges addressed in this work (restitution coefficient: 0.3-0.7,

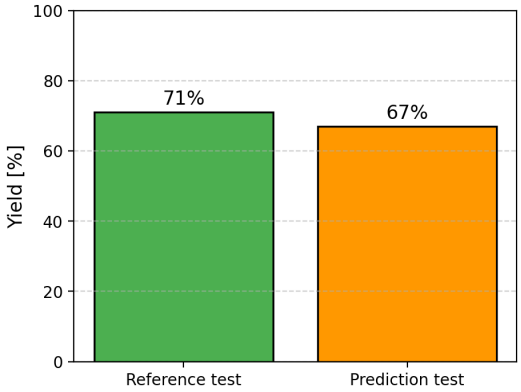


Figure 3.13: Mechanochemical yield in reference case and predicted case

static friction coefficient: 0.3-0.69), we utilize the effective restitution coefficient, ε, as defined by Chialvo and Sundaresan [26].

$$\varepsilon = e - \frac{3}{2}\mu \exp(-3\mu) \tag{3.30}$$

where e is the standard restitution coefficient, and μ is the static friction coefficient. Thus, general equations for constructing universal master curves can be proposed as follows:

$$\tilde{P}_{spec} = \frac{f_{col}\bar{E}}{\rho\omega^3 A^2 d_p^3 n_{ball}} \cdot (1-\varepsilon)^{1/2} \tag{3.31}$$

$$\tilde{f}_{coll} = \frac{f_{col}}{\omega n_{ball}} \cdot (1-\varepsilon)^{1/2} \tag{3.32}$$

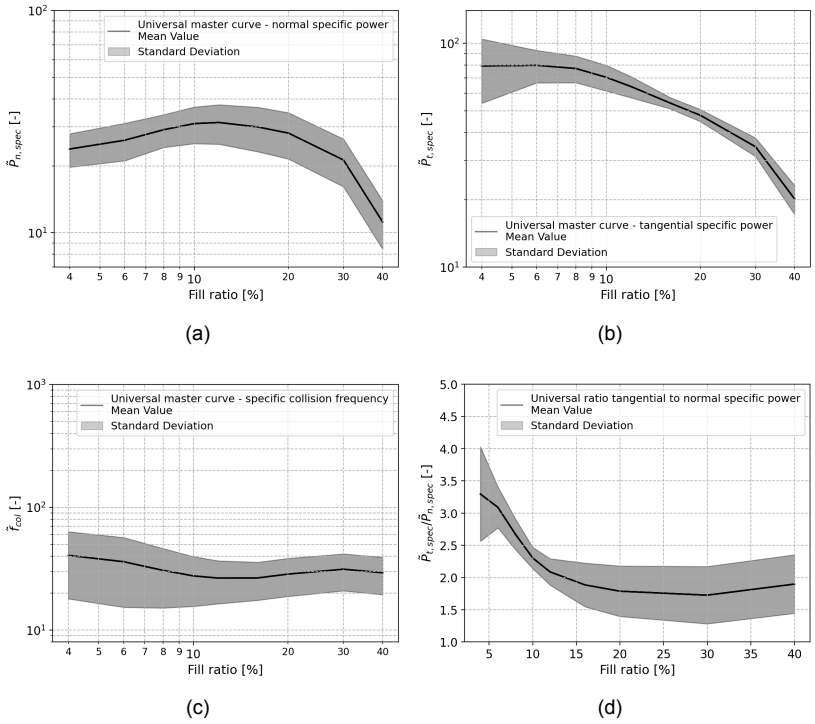


Figure 3.14: (a) Universal master curve - mean specific normal power dissipation per collision $\tilde{P}_{spec,n}$ (b) Universal master curve - mean specific tangential power dissipation per collision $\tilde{P}_{spec,t}$ (c) Universal master curve - mean specific collision frequency $\tilde{f}_{coll,spec}$ (d) Universal ratio between mean specific tangential power dissipation per collision and mean specific normal power dissipation per collision $\frac{\tilde{P}_{spec,t}}{\tilde{P}_{spec,n}}$.

To use these curves effectively, it is necessary to first obtain calibrated values for the restitution coefficient and static friction coefficient to calculate ε. These universal curves are less accurate than the specific-case curves (Figure 3.11 and Figure 3.12) because they do not apply exclusively to a specific set of contact parameters. They aim to capture the dynamic changes caused by adding different processed materials to the milling jar and their impact on the motion of the milling balls. Despite this decrease in accuracy, the universal curves remain a practical and efficient tool for estimating dissipation mechanisms without requiring additional simulations. Moreover, while the utility of this characterization methodology has been tested for the system comprising $NaBO_2 \cdot 4\,H_2O$ and MgH_2, it still needs to be tested for other mechanochemical processes.

3.4. Conclusions

In this chapter, we characterized the internal dynamics of high-energy ball milling for mechanochemistry by distinguishing normal from tangential energy dissipation at ball-ball and ball-jar collisions. Our findings present a clear methodology that can be followed to obtain an in-depth characterization of any milling machine. By following it, the development of green chemical processes can be facilitated, and their generalization and reproducibility across scales and setups becomes viable.

We used the Hertz-Mindlin model and discrete element modeling to accurately characterize the forces involved in ball-ball and ball-wall collisions, allowing us to distinguish the contribution of normal and tangential interactions in the mechanical energy dissipation of the system. Our approach allowed us to identify optimal operational ranges where each form of dissipation can thrive and to characterize the milling process with master curves that predict the performance of the mill under different conditions beyond those presented in this work.

We also found that the contribution of tangential interactions to the mechanical energy dissipation of the system is significant and cannot be neglected, as it dominates the overall dissipation. This is particularly important for use cases where shearing has the potential to be a critical component leading to a successful outcome. Moreover, maximizing or minimizing its significance in the global energy dissipation performance is possible by varying the fill ratio.

The performance of the milling machine is strongly influenced by the milling parameters, such as the ball size, fill ratio, and rotational speed. Our master curves present a valuable tool for designing, optimizing, and predicting mechanochemical processes. Additionally, they enable a direct comparison of key mechanical conditions among different milling machines thanks to the dimensional analysis carried out.

This work presents the first mechanical characterization for the mechanochemical regeneration of $NaBH_4$. By establishing the mean normal energy dissipation, mean tangential energy dissipation, and specific collision frequency as process parameters, we successfully predicted a new set of conditions that replicated a previously achieved result with only a 4% difference.

This work represents a step forward in the understanding of the complex behavior of mechanical forces in high-energy ball milling and their influence on the success of mechanochemical reactions. In the next chapter, we leverage the presented methodology to test whether outcomes are driven by total mechanical energy or by its partitioning between normal and tangential stressing, using targeted experiments to distinguish the two.

References

[1] V. Martinez, T. Stolar, B. Karadeniz, I. Brekalo and K. Užarević. 'Advancing mechanochemical synthesis by combining milling with different energy sources'. In: *Nature Reviews Chemistry* 7.1 (Jan. 2023), pp. 51–65. ISSN: 2397-3358. DOI: 10.1038/s41570-022-00442-1. URL: https://doi.org/10.1038/s41570-022-00442-1.

[2] J.-L. Do and T. Friščić. 'Mechanochemistry: A Force of Synthesis'. In: *ACS Central Science* 3.1 (2017). PMID: 28149948, pp. 13–19. DOI: 10.1021/acscentsci.6b00277. eprint: https://doi.org/10.1021/acscentsci.6b00277. URL: https://doi.org/10.1021/acscentsci.6b00277.

[3] H. X. Nunes, D. L. Silva, C. M. Rangel and A. M. F. R. Pinto. 'Rehydrogenation of Sodium Borates to Close the NaBH4-H2 Cycle: A Review'. In: *Energies* 14.12 (2021). ISSN: 1996-1073. DOI: 10.3390/en14123567. URL: https://www.mdpi.com/1996-1073/14/12/3567.

[4] L. M. Martínez, J. Cruz-Angeles, M. Vázquez-Dávila, E. Martínez, P. Cabada, C. Navarrete-Bernal and F. Cortez. 'Mechanical Activation by Ball Milling as a Strategy to Prepare Highly Soluble Pharmaceutical Formulations in the Form of Co-Amorphous, Co-Crystals, or Polymorphs'. In: *Pharmaceutics* 14.10 (2022). ISSN: 1999-4923. DOI: 10.3390/pharmaceutics14102003. URL: https://www.mdpi.com/1999-4923/14/10/2003.

[5] V. A. Rodriguez, L. Ribas, A. Kwade and L. M. Tavares. 'Mechanistic modeling and simulation of a wet planetary ball mill'. In: *Powder Technology* 429 (2023), p. 118901. ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2023.118901. URL: https://www.sciencedirect.com/science/article/pii/S003259102300685X.

[6] F. Delogu and L. Takacs. 'Information on the mechanism of mechanochemical reaction from detailed studies of the reaction kinetics'. In: *Journal of Materials Science* 53.19 (Oct. 2018), pp. 13331–13342. ISSN: 1573-4803. DOI: 10.1007/s10853-018-2090-1. URL: https://doi.org/10.1007/s10853-018-2090-1.

[7] W. Chen, L. Ouyang, J. Liu, X. Yao, H. Wang, Z. Liu and M. Zhu. 'Hydrolysis and regeneration of sodium borohydride (NaBH 4) – A combination of hydrogen production and storage'. In: *Journal of Power Sources* 359 (Aug. 2017), pp. 400–407. DOI: 10.1016/j.jpowsour.2017.05.075.

[8] M. Broseghini, M. D'Incau, L. Gelisio, N. Pugno and P. Scardi. 'Effect of jar shape on high-energy planetary ball milling efficiency: Simulations and experiments'. In: *Materials & Design* 110 (2016), pp. 365–374. ISSN: 0264-1275. DOI: https://doi.org/10.1016/j.matdes.2016.06.118. URL: https://www.sciencedirect.com/science/article/pii/S0264127516308802.

[9] C. Burmeister, R. Schmidt, K. Jacob, S. Breitung, A. Stolle and A. Kwade. 'Effect of stressing conditions on mechanochemical Knoevenagel synthesis'. In: *Chemical Engineering Journal* 396 (Feb. 2020), p. 124578. DOI: 10.1016/j.cej.2020.124578.

[10] C. F. Burmeister, A. Stolle, R. Schmidt, K. Jacob, S. Breitung-Faes and A. Kwade. 'Experimental and Computational Investigation of Knoevenagel Condensation in Planetary Ball Mills'. In: *Chemical Engineering & Technology* 37.5 (2014), pp. 857–864. DOI: https://doi.org/10.1002/ceat.201300738. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ceat.201300738. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/ceat.201300738.

[11] C.-L. Hsueh, C.-H. Liu, B.-H. Chen, C.-Y. Chen, Y.-C. Kuo, K.-J. Hwang and J.-R. Ku. 'Regeneration of spent-$NaBH_4$ back to $NaBH_4$ by using high-energy ball milling'. In: *International Journal of Hydrogen Energy* 34.4 (2009), pp. 1717–1725. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2008.12.036. URL: https://www.sciencedirect.com/science/article/pii/S0360319908016662.

[12] A. Oliveira, V. Rodriguez, R. de Carvalho, M. Powell and L. Tavares. 'Mechanistic modeling and simulation of a batch vertical stirred mill'. In: *Minerals Engineering* 156 (2020), p. 106487. ISSN: 0892-6875. DOI: https://doi.org/10.1016/j.mineng.2020.106487. URL: https://www.sciencedirect.com/science/article/pii/S0892687520303071.

[13] S. Beinert, G. Fragnière, C. Schilde and A. Kwade. 'Analysis and modelling of bead contacts in wet-operating stirred media and planetary ball mills with CFD–DEM simulations'. In: *Chemical Engineering Science* 134 (2015), pp. 648–662. ISSN: 0009-2509. DOI: https://doi.org/10.1016/j.ces.2015.05.063. URL: https://www.sciencedirect.com/science/article/pii/S0009250915004030.

[14] Altair Engineering Inc. *The Hertz-Mindlin (no slip) model*. URL: https://help.altair.com/EDEM/Creator/Physics/Base_Models/Hertz-Mindlin_(no_slip.htm). (accessed: 08.08.2023).

[15] DCS Computing GmbH, JKU Linz and Sandia Corporation. *Gran model hertz model*. URL: https://www.cfdem.com/media/DEM/docu/gran_model_hertz.html. (accessed: 05.01.2023).

[16] C. Burmeister, M. Hofer, P. Molaiyan, P. Michalowski and A. Kwade. 'Characterization of Stressing Conditions in a High Energy Ball Mill by Discrete Element Simulations'. In: *Processes* 10 (Apr. 2022), p. 692. DOI: 10.3390/pr10040692.

[17] P. Santhanam and E. Dreizin. 'Predicting conditions for scaled-up manufacturing of materials prepared by ball milling'. In: *Powder Technology* 221 (May 2012), pp. 403–411. DOI: 10.1016/j.powtec.2012.01.037.

[18] S. Rosenkranz, S. Breitung-Faes and A. Kwade. 'Experimental investigations and modeling of the ball motion in planetary ball mills'. In: *Powder Technology* 212 (Sept. 2011), pp. 224–230. DOI: 10.1016/j.powtec.2011.05.021.

[19] S. Garrido Nuñez, D. L. Schott and J. T. Padding. 'Optimization of operational parameters in the mechanochemical regeneration of sodium borohydride (NaBH4)'. In: *International Journal of Hydrogen Energy* 97 (2025), pp. 640–648. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2024.11.360. URL: https://www.sciencedirect.com/science/article/pii/S0360319924050511.

[20] C. Jayasundara and H. Zhu. 'Impact energy of particles in ball mills based on DEM simulations and data-driven approach'. In: *Powder Technology* 395 (Sept. 2021). DOI: 10.1016/j.powtec.2021.09.063.

[21] D. Lide. *CRC Handbook of Chemistry and Physics, 90th Edition*. Taylor & Francis, 2009. ISBN: 978-1-4200-9084-0. URL: https://books.google.nl/books?id=0mkbNgAACAAJ.

[22] Lucefin Group. *X46Cr13 technical card*. URL: https://www.lucefin.com/wp-content/files_mf/1.4034a420c25.pdf. (accessed: 29.08.2023).

[23] I. Goldhirsch. 'Introduction to granular temperature'. In: *Powder Technology* 182.2 (2008). Granular Temperature, pp. 130–136. ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2007.12.002. URL: https://www.sciencedirect.com/science/article/pii/S0032591007006109.

[24] R. W. Breault, C. P. Guenther and L. J. Shadle. 'Velocity fluctuation interpretation in the near wall region of a dense riser'. In: *Powder Technology* 182.2 (2008). Granular Temperature, pp. 137–145. ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2007.08.018. URL: https://www.sciencedirect.com/science/article/pii/S0032591007004627.

[25] S. Lommen, D. Schott and G. Lodewijks. 'DEM speedup: Stiffness effects on behavior of bulk material'. In: *Particuology* 12 (2014). Special issue on conveying and handling of particulate solids – Challenges of discrete element simulation, application and calibration, pp. 107–112. ISSN: 1674-2001. DOI: https://doi.org/10.1016/j.partic.2013.03.006. URL: https://www.sciencedirect.com/science/article/pii/S1674200113001387.

[26] S. Chialvo and S. Sundaresan. 'A modified kinetic theory for frictional granular flows in dense and dilute regimes'. In: *Physics of Fluids* 25.7 (July 2013), p. 070603. ISSN: 1070-6631. DOI: 10.1063/1.4812804. eprint: https://pubs.aip.org/aip/pof/article-pdf/doi/10.1063/1.4812804/15710561/070603_1_online.pdf. URL: https://doi.org/10.1063/1.4812804.

4

Impact of tangential to normal stress partitioning on $NaBH_4$ regeneration

Building on the mechanical descriptors of chapter 3, this chapter explores whether reactivity during $NaBH_4$ regeneration (from $NaBO_2 \cdot 4\,H_2O + MgH_2$) is governed by the total mechanical energy delivered or by the way that energy is divided into normal (compressive) versus tangential (shear) components. Using the descriptors, we design two complementary experimental comparisons: matched total dissipation with different normal/tangential partitions and matched partition with different totals. These comparisons separate "how much" from "how it is delivered." As a first approximation to efficiency, we summarize performance with simple, setup-independent metrics to gauge productivity per unit power without committing to a full techno-economic analysis.

Parts of this chapter are adapted from: Garrido Nuñez, S., Schott, D. L. and Padding, J. T. 'Influence of shear and compressive stress regimes on efficient $NaBH_4$ mechanochemical regeneration'. In: Chemical Engineering Journal 525, 170168 (2025).

4.1. Introduction

4.1.1. Sodium borohydride ($NaBH_4$)

Sodium borohydride ($NaBH_4$) is an attractive solid hydrogen carrier given its high theoretical energy density and gravimetric hydrogen storage [1, 2]. The hydrolysis of $NaBH_4$ releases hydrogen and produces sodium metaborate ($NaBO_2 \cdot xH_2O$) as a byproduct.

$$NaBH_4 + (2 + x)H_2O \rightarrow NaBO_2 \cdot xH_2O + 4\,H_2 \tag{4.1}$$

where x is the level of hydration.

The central barrier to the deployment of sodium borohydride at scale is not the hydrogen release, but rather its regeneration. In 2007, the U.S. Department of Energy (DOE) go/no-go reviews concluded that all assessed $NaBH_4$ pathways exceeded cost targets primarily due to high regeneration costs and sodium price sensitivity, leading to a no-go recommendation for on-board hydrolysis [3]. Since then, $NaBH_4$ mechanochemical loops have co-integrated hydrogen release and regeneration within a single, solvent-free process window by capturing the hydrolysate as borates/carbonates and solid-state reducing it back to $NaBH_4$ under ambient milling. The approach bypasses high-pressure H_2, compression, solvent use, and dehydration, consistently achieving ~70–80% regeneration yield [4–9]. Relevant examples in the context of this work are shown in Table 4.1.

Analogous one-step variants have been demonstrated for $LiBH_4$ and $Mg(BH_4)_2$, underscoring that mechanochemistry enables new opportunities for process integration and cost reduction across complex hydrides [16–19]. We also note a related mechanochemical route to $NaBH_4$ from $NaB(OH)_4$ using Mg–Al intermetallics; reported yields remain comparatively low under prolonged milling (<~43%), reflecting different interfacial chemistry than the MgH_2-driven cycles considered here [20]. In chapter 2, we have presented results of the mechanochemical regeneration of $NaBH_4$ from a system comprising $NaBO_2 \cdot 4\,H_2O$ and MgH_2 (Equation 4.2) with yields reaching up to

Table 4.1: Reported yields for the mechanochemical regeneration of $NaBH_4$.

Metaborate	2nd reactant	Yield (%)	Year	Ref.
$NaBO_2$	MgH_2	76	2009	[10]
$NaBO_2$	MgH_2	71	2009	[11]
$NaBO_2$	MgH_2	74	2011	[12]
$NaBO_2$	MgH_2	89	2017	[13]
$NaBO_2 \cdot 2\,H_2O$	Mg	68	2017	[14]
$NaBO_2 \cdot 2\,H_2O$	MgH_2	90	2017	[15]
$NaBO_2 \cdot 4\,H_2O$	MgH_2	88	2017	[15]
$NaBO_2 \cdot 2\,H_2O$	Mg_2Si	78	2017	[5]

90% along with a ranking of the operational variables (rotational speed, milling time, ball-to-powder ratio and molar ratio) based on their relevance to the chemical yield, with milling time being the most significant [21].

$$NaBO_2 \cdot 4\,H_2O + 6\,MgH_2 \rightarrow NaBH_4 + 6\,MgO + 8\,H_2 \tag{4.2}$$

More generally, mechanochemistry harnesses mechanical energy to drive chemical transformations, but its practice often treats the applied stresses simplistically (e.g., as a pressure or scalar force) rather than as a full stress tensor [22, 23]. In reality, mechanical loading in ball milling involves both normal (compressive) and tangential (shear) stress components. The distinct roles of these stress modes, however, have rarely been explored in mechanochemical studies. Apart from a few targeted studies, most notably Kobayashi et al. [24] who mapped the spatial distribution of normal and tangential stresses for a single ball in a planetary ball mill, most reports and mechanistic studies treat mechanical loading as a single pressure or force magnitude, overlooking the fact that shear and compression can influence reaction pathways differently [22, 25]. For example, the Bell model expresses the rate enhancement via a term $\exp\left(-\frac{P\Delta V}{RT}\right)$ involving an isotropic pressure P, but effectively ignores any anisotropy [25]. This simplification means that chemically important distinctions, for example, that tensile stress tends to break bonds, while normal stress often drives bond formation, can be overlooked [26].

Indeed, several authors point out that common methodologies lack independent control of stress modes. For example, studies of boundary lubrication films (ZDDP tribofilm growth) show that prior experiments based on tribometers could not independently control the normal and tangential stress [27]. Quantitative comparisons further underscore the oversight: for ZDDP tribofilm formation, the measured activation volume under shear loading (~0.18 nm^3) is nearly twenty times larger than that under pure compression (~0.01 nm^3), confirming that tangential, not normal stress, drives the mechanochemical reaction [27].

Several factors contribute to this oversimplification. On the experimental side, it is difficult to apply or measure pure shear without collateral normal forces. Most common mechanochemical reactors (ball mills, twin-screw extruders) impose mixed stress states. For instance, planetary and high-energy ball mills induce both impact and shear forces, and continuous extruders force material through narrow channels applying high tangential and normal stress simultaneously [28]. Few laboratory techniques allow independent control of shear versus compression. Traditional pressure devices (diamond-anvil cells, gas- or liquid-pressure cells, AFM tips) primarily deliver uniform normal stress, while tribometers typically combine tangential and normal loading [22]. Even with these limited experimental options, conclusions drawn under such idealized conditions remain difficult to generalize to bulk mechanochemical operations, where complex, varying stress landscapes deviate from bench-top experiments. A consequence of this is that well-defined shear loading is rarely isolated in mechanochemical experiments, so mechanisms driven by sliding or

frictional forces are usually inferred indirectly [28].

In summary, both experimental practice and theoretical frameworks have so far treated mechanical loading as effectively scalar, bypassing the need to decouple components. These simplifications have significant implications. By ignoring stress anisotropy, predictive models may fail to capture important mechanistic pathways. For instance, Jonas et al. showed that under pure sliding (shear) conditions, there exists a finite critical stress below which no reaction occurs, contrary to the continuous behavior predicted by the scalar Bell model [25]. If one calibrates a model on compression-induced data, it may not predict shear-driven kinetics correctly. Similarly, an experimental result obtained under one loading mode (e.g., hydrostatic press) may not translate to another (e.g., milling) if the shear contribution is different, leading to misinterpretation of which bonds are activated. Moreover, the different intrinsic working principles that mills have also produced a wide variation in collision frequencies, impact energies, and shear rates, rendering comparative evaluation across milling platforms complicated [29–33]. This lack of detail can also hamper reactor design when up-scaling becomes relevant; without knowing how shear or normal forces drive a transformation, it is hard to optimize milling media, extruder screw profiles, or frictional conditions to maximize yield or selectivity. In practice, this could mean that some reaction products remain inaccessible simply because the wrong stress component is being applied. Overall, neglecting shear–normal distinctions reduces the predictive power of mechanochemical models and can obscure the true mechanism of activation.

A concrete example relevant to the context of this work is the regeneration of $NaBH_4$. Research has focused on trying different reducing agents or hydration levels while attempting to maximize yield within a limited set of operational parameters (rotational speed, ball-to-powder ratio (BPR), milling time, and molar ratio) [5, 10–15]. Unfortunately, this has unintentionally diverted attention away from deepening the mechanical understanding of the system that defines the reaction's success [34]. This presents several challenges: reproducibility becomes problematic unless identical equipment is used, scaling-up becomes challenging, the connection between macroscopic behavior and molecular transformations is neglected, and predictive models that allow quantitative descriptions of expected chemical conversion are scarce [35]. As reported in chapter 2, the intended replication of identical operational parameters in a different ball mill (for example, going from a shaker ball mill to a high-energy ball mill) led to completely different results in the mechanochemical regeneration of $NaBH_4$. Thus, given the fundamental changes in design, operation, and energy input of different mills, it becomes evident that these operational parameters are not sufficient to accurately characterize a mechanochemical process if the ultimate goals are increasing efficiency, scaling up, and ensuring reproducibility across devices.

To overcome the shortcomings outlined above, namely poor interlaboratory reproducibility, the inability to decouple normal and tangential stress modes, and the resulting obstacles to reliable scale-up, we proposed and validated the methodology presented in chapter 3 which characterizes mechanochemical processes by analyzing

the mechanical interactions occurring inside the milling jar through Discrete Element Method (DEM) modeling [36]. By defining three characteristic values: the mean normal energy dissipation per collision, $\bar{E}_n$; the mean tangential energy dissipation per collision, $\bar{E}_t$; and the specific collision frequency, f_{col}/n_{ball}, we could successfully predict the chemical yield of $NaBH_4$ regeneration under previously untested operational conditions. Together, these metrics capture the distribution of energy dissipated between normal and tangential contacts, providing a practical representation for the underlying stress landscape in a bulk mechanochemical process because the energy released in each collision scales directly with the normal or tangential force components; partitioning that energy therefore mirrors how compressive and shear stresses are delivered in the jar.

Because these descriptors are derived directly from particle–particle and particle–wall interactions, they remain independent of the mill's working principle and design. Consequently, any milling device that can reproduce the triplet $\{\bar{E}_n, \bar{E}_t, f_{col}/n_{ball}\}$ should deliver the expected chemical conversion. Moreover, since the framework explicitly separates the contributions of compression (normal interactions) and shear (tangential interactions), it gives the unique opportunity to deliberately tune their balance in the system while keeping the total mechanical power dissipation constant. Since this power can be expressed as $(\bar{E}_n + \bar{E}_t) f_{col}$, the mechanical descriptors can be manipulated to maintain an invariant mechanical 'budget' even as the fill ratio (i.e. the volume occupied by the balls divided by the jar's total internal volume) changes, and hence the grinding media mass and internal dynamics. The master-curve framework provides the compensating adjustments in rotational speed or amplitude required to hold that power fixed while smoothly shifting the ratio $\bar{E}_t/\bar{E}_n$ from shear-dominated to compression-dominated regimes. This decouples the stress mode from net energy delivery, allowing mechanistic comparisons that are free from confounding effects and isolating the true influence of shear versus normal loading on reaction kinetics. Moreover, projecting any milling device onto the same $\{\bar{E}_n, \bar{E}_t, f_{col}/n_{ball}\}$ space yields a quantitative stress fingerprint that immediately reveals whether the mill intrinsically favors shear or compression, providing a rigorous, transferable basis for equipment benchmarking, optimization, and scale-up.

4.2. Model and materials

The Emax high-energy ball mill, distributed by the German company Retsch, is used for all experiments. The system was set up to allow a maximum temperature of 50 °C. The machine allocates proprietary grinding jars with 125 mL of volume that follow a circular motion with a rotational speed n up to 2000 revolutions per minute with an amplitude (radius) A of 1.7 cm, see Fig. 6.1.

Hydrated sodium metaborate ($NaBO_2 \cdot 4\,H_2O$) (≥ 99%) was purchased from Sigma-Aldrich, while magnesium hydride (MgH_2) (≥ 99.9%, ≤ 50μm) was obtained from Nanoshel. All reactants were used without further purification. The sample preparation for ball milling was carried out in a glovebox under an argon atmosphere, with oxygen and water concentrations maintained below 0.1 ppm. For a detailed description of the

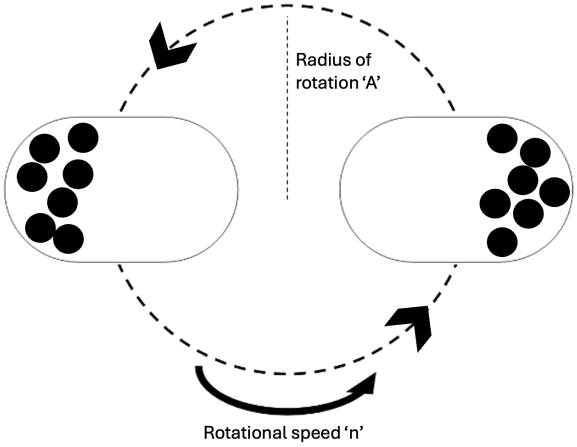


Figure 4.1: Schematic of jar movement

quantification of chemical yield and the cleaning of the equipment to preserve similar conditions for all experimental cases, we refer to chapter 2.

4.2.1. Modeling approach and case definition

The interactions between the milling jar and the grinding media were modeled and simulated using the Discrete Element Method (DEM). For this study, Altair EDEM 2021.2 was used as DEM solver, while Python 3.9.12 was employed for data post-processing. EDEM implements a soft-sphere approach, calculating particle contact forces based on the Hertz-Mindlin model. Refer to chapter 3 for a detailed description of the contact model and the equations governing the particle's motion in the DEM framework, as well as for the calibration and data post-processing carried out to produce the master curves presented in Figure 4.2.

The master curves (Figure 4.2 (a) and (b)) represent the expected mechanical action of the high-energy ball mill in terms of mean specific power dissipation in the normal and tangential directions, respectively. Additionally, the specific collision frequency and ratio between both modes of dissipation are also presented (Figure 4.2 (c) and (d)). These master curves were obtained from a large number of DEM simulations, normalizing the measured dissipated energies and collisions through Equation 4.3 and Equation 4.4.

$$\tilde{P}_{spec} = \frac{f_{col}\bar{E}}{\rho\omega^3 A^2 d_p^3 n_{ball}} \tag{4.3}$$

$$\tilde{f}_{coll} = \frac{f_{col}}{\omega n_{ball}} \tag{4.4}$$

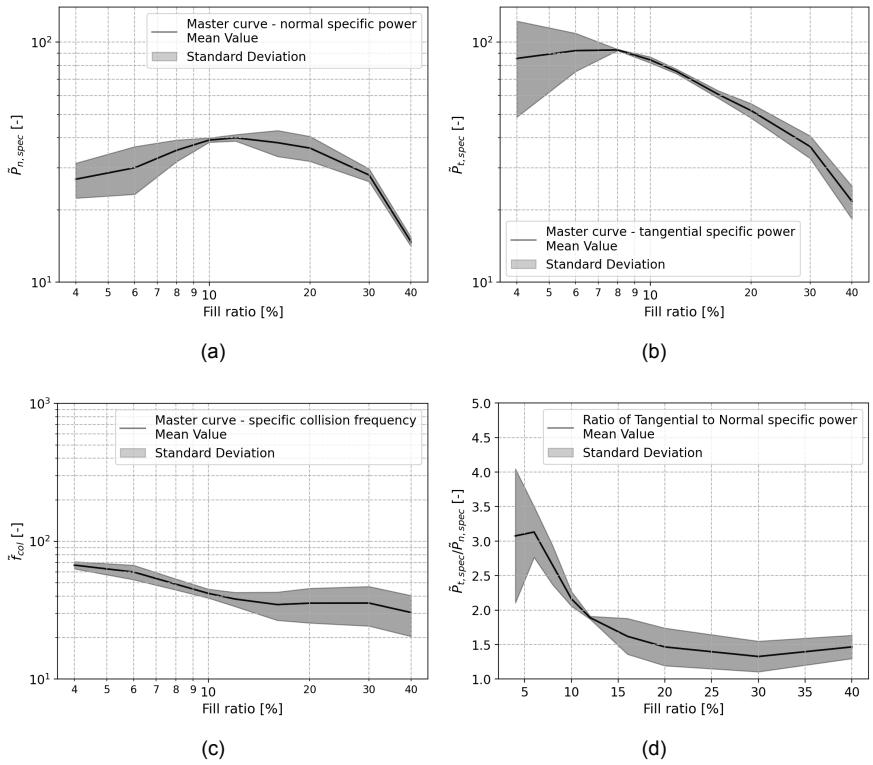


Figure 4.2: (a) Master curve - mean specific normal power dissipation per collision $\tilde{P}_{spec,n}$ (b) Master curve - mean specific tangential power dissipation per collision $\tilde{P}_{spec,t}$ (c) Master curve - mean specific collision frequency $\tilde{f}_{col}$ (d) Ratio between mean specific tangential power dissipation per collision and mean specific normal power dissipation per collision $\tilde{P}_{spec,t}/\tilde{P}_{spec,n}$. Calibrated coefficients (system comprising $NaBO_2 \cdot 4\,H_2O$ and MgH_2).

Where ρ is the milling ball density, ω is the rotational speed of the mill, A is the amplitude of oscillation, d_p is the diameter of the milling balls, and n_{ball} is the number of milling balls in the system.

The combination of the master curves along with Equation 4.3 and Equation 4.4 enables the estimation of the triplet of characteristic values ($\bar{E}_n$, $\bar{E}_t$ and f_{col}/n_{ball}) for this mechanochemical process without the need for any simulations. Then, it is possible to calculate the total power following Equation 4.5.

$$P_{\text{tot}} = f_{\text{col}} \left(\bar{E}_{\text{n}} + \bar{E}_{\text{t}} \right) \tag{4.5}$$

Looking at the ratio between tangential and normal dissipation (Figure 4.2 (d)), it becomes possible to define three distinct operational regimes that strictly depend on the fill ratio. The first goes from 4% to 9% fill ratio and favors tangential dissipation. At 10%, an elbow can be identified in the curve, meaning that this fill ratio provides the best balance between normal and tangential dissipation in this mill. Then, the second zone, which extends up to 20%, increases the relevance of normal dissipation. Lastly, the third zone, extending beyond 20%, provides no benefit as the ratio can no longer achieve values beyond those achievable at lower fill ratios. Thus, to test the effect of each dissipation regime on the mechanochemical yield, the cases presented in Table 4.2 are selected.

The five cases in Table 4.2 are designed with the following rationale: a reference case, with a fill ratio of 10%, achieves the best balance in dissipation between tangential and normal components in the context of the Emax mill (see Figure 4.2 (d)). Cases favoring

Table 4.2: Comparison of milling scenarios at constant rotational speed versus constant total power. Two biased cases are shown for each operating mode: one that favours tangential dissipation and one that favours normal dissipation.

	Reference	Constant rot. speed		Constant total power	
Parameter	case	*Favor tang.*	*Favor norm.*	*Favor tang.*	*Favor norm.*
Fill ratio (%)	0.10	0.06	0.17	0.06	0.17
Rot. speed (Hz)	10.0	10.0	10.0	13.14	9.03
Spec. power, normal	39.7	23	43	23	43
Spec. power, tangential	90	75	60	75	60
Spec. collision frequency	40	70	25	70	25
Collision frequency (s^{-1}/n_{ball})	400	700	250	920	226
P_{total} (W)	6.9	3.0	9.4	6.9	6.9
Normal energy (10^{-4} J)	2.21	0.73	3.82	1.26	3.12
Tangential energy (10^{-4} J)	5.00	2.38	5.34	4.11	4.35
Ratio tang to norm.	2.26	3.26	1.40	3.26	1.39
Yield (%)	71	84	43	94	40

tangential dissipation use a fill ratio of 6%, while those favoring normal dissipation use a fill ratio of 17%. It is worth highlighting that, until now, the influence of the fill ratio on the mechanochemical regeneration of $NaBH_4$ has not been investigated at all, let alone with a mechanistic rationale [21]. In broader mechanochemical research, this parameter is shifted for pragmatic or empirical reasons, making it unclear whether the reported results on mechanochemical yield arise from altered power input, modified collision modes, or both [37, 38]. To account for the effects of these different fill ratios on the system's dynamics and mechanical behavior, two approaches are introduced: constant rotational speed and constant total power. Using both approaches, we can introduce a systematic exploration of the effects of the fill ratio in a mechanochemical reaction by enforcing power equivalence, providing a controlled platform for mechanistic interpretation.

State-of-the-art mechanochemical studies are typically carried out using only the constant rotational speed approach, where the mill speed stays fixed while the fill ratio is adjusted. Varying the fill ratio changes the total power dissipation because both the mass in the jar (number of milling balls) and the collision frequency shift. As stated above, the problem is that any change in reaction outcome now combines two effects: tangential vs. normal distribution and total power, making it impossible to determine which factor drives the yield. To resolve this confounding, the constant power strategy can be used.

In the constant power approach, the machine's rotational speed is carefully adjusted using Equation 4.3 and Equation 4.4 to maintain constant total power dissipation. Since ultimately the goal is to set $P_{tot,ref} = P_{tot,i}$, the aforementioned equations can be manipulated to arrive at Equation 4.6. This step assumes that the ball material and diameter, as well as the machine's amplitude, remain unchanged. Nonetheless, the influence of these parameters can be readily implemented back into the expression if they are varied.

$$\omega_i = \omega_{\text{ref}} \left(\frac{n_{\text{ball,ref}} \tilde{P}_{\text{spec,ref}}}{n_{\text{ball},i} \tilde{P}_{\text{spec},i}} \right)^{1/3} \tag{4.6}$$

In both strategies, the tangential to normal dissipation ratio remains constant because it is set solely by the fill ratio. What changes is how the total power is delivered to the reactants. By keeping that power constant, any shift in yield reflects how the power is supplied, rather than how much the system receives (see Table 4.2). Furthermore, all cases have the following previously optimized chemical operational parameters: BPR of 30, molar ratio of 66% above the stoichiometric value (see Equation 4.2), milling time of 12.5 hours, and stainless steel balls (1.4034 G100 DIN 5401) with a diameter of 10 mm [21]. Each case was carried out three times to ensure reproducibility.

4.3. Results and Discussion

This section presents the results obtained when the relative contribution of tangential and normal collision modes inside the jar is systematically varied. Two complementary

approaches were applied to isolate mechanical effects. The first follows constant rotational speed, and the second follows constant total mechanical power. Alongside the absolute regeneration yield, each run is evaluated through its specific yield (W^{-1}), which relates fractional conversion yield directly to the mechanical power consumed, and through the converted mass per Watt ($g \cdot W^{-1}$), which relates net product output directly to the mechanical power consumed. Finally, we estimate mechanochemical energy leverage metrics λ_{mc}, which measure the fraction of mechanical energy converted into recoverable chemical energy.

4.3.1. Tangential to normal dissipation ratio

The results in Figure 4.3 depict how the conversion yield changes with fill ratio, which dictates the ratio between tangential and normal stressing events in the milling machine. In both approaches, the lowest fill (6%) leading to the highest tangential contribution gives the best performance. Under constant power operation, this achieves a regeneration yield of 94%, the highest value reported to date for the $NaBO_2 \cdot 4\,H_2O$ and MgH_2 system. Specifically, this is achieved while reducing the milling time by 38%, the ball-to-powder ratio by 40% (meaning more powder can be treated within the same batch), and the rotational speed by 34% (see Table 4.3) [15]. These gains demonstrate the potential for optimizing mechanochemical reactions with a deeper understanding of

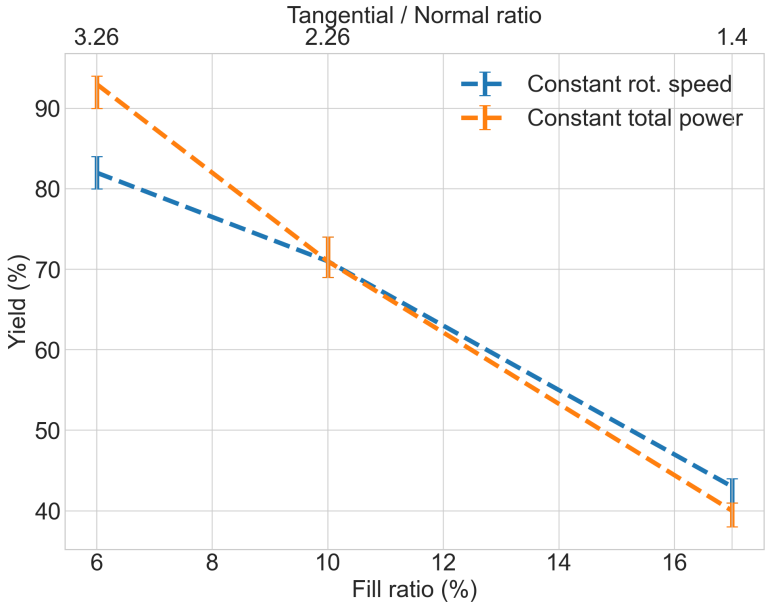


Figure 4.3: Yield versus fill ratio for experiments run (i) at constant rotational speed (blue) and (ii) at constant total power (orange). The secondary (top) axis converts each fill ratio to the corresponding tangential/normal dissipation ratio inside the jar. Error bars show the variability in yield.

interactions among variables and the influence of mechanical conditions on the system.

When the system is held at constant speed, the tangential biased case still leads, but the yield is about ten percentage points lower than in the constant power approach. For normal-dominant cases, the opposite trend appears; constant speed offers a small advantage of less than five percentage points. A simple first-order model captures this behavior with high fidelity. Linear regression of yield against fill ratio gives $R^2 > 0.99$ for both approaches. The fitted slopes quantify the penalty for reducing tangential influence: yield falls by ≈ 0.038 per percentage of fill ratio at constant speed and by ≈ 0.049 per percentage of fill ratio at constant power, the latter being about 30% steeper. Thus, increasing fill ratio from 6% to 17% lowers the yield by roughly 0.4–0.5 (40–50 percentage points) regardless of the strategy.

For $NaBO_2 \cdot xH_2O + MgH_2 \longrightarrow NaBH_4$, solid-state studies support a stepwise interfacial hydride-substitution pathway: $[B(OH)_4]^-$ converts via a borohydride–hydroxy intermediate ($BH_3(OH)^-$/'$NaBH_3OH$') before full formation of $BH_4{}^-$, with ^{11}B MAS NMR directly detecting the intermediate under high-energy milling [15, 39]. Accordingly, milling modes that maximize shear contact at reactive interfaces, continually renewing MgH_2/borate surfaces and abrading passivation layers (e.g., MgO, $Mg(OH)_2$), are more productive per Watt than brief head-on impacts (see Table 4.4 and Figure 4.4). The advantage of a tangential-rich regime is consistent with known shear-activation channels in mechanochemistry [40]; it increases defect density and freshly created surface area, accelerates intimate mixing at reactive interfaces, and thereby can lower effective barriers for B – H bond-forming steps toward $BH_4{}^-$.

Across the different operating conditions tested, the highest specific yield performance is achieved when tangential dissipation dominates. Under the constant rotational speed regime, the mill reaches a specific yield of 0.28 (W^{-1}), which is 2.7 times higher than the balanced reference. Even when total mechanical power is held fixed at 6.9 W, simply reorienting collisions from normal to tangential raises the specific yield from 0.058 to 0.136 (W^{-1}), delivering a 134% jump in productivity without any additional energy input. By contrast, normal biased operation is doubly penalized as it demands more power (9.4 W versus 6.9 W) yet still depresses specific yield from 0.058 to 0.046 (W^{-1}).

It is also worth noting that, once tangential bias has driven the yield into

Table 4.3: Mechanochemical regeneration of $NaBH_4$ from $NaBO_2 \cdot 4H_2O + MgH_2$: prior art (Chen et al., 2017[15]) versus this work.

Metric	Chen et al., 2017[15]	This work	Change
Regeneration yield (%)	88	94	+6%
Milling time (h)	20	12.5	−38%
Ball-to-powder ratio (BPR)	50	30	−40%
Rotational speed (Hz)	20	13.14	−34%

Table 4.4: Yield normalised by total mechanical power input. The lower block continues with the mass-based metrics for the same five conditions.

Condition (orientation · fill · # balls)	P_{tot} (W)	Yield	Specific-yield (yield·W^{-1})	Gain vs. ref. (%)
Reference – balanced · 10 % · 24	6.9	0.71	0.103	–
Tangential favored – 6 % · 14	3.0	0.84	0.280	+172
Normal favored – 17 % · 41	9.4	0.43	0.046	-56
Tangential favored – 6 % · 14	6.9	0.94	0.136	+32
Normal favored – 17 % · 41	6.9	0.40	0.058	-44

Condition (orientation · fill · # balls)	$NaBO_2 4H_2O$ mass (g)	Converted mass per W (g·W^{-1})	Gain vs. ref. (%)
Reference – balanced · 10 % · 24	1.10	0.113	–
Tangential favored – 6 % · 14	0.66	0.185	+64
Normal favored – 17 % · 41	1.87	0.086	-24
Tangential favored – 6 % · 14	0.66	0.090	-20
Normal favored – 17 % · 41	1.87	0.108	-4

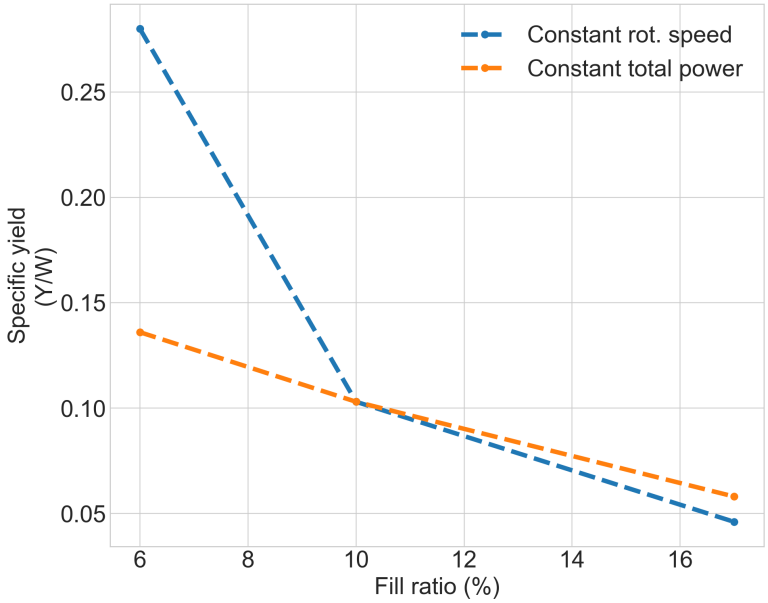


Figure 4.4: Specific yield (W^{-1}) versus fill ratio for experiments run at constant rotational speed (blue) and at constant total mechanical power (orange).

the mid-80% range, further gains become increasingly expensive. Raising the yield from 84% (3 *W*) to 94% (6.9 *W*) requires more than doubling the mechanical power input, and the specific yield is reduced by 55%, illustrating clear diminishing returns as the process approaches full conversion. Therefore, expressing performance as converted mass per Watt ($g \cdot W^{-1}$) (see Table 4.4) gives a direct, quantitative metric to evaluate these diminishing returns. Because the ball-to-powder ratio is constant, increasing the fill ratio processes more total reactant mass per batch, so even if a higher fill ratio yields a lower percentage conversion, the mass converted per Watt can still be greater. Framing conversion this way automatically penalizes those marginal, last percentage point gains that demand disproportionately more power; this can be visualized in Figure 4.5.

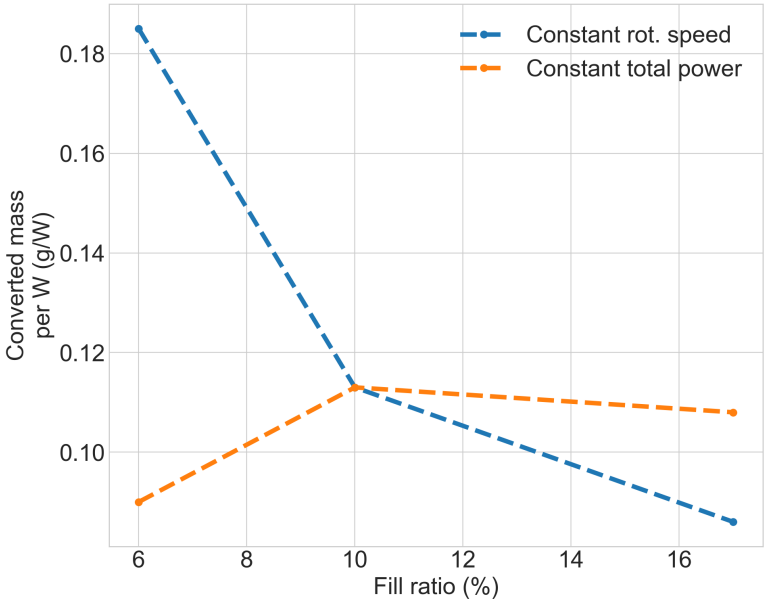


Figure 4.5: Converted mass per W (g/W) versus fill ratio for experiments run (i) at constant rotational speed (blue) and (ii) at constant total power (orange).

While the constant rotational speed tangential run retains the highest conversion, its advantage shrinks to 64% compared to the reference case once the smaller batch mass is accounted for (see Table 4.4). Critically, in the constant mechanical power approach, the normal mode run overtakes its tangential counterpart on a mass per Watt basis (0.108 vs. 0.090 $g \cdot W^{-1}$), and both perform worse than the balanced mode at 0.113 $g \cdot W^{-1}$. This result emphasizes a trade-off typically ignored in mechanochemical processes: while the net conversion efficiency reaches its peak (94%) under these tangential-rich conditions, the mechanical energetic cost to achieve the conversion renders the mass per Watt less attractive. The choice of operational conditions, therefore, depends on whether the goal is to maximize absolute percentage conversion or mass per Watt conversion.

4.3.2. Mechanochemical energy leverage

By quantifying the mechanical energy dissipated during the mechanochemical process, we can estimate the theoretical maximum of recaptured chemical energy in the regenerated $NaBH_4$, and the liberated H_2 (see Equation 4.2). For this, we define mechanochemical energy leverage (λ_{mc}) in Equation 4.7:

$$\lambda_{mc} = \frac{E_{\text{chem}}}{E_{\text{mech}}} \tag{4.7}$$

where E_{mech} is the available mechanical energy defined as $P_{tot} \cdot t_{mill}$, with P_{tot} calculated via Equation 4.5 and t_{mill} the milling time (in our case 12.5 hours). E_{chem} is the chemical energy associated with $NaBH_4$ produced, evaluated on three baselines: the reaction enthalpy ΔH (-1226 kJ/mol), the Gibbs free energy ΔG (-1510 kJ/mol at 298 K), or the lower heating value LHV of H_2 available per cycle (2908 kJ/mol). Note that because λ_{mc} is referenced to the mechanical power dissipated inside the jar, it will overestimate the wall-plug efficiency. A formal techno-economic analysis is outside the scope of this study; nonetheless, the corresponding wall-plug specific energy intensity ($\bar{E}_{\text{plug}}$) and cost per kilogram of $NaBH_4$ (C_{NaBH_4}) can be estimated with Equation 4.8 and Equation 4.9.

$$\bar{E}_{\text{plug}}\,[\text{kWh kg}^{-1}] = \frac{E_{mech}}{\eta_{\text{plug}}\, m_{NaBH_4}}, \tag{4.8}$$

$$C_{NaBH_4}\,[€\ \text{kg}^{-1}] = \bar{E}_{\text{plug}}\, p_{\text{elec}}, \tag{4.9}$$

where m_{NaBH_4} is the mass produced (kg), η_{plug} the electromechanical efficiency, and p_{elec} the electricity price (€/kWh).

The three comparison baselines form an energy quality hierarchy. The reaction enthalpy represents low-grade heat recoverable at room temperature; the Gibbs free energy is 1.23 × larger because it includes the maximum reversible work associated with the entropy gain, and the LHV is 2.37 × larger than $|\Delta H|$ because it accounts for the usable fuel energy stored in the liberated hydrogen and the regenerated $NaBH_4$. With E_{mech} fixed, λ_{mc} scales by the same factors (see Table 4.5). This shows how much of the invested work can be exploited as the process moves from heat recovery towards full fuel utilization.

Table 4.5: Mechanochemical energy leverage λ_{mc} for each milling condition.

Operating conditions	$\lambda_{mc}(\Delta H)$ (%)	$\lambda_{mc}(\Delta G)$ (%)	λ_{mc}(LHV) (%)
Tangential favored – 6 % · 3 W	3.65	4.49	8.66
Reference balanced – 10 % · 6.9 W	2.24	2.75	5.30
Normal favored – 17 % · 9.4 W	1.69	2.08	4.00
Tangential favored – 6 % · 6.9 W	1.78	2.19	4.21
Normal favored – 17 % · 6.9 W	2.14	2.64	5.08

As expected, the ranking of operating conditions coincides with the mass per Watt conversion ranking. The mechanochemical energy leverage values reveal that only 1.7–3.7% of the input mechanical work is recovered when evaluated on an enthalpy basis, 2.1–4.4% on a Gibbs free energy basis, and 4–8.7% when considering the fuel value of all available hydrogen (both directly liberated H_2 and that stored in $NaBH_4$). The best performance across all metrics is achieved under tangential milling at low power. While these LHV efficiencies account for the total hydrogen yield, they represent theoretical maxima assuming costless extraction of hydrogen from $NaBH_4$ via hydrolysis. In practice, the energy costs associated with hydrolysis will reduce these leverage values below the reported range, making the actual recoverable energy storage efficiency lower than these estimates indicate.

For context, mechanocatalytic water splitting over NiO achieves ≈ 4% enthalpy-based efficiency [41], whereas planetary ball mill reduction of water by metallic Ti consumes 1.72 kWh m^{-3} [42]. That equates to ≈ 52 g kW^{-1} h^{-1}, or 0.052 $g \cdot W^{-1}$ on a one-hour basis, which is two-to-four times lower than the 0.090-0.185 $g \cdot W^{-1}$ obtained here for $NaBH_4$ regeneration. While the processes are different, the scarcity of performance data for $NaBH_4$ regeneration makes any deeper comparison impossible.

4.3.3. Conclusions

In this chapter, we present a methodology to steer the contributions of tangential (shear) and normal (compressive) loading events during mechanochemical ball milling, specifically for the regeneration of $NaBH_4$. By combining DEM-derived mechanical descriptors with two complementary experimental approaches, one at constant rotational speed and one at constant total power, we separate changes in the energy dissipation distribution between tangential and normal modes from changes in overall collision dynamics. Because the dissipated energy in each mode scales with the corresponding contact forces, this distribution provides a practical and quantitative representation for the underlying stress landscape. Benchmarking each scenario through the specific yield (W^{-1}), the converted mass per Watt ($g \cdot W^{-1}$), and the mechanochemical energy leverage (λ_{mc}) offers relevant metrics for optimization and scale-up.

A low fill ratio of 6%, which maximizes tangential dissipation, is the clear optimum across both approaches if absolute percentage conversion is to be maximized. Under constant-power operation, it delivers a record 94% regeneration yield while reducing milling time by 38%, the ball-to-powder ratio by 40%, and the rotational speed by 34% compared to prior literature. The same tangential-rich state achieved at around half the power (3 W) attains 84% yield, giving a specific yield of 0.28 W^{-1}, five times the normal-dominant cases. Yield falls linearly with fill ratio ($R^2 > 0.99$), with gradients of 0.038 per percentage point at constant speed and 0.049 per point at constant power; thus, raising the fill from 6% to 17% costs 0.4–0.5 of fractional yield (40–50 percentage points) regardless of experimental approach. Together, these trends indicate that shear-dominated contact mechanics, rather than normal impacts, govern the rate-limiting interfacial chemistry. To mechanistically ground this, Density Functional Theory (DFT) derived energetics for the elementary steps could seed a microkinetic

model that quantitatively links mechanical stress components (normal vs tangential) to interfacial reaction rates, enabling prediction of optimal operating windows.

Our analysis further reveals that while tangential bias maximizes absolute conversion yield, mechanical energy cost and BPR diminish per-Watt mass conversion efficiency. Under constant mechanical power conditions, the balanced regime (10% fill) yields the highest converted mass per Watt (0.113 $g \cdot W^{-1}$) compared to both tangential (0.090 $g \cdot W^{-1}$) and normal (0.108 $g \cdot W^{-1}$) biased cases. This trade-off highlights that the optimal operating point depends on the prioritization of absolute conversion efficiency or normalized productivity.

The mechanochemical energy leverage of the process reveals theoretical maxima of 1.7–3.7% on an enthalpy basis, 2.1–4.4% on a Gibbs free energy basis, and 4–8.7% when considering the fuel value of all available hydrogen. These values indicate that the majority of input mechanical work is dissipated as heat and deformation rather than being converted to useful chemical energy. This inefficiency suggests opportunities for improving the process via thermal management, catalytic additives, or alternative reactor geometries that better harness mechanical work for chemical activation.

Taken together, these findings demonstrate the critical role of stress mode partitioning in mechanochemical reactions. Because the three mechanical descriptors $\{\bar{E}_n, \bar{E}_t, f_{\text{col}}/n_{\text{ball}}\}$ are transferable between mills, any reactor that reproduces the tangential-rich scenarios identified here is expected to achieve comparable yields, provided other thermochemical factors (jar atmosphere, contamination, etc.) are controlled. By targeting desired mechanical conditions, practitioners can rationally design milling protocols to maximize yield, throughput, and energy efficiency.

References

[1] A. A. K. Erin S. van Rheenen Johan T. Padding and K. Visser. 'Comparative energy analysis of hydrogen carriers as energy source on ships'. In: *Journal of Marine Engineering & Technology* 0.0 (2025), pp. 1–15. DOI: 10.1080/20464177.2024.2448057.

[2] E. Y. Marrero-Alfonso, A. M. Beaird, T. A. Davis and M. A. Matthews. 'Hydrogen Generation from Chemical Hydrides'. In: *Industrial & Engineering Chemistry Research* 48.8 (2009), pp. 3703–3712. DOI: 10.1021/ie8016225. eprint: https://doi.org/10.1021/ie8016225. URL: https://doi.org/10.1021/ie8016225.

[3] U. Demirci, O. Akdim and P. Miele. 'Ten-year efforts and a no-go recommendation for sodium borohydride for on-board automotive hydrogen storage'. In: *International Journal of Hydrogen Energy* 34.6 (2009), pp. 2638–2645. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2009.01.038. URL: https://www.sciencedirect.com/science/article/pii/S0360319909000603.

[4] Y. Zhu, L. Ouyang, H. Zhong, J. Liu, H. Wang, H. Shao, Z. Huang and M. Zhu. 'Closing the Loop for Hydrogen Storage: Facile Regeneration of NaBH4 from its Hydrolytic Product'. In: *Angewandte Chemie International Edition* 59.22 (2020), pp. 8623–8629. DOI: https://doi.org/10.1002/anie.201915988. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.201915988. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.201915988.

[5] H. Zhong, L. Ouyang, J. Ye, J. Liu, H. Wang, X. Yao and M. Zhu. 'An one-step approach towards hydrogen production and storage'. In: *Energy Storage Materials* 7 (Mar. 2017). DOI: 10.1016/j.ensm.2017.03.001.

[6] H. X. Nunes, D. L. Silva, C. M. Rangel and A. M. F. R. Pinto. 'Rehydrogenation of Sodium Borates to Close the NaBH4-H2 Cycle: A Review'. In: *Energies* 14.12 (2021). ISSN: 1996-1073. DOI: 10.3390/en14123567. URL: https://www.mdpi.com/1996-1073/14/12/3567.

[7] H. N. Abdelhamid. 'A review on hydrogen generation from the hydrolysis of sodium borohydride'. In: *International Journal of Hydrogen Energy* 46.1 (2021), pp. 726–765. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2020.09.186. URL: https://www.sciencedirect.com/science/article/pii/S0360319920336260.

[8] H. Zhong, L. Ouyang, M. Zeng, J. Liu, H. Wang, H. Shao, M. Felderhoff and M. Zhu. 'Realizing facile regeneration of spent NaBH4 with Mg–Al alloy'. In: *J. Mater. Chem. A* 7 (17 2019), pp. 10723–10728. DOI: 10.1039/C9TA00769E. URL: http://dx.doi.org/10.1039/C9TA00769E.

[9] M. Huang, H. Zhong, L. Ouyang, C. Peng, X. Zhu, W. Zhu, F. Fang and M. Zhu. 'Efficient regeneration of sodium borohydride via ball milling dihydrate sodium metaborate with magnesium and magnesium silicide'. In: *Journal of Alloys and Compounds* 729 (2017), pp. 1079–1085. ISSN: 0925-8388. DOI: https://doi.org/10.1016/j.jallcom.2017.09.262. URL: https://www.sciencedirect.com/science/article/pii/S0925838817333121.

[10] C.-L. Hsueh, C.-H. Liu, B.-H. Chen, C.-Y. Chen, Y.-C. Kuo, K.-J. Hwang and J.-R. Ku. 'Regeneration of spent-NaBH 4 back to NaBH 4 by using high-energy ball milling'. In: *International Journal of Hydrogen Energy - INT J HYDROGEN ENERG* 34 (Feb. 2009), pp. 1717–1725. DOI: 10.1016/j.ijhydene.2008.12.036.

[11] L. Kong, C. Xinyu, H. Jin, J. Wu, H. Du and T. Xiong. 'Mechanochemical Synthesis of Sodium Borohydride by Recycling Sodium Metaborate'. In: *Energy & Fuels* 23 (Sept. 2009). DOI: 10.1021/ef900619y.

[12] C. Cakanyildirim and M. Gürü. 'Processing of NaBH4 from NaBO2 with MgH2 by ball milling and usage as hydrogen carrier'. In: *Renewable Energy* 35 (Sept. 2010), pp. 1895–1899. DOI: 10.1016/j.renene.2010.01.001.

[13] C. Lang, Y. Jia, J. Liu, H. Wang, L. Ouyang, M. Zhu and X. Yao. 'NaBH 4 regeneration from NaBO 2 by high-energy ball milling and its plausible mechanism'. In: *International Journal of Hydrogen Energy* 42 (Apr. 2017). DOI: 10.1016/j.ijhydene.2017.04.014.

[14] L. Ouyang, W. Chen, J. Liu, M. Felderhoff, H. Wang and M. Zhu. 'Enhancing the Regeneration Process of Consumed NaBH 4 for Hydrogen Storage'. In: *Advanced Energy Materials* 7 (June 2017), p. 1700299. DOI: 10.1002/aenm.201700299.

[15] W. Chen, L. Ouyang, J. Liu, X. Yao, H. Wang, Z. Liu and M. Zhu. 'Hydrolysis and regeneration of sodium borohydride (NaBH 4) – A combination of hydrogen production and storage'. In: *Journal of Power Sources* 359 (Aug. 2017), pp. 400–407. DOI: 10.1016/j.jpowsour.2017.05.075.

[16] S. Shen and L. Ouyang. 'From Borate to Borohydride: Achieving a High-Yield Synthesis of Magnesium Borohydride Enabled by Cation Tuning'. In: *ACS Sustainable Chemistry & Engineering* 13.7 (2025), pp. 2933–2942. DOI: 10.1021/acssuschemeng.4c09817. eprint: https://doi.org/10.1021/acssuschemeng.4c09817. URL: https://doi.org/10.1021/acssuschemeng.4c09817.

[17] K. Chen, H. Zhong, L. Ouyang, F. Liu, H. Wang, J. Liu, H. Shao and M. Zhu. 'Achieving a novel solvent-free regeneration of LiBH4 combining hydrogen storage and production in a closed material cycle'. In: *Journal of Magnesium and Alloys* 11.5 (2023), pp. 1697–1708. ISSN: 2213-9567. DOI: https://doi.org/10.1016/j.jma.2021.08.005. URL: https://www.sciencedirect.com/science/article/pii/S2213956721002036.

[18] K. Chen, L. Ouyang, H. Zhong, J. Liu, H. Wang, H. Shao, Y. Zhang and M. Zhu. 'Converting H+ from coordinated water into H− enables super facile synthesis of LiBH4'. In: *Green Chem.* 21 (16 2019), pp. 4380–4387. DOI: 10.1039/C9GC01897B. URL: http://dx.doi.org/10.1039/C9GC01897B.

[19] L. Ouyang, J. Jiang, K. Chen, M. Zhu and Z. Liu. 'Hydrogen Production via Hydrolysis and Alcoholysis of Light Metal-Based Materials: A Review'. In: *Nano-Micro Letters* 13.1 (2021), p. 134. DOI: 10.1007/s40820-021-00657-9. URL: https://doi.org/10.1007/s40820-021-00657-9.

[20] H. Zhong, K. Chen, C. Qin, C. Lang, J. Liu, H. Wang, J. Zhang and L. Ouyang. 'Enhancing NaBH4 regeneration using an Al-rich alloy'. In: *Journal of Alloys and Compounds* 976 (2024), p. 173160. ISSN: 0925-8388. DOI: https://doi.org/10.1016/j.jallcom.2023.173160. URL: https://www.sciencedirect.com/science/article/pii/S0925838823044638.

[21] S. Garrido Nuñez, D. L. Schott and J. T. Padding. 'Optimization of operational parameters in the mechanochemical regeneration of sodium borohydride (NaBH4)'. In: *International Journal of Hydrogen Energy* 97 (2025), pp. 640–648. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2024.11.360. URL: https://www.sciencedirect.com/science/article/pii/S0360319924050511.

[22] J. Batteas, K. G. Blank, E. Colacino, F. Emmerling, T. Friščić, J. Mack, J. Moore, M. E. Rivas and W. Tysoe. 'Moving mechanochemistry forward'. In: *RSC Mechanochem.* 2 (1 2025), pp. 10–19. DOI: 10.1039/D4MR90021A. URL: http://dx.doi.org/10.1039/D4MR90021A.

[23] S. V. Sukhomlinov, G. Kickelbick and M. H. Müser. 'Mechanochemical Ionization: Differentiating Pressure-, Shear-, and Temperature-Induced Reactions in a Model Phosphate'. In: *Tribology Letters* 70.4 (Aug. 2022), p. 102. ISSN: 1573-2711. DOI: 10.1007/s11249-022-01644-w. URL: https://doi.org/10.1007/s11249-022-01644-w.

[24] H. Kobayashi, Y. Suzuki, T. Sagawa, K. Kuroki, J.-y. Hasegawa and A. Fukuoka. 'Impact of tensile and compressive forces on the hydrolysis of cellulose and chitin'. In: *Phys. Chem. Chem. Phys.* 23 (30 2021), pp. 15908–15916. DOI: 10.1039/D1CP01650D. URL: http://dx.doi.org/10.1039/D1CP01650D.

[25] R. Rana, N. Hopper, F. Sidoroff and W. T. Tysoe. 'Critical stresses in mechanochemical reactions'. In: *Chem. Sci.* 13 (43 2022), pp. 12651–12658. DOI: 10.1039/D2SC04000J. URL: http://dx.doi.org/10.1039/D2SC04000J.

[26] F. H. Bhuiyan, Y.-S. Li, S. H. Kim and A. Martini. 'Shear-activation of mechanochemical reactions through molecular deformation'. In: *Scientific Reports* 14.1 (Feb. 2024), p. 2992. ISSN: 2045-2322. DOI: 10.1038/s41598-024-53254-2. URL: https://doi.org/10.1038/s41598-024-53254-2.

[27] L. Fang, S. Korres, W. A. Lamberti, M. N. Webster and R. W. Carpick. 'What stress components drive mechanochemistry? A study of ZDDP tribofilm formation'. In: *Faraday Discussions* 241 (2023), p. 394. DOI: 10.1039/D2FD00123C. URL: https://pubs.rsc.org/en/content/articlelanding/2023/FD/D2FD00123C.

[28] J. L. Howard, Q. Cao and D. L. Browne. 'Mechanochemistry as an emerging tool for molecular synthesis: what can it offer?' In: *Chemical Science* 9 (2018). Open Access, pp. 3080–3094. DOI: 10.1039/C7SC05371A. URL: https://pubs.rsc.org/en/content/articlehtml/2018/sc/c7sc05371a.

[29] S. Rosenkranz, S. Breitung-Faes and A. Kwade. 'Experimental investigations and modeling of the ball motion in planetary ball mills'. In: *Powder Technology* 212 (Sept. 2011), pp. 224–230. DOI: 10.1016/j.powtec.2011.05.021.

[30] C. Burmeister, M. Hofer, P. Molaiyan, P. Michalowski and A. Kwade. 'Characterization of Stressing Conditions in a High Energy Ball Mill by Discrete Element Simulations'. In: *Processes* 10 (Apr. 2022), p. 692. DOI: 10.3390/pr10040692.

[31] C. Burmeister, R. Schmidt, K. Jacob, S. Breitung, A. Stolle and A. Kwade. 'Effect of stressing conditions on mechanochemical Knoevenagel synthesis'. In: *Chemical Engineering Journal* 396 (Feb. 2020), p. 124578. DOI: 10.1016/j.cej.2020.124578.

[32] C. F. Burmeister, A. Stolle, R. Schmidt, K. Jacob, S. Breitung-Faes and A. Kwade. 'Experimental and Computational Investigation of Knoevenagel Condensation in Planetary Ball Mills'. In: *Chemical Engineering & Technology* 37.5 (2014), pp. 857–864. DOI: https://doi.org/10.1002/ceat.201300738. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ceat.201300738. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/ceat.201300738.

[33] A. Krusenbaum, S. Grätz, G. T. Tigineh, L. Borchardt and J. G. Kim. 'The mechanochemical synthesis of polymers'. In: *Chemical Society Reviews* 51.7 (2022), pp. 2873–2905. DOI: 10.1039/d1cs01093j. URL: https://doi.org/10.1039/d1cs01093j.

[34] S. Pagola. 'Outstanding Advantages, Current Drawbacks, and Significant Recent Developments in Mechanochemistry: A Perspective View'. In: *Crystals* 13.1 (2023). ISSN: 2073-4352. URL: https://www.mdpi.com/2073-4352/13/1/124.

[35] R. T. O'Neill and R. Boulatov. 'The many flavours of mechanochemistry and its plausible conceptual underpinnings'. In: *Nature Reviews Chemistry* 5.3 (2021), pp. 148–167. ISSN: 2397-3358. DOI: 10.1038/s41570-020-00249-y. URL: https://doi.org/10.1038/s41570-020-00249-y.

[36] S. Garrido Nuñez, D. L. Schott and J. T. Padding. 'Predictive models for energy dissipation in mechanochemical ball milling'. In: *Powder Technology* 457, 120919 (2025). ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2025.120919. URL: https://www.sciencedirect.com/science/article/pii/S0032591025003146.

[37] O. Jafter, S. Lee, J. Park, C. Cabanetos and D. Lungerich. 'Navigating Ball Mill Specifications for Theory-to-Practice Reproducibility in Mechanochemistry'. In: *Angewandte Chemie International Edition* 63 (Oct. 2024), e202409731. DOI: 10.1002/anie.202409731.

[38] W. Pickhardt, S. Grätz and L. Borchardt. 'Direct Mechanocatalysis: Using Milling Balls as Catalysts'. In: *Chemistry – A European Journal* 26.57 (2020), pp. 12903–12911. DOI: https://doi.org/10.1002/chem.202001177. eprint: https://chemistry-europe.onlinelibrary.wiley.com/doi/pdf/10.1002/chem.202001177. URL: https://chemistry-europe.onlinelibrary.wiley.com/doi/abs/10.1002/chem.202001177.

[39] L. Ouyang, H. Zhong, H.-W. Li and M. Zhu. 'A Recycling Hydrogen Supply System of NaBH4 Based on a Facile Regeneration Process: A Review'. In: *Inorganics* 6.1 (2018). ISSN: 2304-6740. DOI: 10.3390/inorganics6010010. URL: https://www.mdpi.com/2304-6740/6/1/10.

[40] J. Lyu, A. Lider and V. Kudiiarov. 'Using Ball Milling for Modification of the Hydrogenation/Dehydrogenation Process in Magnesium-Based Hydrogen Storage Materials: An Overview'. In: *Metals* 9.7 (2019). ISSN: 2075-4701. DOI: 10.3390/met9070768. URL: https://www.mdpi.com/2075-4701/9/7/768.

[41] K. Domen, J. N. Kondo, M. Hara and T. Takata. 'Photo- and Mechano-Catalytic Overall Water Splitting Reactions to Form Hydrogen and Oxygen on Heterogeneous Catalysts'. In: *Bulletin of the Chemical Society of Japan* 73.6 (Feb. 2001), pp. 1307–1331. ISSN: 0009-2673. DOI: 10.1246/bcsj.73.1307. eprint: https://academic.oup.com/bcsj/article-pdf/73/6/1307/56193956/bcsj.73.1307.pdf. URL: https://doi.org/10.1246/bcsj.73.1307.

[42] T. Yamamoto, S. Ashida, N. Inubuse, S. Shimizu, Y. Miura, T. Mizutani and K.-i. Saitow. 'Room-temperature thermochemical water splitting: efficient mechanocatalytic hydrogen production'. In: *J. Mater. Chem. A* 12 (44 2024), pp. 30906–30918. DOI: 10.1039/D4TA04650A. URL: http://dx.doi.org/10.1039/D4TA04650A.

5

Linking mechanics and chemistry: Machine learning for yield prediction

The results of the previous chapters have shown that the design space of a mechanochemical process is large, interactions are non-linear, and internal dynamics can be summarized by mechanical descriptors related to machine-level variables. However, chemical and mechanical components remain uncoupled so far. Because exhaustive testing of every condition is costly and time-consuming, we couple those descriptors with reaction stoichiometry to build data-efficient machine learning (ML) models that predict $NaBH_4$ yield, aid interpretability, and quantify uncertainty. This work establishes a framework for using ML to optimize mechanochemical processes, reducing experimental cost and offering a method to link mechanical conditions to chemical outcomes, thereby enabling predictive mechanochemistry.

Parts of this article are adapted from: Garrido Nuñez, S.; Schott, D. L.; Padding, J. T. “Linking mechanics and chemistry: machine learning for yield prediction in $NaBH_4$ mechanochemical regeneration”. In RSC Mechanochemistry (2025).

5.1. Introduction

5.1.1. Mechanochemical reactions via high-energy ball milling

The advancement of mechanochemistry in the last two decades has seen the application and innovation of multiple tools and processes to achieve chemical and material synthesis that align with the principles of green chemistry [1]. Typically, mechanochemical processes at the lab scale rely on ball mills to supply the (mechanical) energy required to achieve a desired chemical reaction, although different methods have been explored to combine this with additional sources of energy, such as thermal energy, acoustic energy, or electrical energy [2]. Pure mechanochemical ball milling is often characterized by intuitive process parameters that any ball mill can readily account for, namely rotational speed, filling ratio, ball-to-powder ratio (BPR), milling time, and additional physical material properties such as density of the milling balls [3–7]. Although these parameters certainly steer the overall behavior of the process, it has been observed that they are not sufficient to accurately characterize mechanochemical processes, leading to significant challenges in reproducibility and scaling up given the intrinsic differences in working principle that different machines have [8, 9].

It becomes clear that mechanochemistry involves a series of complex interactions that must be investigated systematically before layering on additional, non-intuitively tunable energy inputs, especially because both mechanical and chemical variables fundamentally dictate high yields. However, due to the relative novelty of the field, research has remained largely exploratory, employing one-variable-at-a-time (OVAT) studies that prove inadequate once scale-up or efficiency optimization becomes the goal [10–12].
To tackle this challenge, the Discrete Element Method (DEM) has been employed to accurately characterize a high-energy ball mill's internal dynamics, effectively bypassing the dependency of the utilized mill or the aforementioned process variables. This is done by defining three key mechanical characterization properties: the mean normal energy dissipation per collision $\bar{E}_n$, the mean tangential energy dissipation per collision $\bar{E}_t$, and the specific collision frequency per ball $\frac{f_{col}}{n_{ball}}$ [8]. This methodology can be applied to any milling machine of any scale, reducing the challenges in reproducibility and providing guidelines for the specifications needed in larger-scale equipment. Regardless, this numerical characterization remains mechanical and thus cannot include the influence of the chemical variables of the system, such as the molar ratio, BPR, milling time, and their confounded influence with the rest of the mechanical variables.

The chemical characterization of the system can only be accomplished experimentally. In our target reaction, the mechanochemical regeneration of $NaBH_4$ from $NaBO_2 \cdot 4\,H_2O$ (see Equation 5.1), the dependence on molar ratio is non-linear, while the influence of milling time is effectively linear within the investigated range [3]. Furthermore, the interaction between the BPR and molar ratio is statistically significant, indicating a complex interaction among operating parameters. These experiments were carried out under constant, albeit optimized mechanical conditions from the pure

perspective of energy dissipation, ignoring the effect that changing the distribution of shear and normal stress can have on the system [8]. Thus, while DEM simulations can facilitate a mechanical characterization, their effectiveness can only be tested experimentally. To overcome this limitation, we investigate the use of different machine learning (ML) algorithms to predict the conversion yield, reducing the need for trial-and-error experiments.

$$NaBO_2 \cdot 4\,H_2O + 6\,MgH_2 \rightarrow NaBH_4 + 6\,MgO + 8\,H_2 \tag{5.1}$$

In other applications, Anglou et al. [13] employed linear regression to link DEM outputs (collision frequency and average kinetic energy of a milling ball) to the depolymerization of PET, obtaining a good fit (R^2 = 0.966). This result, however, holds only within a range of total energy given to the system before the linear condition is lost. The authors accurately point out that a non-linear model could be trained, but the lack of data prohibits this. Furthermore, this study made use of a single milling ball in a 25 mL jar where only the milling frequency was varied. This configuration effectively simplifies many other operational parameters that lab-scale and industrial-scale ball milling processes can have.

Similarly, Yu et al. [14] utilized polynomial regression to analyze different milling parameters and predict target particle sizes while ball milling alumina ceramics. Although no chemical processes were involved in this study, the authors point out the same challenge mentioned before: most studies focus on optimizing milling parameters, varying only one process variable and keeping the rest constant, which severely limits the applicability of data-based methods to gain a more profound understanding of their impact on the process. In the same context of pure milling, Lit et al. [15] trained a convolutional neural network (CNN) to predict the grinding rate and size distribution of a rotating drum mill, achieving high accuracy (R^2 > 0.95) and good transfer learning results. This indicates that deep neural networks can capture the complex physics of milling when sufficient training data exists.
However, deep models such as these remain unviable when applied directly to experimental mechanochemical data, simply because such large, labeled datasets do not exist. Furthermore, the creation of these datasets requires extensive experimental work that necessitates a significant amount of time. For instance, a typical experiment involving the regeneration of $NaBH_4$ takes at least 72 hours from sample preparation to yield quantification.

To address these issues, a shared mechanochemical reaction database has been created [16], allowing researchers to pool results and push machine learning approaches that connect milling conditions with chemical outcomes, something experts believe could revolutionize the field [17]. However, because different groups study different reactions in different mills, detailed data for any one process remain scarce, and most characterization methods only work on the specific equipment for which they were developed as stated before. Moreover, the ”black-box” nature of ML models adds another limitation. While some methods, like random forests, offer feature-importance insights, other methods, like deep neural networks and support-vector machines, hardly explain why a given parameter set succeeds or fails. This makes it hard to

build a mechanistic understanding or plan experiments beyond the model's training scope. Furthermore, while previous applications of machine learning in milling have often focused on either purely physical outcomes (e.g., particle size prediction) or utilized limited operational parameters for chemical yield, a comprehensive approach integrating detailed, DEM-derived mechanical descriptors with a broader set of chemical process variables to predict yield for complex reactions remains unexplored. Finally, practical challenges persist: producing large, high-quality datasets demands extensive experimentation, run-to-run variability can introduce noise, and fitting sensors inside a sealed milling jar to gather real-time data is technically difficult [18–20]. Altogether, these five factors keep ML-driven ball milling mechanochemistry at a very early stage.

Within the broader landscape of data-driven reaction discovery and optimization, machine learning has not only accelerated condition search but also changed how chemists learn from experiments. In solution-phase synthesis, high-throughput experimentation (HTE) and automation provide the dense, standardized datasets that enable multivariate modeling and closed-loop optimization [21, 22]. Multivariate linear models extract quantitative structure–reactivity/selectivity relationships that rank which variables matter and why, enabling prospective design [23]. Orchestration and active-learning platforms (e.g., ChemOS; LabMate.ML) close the loop between Bayesian decision-making and automated execution, reaching high-yielding conditions in tens of experiments while handling mixed categorical/continuous spaces [24, 25]. Beyond single substrates, closed-loop protocols now optimize for generality across substrate matrices, identifying condition sets that transfer across chemotypes [26]. Recent systems show that optimization can produce knowledge on-the-fly, integrating interpretable/physics-informed models with automation to uncover mechanistic factors during optimization [27, 28]. These developments motivate our study, but also highlight two distinctions specific to mechanochemistry: data throughput is typically much lower than in plate- or flow-based solution platforms [21], and controllable variables necessarily include mechanical/process descriptors of mechanical stressing and energy transfer, which are absent from most solution phase models [8, 29].

Motivated by recent ML-driven progress in solution-phase optimization, we lay the groundwork for a mechanochemistry-specific framework. We take advantage of a DEM-based mechanical characterization that establishes a commonality between mills, enabling unified datasets. We compare modeling families and map their strengths and limitations to use cases in small-data, high-cost regimes. Ultimately, we show that combining mechanical and chemical operating variables can accurately predict the mechanochemical yield. The dataset spans 27 experiments with wide ranges in both chemical and mechanical factors and, although compact, constitutes the most extensive open-access operating space for $NaBH_4$ regeneration to date, positioning this study as a practical starting point for predictive mechanochemistry.

5.2. Methodology

This section details the methodology employed to predict the experimental yield using machine learning techniques. The workflow encompasses data acquisition, feature engineering, model training, hyperparameter optimization, and evaluation. All analyses were performed using Python 3.9.

5.2.1. Data acquisition

The dataset (Table 5.1) utilized in this work combines two previously published components: experimental yields for regeneration of $NaBH_4$ using the Emax high-energy ball mill [3], and a DEM-based methodology to mechanically characterize ball milling conditions [8]. In the present study, we derive device-independent descriptors for all experimental cases and assemble an ML-ready dataset that supports comparison and transfer across ball-milling devices. This is achieved by defining three key parameters: the mean normal energy dissipation per collision $\bar{E}_n$, the mean tangential energy dissipation per collision $\bar{E}_t$, and the specific collision frequency per ball $\frac{f_{col}}{n_{ball}}$. We note that the variables modeled in solution phase yield prediction studies typically comprise solvent, base, ligand/catalyst, temperature, concentrations, and time, often explored at scale via HTE or flow with inline analytics. In mechanochemistry, outcome-relevant variables also include mill type, jar/ball materials and sizes, ball-to-powder ratio, fill ratio, and milling frequency, and thus require abstraction via the aforementioned mechanical descriptors of energy transfer to compare between devices. The results presented in our previous work can be readily used to arrive at these key parameters in the Emax, but the methodology can be applied to any ball mill [8].

Experimentally, hydrated sodium metaborate ($NaBO_2 \cdot 4\,H_2O$) (≥ 99%) was purchased from Sigma-Aldrich, while magnesium hydride (MgH_2) (≥ 99.9%, ≤ 50μm) was sourced from Nanoshel. All reactants were used without further purification. The sample preparation for all ball milling experiments was carried out in a glove box under an argon atmosphere, with oxygen and water concentrations maintained below 0.1 ppm. For a detailed description of the quantification of the chemical yield and equipment cleaning to preserve similar conditions for all experimental cases, we refer to our previous work [3].

5.2.2. Feature engineering

To facilitate the capture of non-linear relationships and interactions, several feature engineering steps were applied to the initial feature set:

1. **Quadratic Term:** Our previous results indicate that the molar ratio has a significant quadratic relationship with the chemical yield [3]. Thus, a new feature containing this quadratic term was added.

2. **Interaction Term:** We have found a strong interaction between the BPR and the molar ratio [3]. Therefore, we also include an additional feature composed of the product of these 2 variables.

Table 5.1: $NaBH_4$ regeneration dataset. BPR = ball–powder mass ratio; "Mol. ratio" = $NaBH_4$: MgH_2; $\bar{E}_n$, $\bar{E}_t$ are mean normal and tangential energy dissipated per collision from DEM; $f_{\mathrm{col}}/n_{\mathrm{ball}}$ is the specific collision-frequency per ball. Experimental details: [3]; DEM details: [8].

Case	BPR	Mol. ratio	Time (h)	$\bar{E}_n$ (µJ)	$\bar{E}_t$ (µJ)	$f_{\mathrm{col}}/n_{\mathrm{ball}}$ (s^{-1})	Yield (%)
0	10	8	5.0	221	500	400	12
1	10	8	12.5	382	888	533	22
2	10	8	20.0	613	1391	667	30
3	10	10	5.0	221	500	400	28
4	10	10	12.5	613	1391	667	39
5	10	10	20.0	221	500	400	45
6	10	12	5.0	613	1391	667	40
7	10	12	12.5	221	500	400	61
8	10	12	20.0	382	888	533	73
9	30	8	5.0	382	888	533	26
10	30	8	12.5	613	1391	667	37
11	30	8	20.0	221	500	400	42
12	30	10	5.0	613	1391	667	50
13	30	10	12.5	221	500	400	71
14	30	10	20.0	382	888	533	88
15	30	12	5.0	221	500	400	21
16	30	12	12.5	382	888	533	32
17	30	12	20.0	613	1391	667	49
18	50	8	5.0	613	1391	667	25
19	50	8	12.5	221	500	400	62
20	50	8	20.0	382	888	533	74
21	50	10	5.0	221	500	400	31
22	50	10	12.5	382	888	533	73
23	50	10	20.0	613	1391	667	90
24	50	12	5.0	382	888	533	41
25	50	12	12.5	613	1391	667	62
26	50	12	20.0	221	500	400	57

3. **Sigmoid Transformation:** To account for potential saturation effects, sigmoid transformations ($1/(1 + \exp(-x))$) were applied to the 'Time', and 'BPR' features, creating new features while retaining the original features. This allows the model to plateau rather than grow indefinitely (or turn negative) as these variables change.

The resulting set of features constituted the final engineered feature matrix used for model training.

5.2.3. Train-Test split and feature scaling

The dataset, comprising the engineered features and the target yield, was divided into training (80%) and testing (20%) sets. A fixed random state was used to ensure reproducibility of the split. With 27 experimental cases, this corresponds to 21 training and 6 test samples. To ensure every algorithm is evaluated on the same examples, we used a single, predetermined 21/6 partition created by shuffling once and then locking that partition for all analyses. All model fitting and hyperparameter selection used only the training data; the test set was held back until the final evaluation.

Feature scaling was applied to ensure that features with larger ranges did not disproportionately influence the model's sensitivity to feature magnitude, such as distance-based algorithms (Support Vector Regression (SVR), Gaussian Process Regression (GPR)) and linear models. Specifically, standardization was employed, where each feature was transformed to have zero mean and unit variance according to Equation 5.2:

$$x_{\text{scaled}} = \frac{x - \mu_{\text{train}}}{\sigma_{\text{train}}} \tag{5.2}$$

where x is the original feature value and μ_{train} and σ_{train} are the mean and standard deviation of that feature calculated exclusively from the training data partition.

The train-test split ensures that no information from the test set influences the transformation applied during the training phase (preventing data leakage) and preserves the integrity of the test set for unbiased model evaluation. The same training set parameters ($\mu_{\text{train}}, \sigma_{\text{train}}$) were then used to standardize the corresponding features of the training set and the test set. Models requiring scaled data (Linear Regression, GPR, SVR) utilized these standardized features for both training and prediction. In contrast, tree-based models (Random Forest, XGBoost), which are less sensitive to feature scaling, were trained and evaluated using the original, unscaled engineered features.

5.2.4. Weighted loss function

To prioritize accurate prediction of higher yields, which are often of greater experimental interest, a custom weighted mean squared error (MSE) loss function was defined:

$$\text{Weighted MSE} = \frac{1}{N}\sum_{i=1}^{N} w_i (y_{\text{true},i} - y_{\text{pred},i})^2 \tag{5.3}$$

where N is the number of samples, $y_{\text{true},i}$ and $y_{\text{pred},i}$ are the true and predicted yields for sample i, respectively. The weight w_i was set to 2.0 if $y_{\text{true},i} > 70\%$, and 1.0 otherwise. This weighted MSE was used as the primary scoring metric during hyperparameter optimization and for comparing model performance.

5.2.5. Modeling approach motivation

Informed by previous research [3], the milling time feature alone was found to account for approximately 50% of the observed variance in yield. A primary concern was that this dominant predictor could mask the influence of the remaining process parameters. To address this potential overshadowing effect and better capture the contributions of the remaining features, we implemented a specialized two-step modeling strategy. The approach involves:

1. Training a first model using only the 'time' feature to predict the yield.
2. Calculating the residuals (actual yield minus the first model's prediction) on the training data.
3. Training a second model using all other engineered features (excluding 'time') to predict these residuals.
4. The final prediction is the sum of the predictions from the time model and the residual model.

This allows the residual model to focus on explaining the yield variation not captured by the primary time trend. To rigorously assess the benefit of this specialized strategy, we also trained corresponding models directly on the full set of engineered features for direct performance comparison against their two-step counterparts. The general methodology is visualized in Figure 5.1.

5.2.6. Machine learning algorithms and hyperparameter optimization

The machine learning algorithms described below were selected on the basis of their suitability and applicability to the current state of available experimental mechanochemistry data. It is worth highlighting that the feature engineering and two-step approach described in previous sections are omitted for the linear regression model, which is included solely as a baseline for comparison. For the GPR (Gaussian Process Regression), RF (Random Forest), SVR (Support Vector Regression), and XGBoost models, hyperparameter optimization was performed using the Tree-structured Parzen Estimator (TPE) algorithm implemented in the Hyperopt library. The objective was to minimize the weighted MSE, evaluated using repeated k-fold cross-validation of the training data with $k = 5$ splits and $n = 3$ repeats. We chose $k = 5$ as a pragmatic bias–variance compromise [30]. Given 21 samples, 5-fold yields validation folds of 4–5 samples (training folds of 16–17), whereas 10-fold would leave 2–3 per validation, and leave-one-out CV only 1, both of which increase variance in hyperparameter

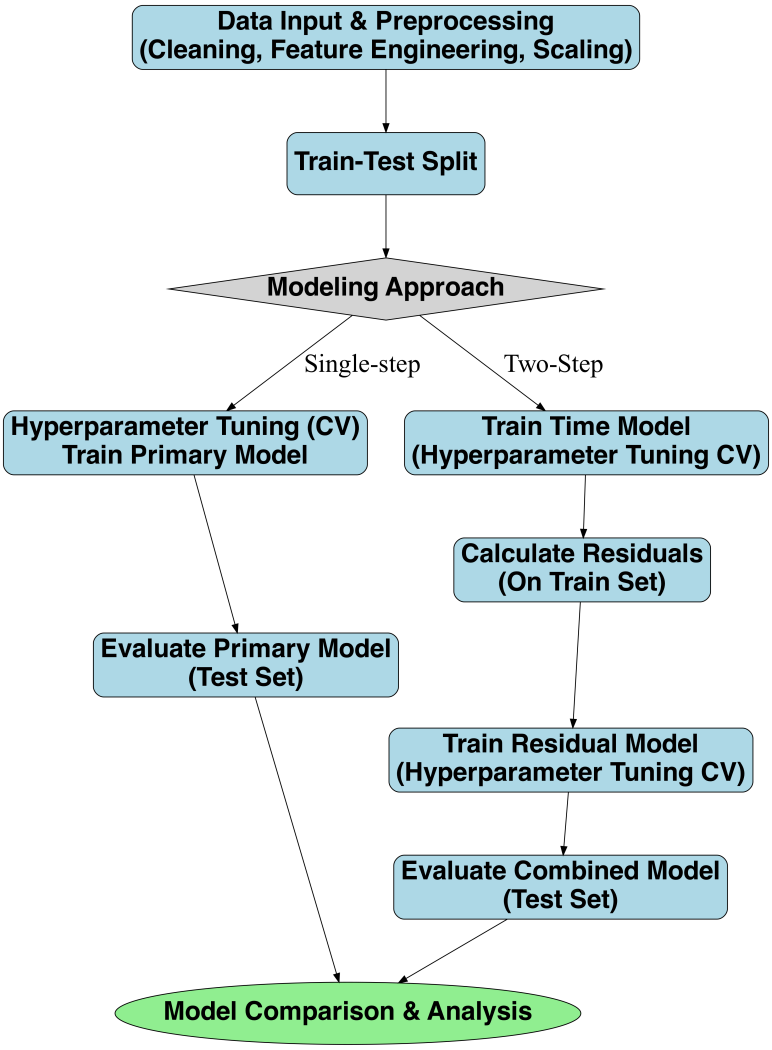


Figure 5.1: Methodology workflow overview.

comparisons. A total of 100 function evaluations were assigned for most models, while SVR, known to be potentially slower to tune, was assigned 500 evaluations to ensure a thorough search [31]. The best hyperparameters found during this process were used to train the final model on the entire training set.

Linear regression

Linear regression is included, given its simplicity and the ability to assess how well the relationship between input and output variables can be captured with a linear relationship. Linear models have been widely used to relate physical-organic descriptors to outcomes and selectivity in reaction development [23, 32]. The relationship between target Y and input variables X_i can be described as [33]:

$$Y = \beta_0 + \sum_{i=1}^{p} X_i \beta_i \tag{5.4}$$

Gaussian Process Regression (GPR)

Gaussian Process Regression (GPR) is a supervised learning method that models a distribution over possible functions rather than fitting a single function directly [34]. At its core, GPR assumes that any set of observed points is drawn from a joint Gaussian distribution characterized by a mean function and a covariance (kernel) function. The kernel function defines how closely related any two points are, which in turn governs the smoothness and shape of the functions in the model. Thus, GPR is an attractive alternative, as it produces not only a prediction value but also a distribution, effectively giving confidence intervals for the outcome. Moreover, these functions can adapt as more data is collected for training, making it particularly applicable for small data sets. Given the scarce data available currently in mechanochemistry, it is a clear candidate until more data can be collected for deep models. For background on GPR in chemistry, including kernel design and uncertainty quantification in small-data settings, see the general overview in Chemical Reviews [35] and recent catalysis-focused reviews [36]. In our formulation, we assume a zero-mean function. This common choice is made when no strong prior knowledge about the mean exists; any systematic trends are then captured by the covariance (kernel) function, allowing the model to focus on the underlying correlation structure. Additionally, we use a composite kernel (Equation 5.5) consisting of a constant scaling factor (C), a Matérn kernel (with smoothness parameter ν fixed at 1.5) since it is effective in modeling physical processes [37, 38], and a white noise kernel (σ_n^2) to account for observation noise.

$$k = Ck_{\text{Matern}}(\nu = 1.5) + k_{\text{WhiteKernel}}(\sigma_n^2), \tag{5.5}$$

$$k_{Matern}(r) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}\, r}{l} \right)^{\nu} K_\nu \left(\frac{\sqrt{2\nu}\, r}{l} \right), \tag{5.6}$$

where $r = |x - x'|$ is the distance between two inputs, l is the length scale, σ^2 is the signal variance, ν controls the smoothness, $\Gamma(\cdot)$ is the Gamma function, and $K_\nu(\cdot)$ is the

modified Bessel function of the second kind.
Given a training set $\mathbf{X} = \{x_1, \dots, x_N\}$ with outputs $\mathbf{y} = \{y_1, \dots, y_N\}$ and a test set $\mathbf{X}^* = \{x_1^*, \dots, x_m^*\}$, the joint distribution of the training outputs and the latent function values $\mathbf{f}^*$ at the test points is modeled as:

$$\begin{pmatrix} \mathbf{y} \\ \mathbf{f}^* \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} K(\mathbf{X},\mathbf{X}) + \sigma_n^2\mathbf{I} & K(\mathbf{X},\mathbf{X}^*) \\ K(\mathbf{X}^*,\mathbf{X}) & K(\mathbf{X}^*,\mathbf{X}^*) \end{pmatrix}\right), \tag{5.7}$$

where σ_n^2 denotes the noise variance. K denotes the covariance function computed from the composite kernel and is used to construct the covariance matrices for both the training data and the test data.
Conditioning on the training data, the predictive distribution for the latent function values at the test points is Gaussian with mean and covariance given by

$$\bar{\mathbf{f}}^* = K(\mathbf{X}^*,\mathbf{X})\left[K(\mathbf{X},\mathbf{X}) + \sigma_n^2\mathbf{I}\right]^{-1}\mathbf{y}, \tag{5.8}$$

$$\operatorname{cov}(\mathbf{f}^*) = K(\mathbf{X}^*,\mathbf{X}^*) - K(\mathbf{X}^*,\mathbf{X})\left[K(\mathbf{X},\mathbf{X}) + \sigma_n^2\mathbf{I}\right]^{-1} K(\mathbf{X},\mathbf{X}^*). \tag{5.9}$$

The kernel hyperparameters (l, σ_n^2, ν) and the noise variance σ_n^2 are estimated by maximizing the log marginal likelihood:

$$\log p(\mathbf{y}|\mathbf{X}) = -\frac{1}{2}\mathbf{y}\left[K(\mathbf{X},\mathbf{X}) + \sigma_n^2\mathbf{I}\right]^{-1}\mathbf{y} - \frac{1}{2}\log\left|K(\mathbf{X},\mathbf{X}) + \sigma_n^2\mathbf{I}\right| - \frac{N}{2}\log(2\pi). \tag{5.10}$$

This structure allows flexibility in modeling the signal variance, smoothness, feature relevance, and noise level. The key hyperparameters optimized via Hyperopt are detailed in Table 5.2.

Table 5.2: Hyper-parameter search space used for Gaussian-process regression (GPR). $\mathcal{U}(a, b)$ denotes a continuous uniform prior on the interval (a, b).

Hyper-parameter	Search space / value
Constant scaling C	$\mathcal{U}(0.1, 10^3)$
Base length-scale l	$\mathcal{U}(0.05, 10)$
Noise variance σ_n^2	$\mathcal{U}(10^{-5}, 1.5)$
Matérn smoothness ν	fixed at 1.5

Random Forest

A random forest (RF) is an ensemble algorithm that makes use of multiple decision trees to enhance performance and reduce over-fitting. Each tree is fed with different samples of the training data (i.e., a bootstrap), and at each node, a random subset of features is used for decision making [39]. This introduces variability across trees and thus, errors made across different trees are averaged out in the final prediction. Given the confounded nature of variables in mechanochemical processes, tree-based algorithms are appealing due to the 'if-then' working principle, which can capture non-linear relationships. For instance, tree-based ensembles such as random forest are

standard in chemoinformatics [40], and QSAR [41]. The overall prediction is given by Equation 5.11 [42].

$$\hat{f}(x) = \frac{1}{M} \sum_{m=1}^{M} h(x; \Theta_m), \tag{5.11}$$

where $\hat{f}(x)$ is the ensemble prediction for x, M is the number of trees in the forest, and $h(x; \Theta_m)$ denotes the prediction of the m-th tree. Here, Θ_m represents the random factors, such as the bootstrap sample and random feature selection, used in constructing the m-th tree.
Key hyperparameters were tuned to optimize performance as detailed in Table 5.3.

Table 5.3: Hyper-parameter search space used for random-forest (RF) regression. $\mathcal{U}(a, b)$ denotes a discrete uniform prior on the integer range $\{a, \ldots, b\}$.

Hyper-parameter	Search space / value
Number of trees $n_{\text{estimators}}$	$\mathcal{U}(50, 300)$
Maximum tree depth	$\mathcal{U}(5, 30)$
Minimum samples split	$\mathcal{U}(2, 20)$
Minimum samples leaf	$\mathcal{U}(1, 10)$

Support Vector Regression

A support Vector Regression (SVR) aims to find a function $f(x)$ that deviates from the target values y_i by a value no greater than ϵ for all training points, while remaining as flat as possible [43]. The resulting regression function takes the form:

$$\hat{f}(x) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) K(x_i, x) + b, \tag{5.12}$$

where x_i are the training points (support vectors), α_i, α_i^* are non-negative Lagrange multipliers determined during optimization, and b is a bias term. The choice of the kernel function $K(x_i, x)$ allows capturing non-linear relationships. The optimization process finds these multipliers subject to constraints, including the crucial box constraint $0 \leq \alpha_i, \alpha_i^* \leq C$, where C is the regularization hyperparameter. This parameter $C > 0$ controls the trade-off between the flatness of $f(x)$ and the tolerance for errors larger than ϵ; a larger C allows less error but potentially a more complex function.
In SVR, the support vectors are defined by training data points that lie on or outside the boundary of the ϵ-insensitive tube. This characteristic of SVR is particularly advantageous in the current state of mechanochemical processes, where experimental data is scarce, and studies typically explore the effect of only one or two parameters on yield. By concentrating on the most informative data points, it can uncover subtle nonlinear dependencies among multiple process parameters. SVR has long been part of the chemometrics toolkit for nonlinear calibration and classification [44, 45]. The

Radial Basis Function (RBF) kernel was employed:

$$K(x_i, x) = \exp(-\gamma \|x_i - x\|^2) \tag{5.13}$$

where the kernel parameter γ controls the influence of a single training example. The key hyperparameters C, ϵ, and γ, which influence the model's complexity, error tolerance, and kernel shape, respectively, were optimized using Hyperopt as detailed in Table 5.4.

Table 5.4: Hyper-parameter search space used for support-vector regression (SVR). $\mathcal{U}(a, b)$ denotes a continuous uniform prior on (a, b), while $\log \mathcal{U}(a, b)$ is uniform in $\log_{10}$ space.

Hyper-parameter	Search space / value
Regularization C	$\log \mathcal{U}(0.1, 100)$
Epsilon ε	$\mathcal{U}(0.001, 1)$
Kernel scale γ	$\log \mathcal{U}(0.001, 1)$

XGBoost

XGBoost is another tree-based algorithm that, in contrast to RF, which constructs independent trees and averages them, builds an ensemble of regression trees in a sequential form [46]. Its efficiency, ability to capture complex non-linear relationships and feature interactions, and sophisticated regularization techniques make it a powerful choice for predictive modeling tasks, particularly with structured or tabular data often encountered in chemical process optimization. The final prediction $\hat{f}(x)$ is the sum of the predictions from all M trees:

$$\hat{f}(x) = \sum_{m=1}^{M} f_m(x), \tag{5.14}$$

where $f_m(x)$ represents the prediction of the m-th tree, and M corresponds to the number of estimators.
The training process iteratively adds trees by minimizing an objective function $\mathcal{L}(\phi)$ that combines a loss term (measuring the difference between predictions and actual values) and a regularization term Ω (penalizing model complexity), summed over all trees:

$$\mathcal{L}(\phi) = \sum_{i=1}^{N} l\left(y_i, \hat{y}_i^{(M)}\right) + \sum_{m=1}^{M} \Omega(f_m), \tag{5.15}$$

where $l(y_i, \hat{y}_i^{(M)})$ is the loss for sample i after M trees (e.g., squared error for regression), and $\hat{y}_i^{(M)}$ is the cumulative prediction. The regularization term for a single tree f is defined as:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \|w\|^2. \tag{5.16}$$

Here, T is the number of leaves in the tree, and w is the vector of scores (weights) at the leaves. The hyperparameter γ represents the minimum loss reduction required to make a further partition on a leaf node, acting as a tree pruning mechanism. The term $\frac{1}{2}\lambda\|w\|^2$ is an L2 regularization on the leaf weights, where λ (typically fixed, e.g., $\lambda = 1$ by default in XGBoost, and not tuned in this study) helps to prevent overfitting by shrinking the leaf scores.
The structural complexity of each tree f_m is primarily controlled by its maximum depth. The boosting process, which dictates how the ensemble is built, is further refined by several key hyperparameters: the learning rate (often denoted as η) scales the contribution of each new tree, reducing the impact of individual trees and preventing overfitting. The subsample specifies the fraction of training instances randomly sampled to grow each tree, introducing stochasticity and improving generalization. Sample-by-tree defines the fraction of features randomly sampled when constructing each tree (or each split), which further diversifies the trees and helps manage feature collinearity. These parameters, along with the number of estimators and γ, were optimized via Hyperopt. This careful tuning of the gradient boosting process, combined with its inherent regularization strategies, allows XGBoost to achieve high accuracy while effectively mitigating overfitting [47]. For chemical best practices with XGBoost on tabular reaction/molecular data, we refer to the Journal of Cheminformatics guidelines [48] and recent domain reviews in catalysis science [36]. The main hyperparameters tuned are listed in Table 5.5.

Table 5.5: Hyper-parameter search space used for extreme-gradient boosting (XGBoost). $\mathcal{U}(a, b)$ is a continuous uniform prior on (a, b), $\log \mathcal{U}(a, b)$ is uniform in $\log_{10}$ space, and $\mathcal{U}_{\text{int}}(a, b)$ is a discrete uniform prior on the integers $\{a, \ldots, b\}$.

Hyper-parameter	Search space / value
Maximum depth $d_{\max}$	$\mathcal{U}_{\text{int}}(3, 10)$
Learning rate η	$\log \mathcal{U}(0.01, 0.3)$
Number of estimators M	$\mathcal{U}_{\text{int}}(100, 500)$
Row subsample (`subsample`)	$\mathcal{U}(0.5, 1)$
Column subsample (`colsample_bytree`)	$\mathcal{U}(0.5, 1)$
Minimum loss-reduction γ	$\mathcal{U}(0, 5)$

5.3. Results and Discussion

The primary objective of this study is to develop accurate predictive models that can link mechanical and chemical operational parameters to experimental mechanochemical yield. To assess this, we evaluate several machine learning algorithms using two distinct modeling strategies:

1. A primary modeling approach, where each algorithm was trained directly on the full set of engineered features (either scaled or unscaled, as appropriate for the specific model).

2. A two-step modeling approach, designed to address the potentially dominant influence of the 'time' feature. This involved first modeling the yield based on 'time' alone, and then modeling the residuals (the difference between actual yield and the time model's prediction) using the remaining engineered features. The final prediction was the sum of the outputs from these two component models.

This dual approach allows for a comprehensive assessment of how well different algorithms capture the underlying relationships in the data, particularly concerning the prominent role of reaction time. It should be reiterated that the weighted MSE calculations discussed here reflect a configuration in which the yields above 70% were given a weight of 2.0, and all other yields a weight of 1.0.

5.3.1. Model Performance Evaluation

The performance of all trained models was evaluated on a held-out test set. To provide a multifaceted view of predictive accuracy, several standard regression metrics were employed, in addition to the Weighted MSE already defined in Eq. (5.3). These metrics are:

- **Root Mean Squared Error (RMSE):** This metric calculates the square root of the average of the squared differences between predicted and actual values. It is sensitive to large errors due to the squaring term. The RMSE is given by:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_{\text{true},i} - y_{\text{pred},i})^2} \tag{5.17}$$

  where N is the total number of samples in the test set, $y_{\text{true},i}$ is the actual yield for sample i, and $y_{\text{pred},i}$ is the predicted yield for sample i. Lower RMSE values indicate a better fit, and the metric shares the same units as the target variable (yield).

- **Mean Absolute Error (MAE):** MAE measures the average magnitude of errors in a set of predictions. It is the average over the test sample of the absolute differences between prediction and actual observation.

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|y_{\text{true},i} - y_{\text{pred},i}\right| \tag{5.18}$$

  MAE is less sensitive to outliers compared to RMSE and provides a straightforward interpretation of the average error magnitude, also in the units of the target variable.

- **Mean Absolute Percentage Error (MAPE):** MAPE expresses the average absolute difference between predicted and actual values as a percentage of actual values. This makes it a scale-independent metric, useful for comparing performance across datasets or models with different output scales.

$$\text{MAPE} = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{y_{\text{true},i} - y_{\text{pred},i}}{y_{\text{true},i}}\right| \tag{5.19}$$

where $y_{\text{true},i} \neq 0$. Lower MAPE values are desirable.

- **Coefficient of Determination (R^2):** The R^2 score indicates the proportion of the variance in the dependent variable (yield) that is predictable from the independent variables (features).

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_{\text{true},i} - y_{\text{pred},i})^2}{\sum_{i=1}^{N}(y_{\text{true},i} - \bar{y}_{\text{true}})^2} \tag{5.20}$$

where $\bar{y}_{\text{true}}$ is the mean of the true yield values in the test set. R^2 values range from $-\infty$ to 1, where 1 indicates a perfect fit, 0 indicates the model performs no better than predicting the mean of the target, and negative values indicate poorer performance than predicting the mean.

5.3.2. Comparison of modeling strategies and algorithm performance

The performance metrics for all evaluated models on the test set are summarized in Table 5.6, and the predictions can be visualized in Figure 5.2.
The Linear Regression model, utilizing a selected subset of scaled features, registered a weighted MSE of 395.29 and an R^2 of 0.53. While simple and interpretable, its linear nature inherently limits its ability to capture the complex, non-linear dynamics typical of chemical reactions, including those in mechanochemistry. Nonetheless, it should be noted that when examining the primary (single-stage) versions of the more complex algorithms, most struggled to significantly outperform this baseline. For instance, the GPR (Primary) model achieved a weighted MSE of 260.82 and an R^2 of 0.51, while RF (Primary) yielded a weighted MSE of 354.31 and R^2 of 0.52, and SVR (Primary) resulted in a weighted MSE of 346.48 and R^2 of 0.52. On average, these primary models offered only a modest reduction in weighted MSE (approximately 20-35% improvement over baseline) and showed R^2 values very close to, or even slightly below, that of the

Table 5.6: Predictive performance of the evaluated regression models on the test set. For RMSE, MAE, MAPE, and weighted MSE (W-MSE), lower values indicate better performance; for R^2 higher is better.

Model	RMSE	MAE	MAPE (%)	R^2	W-MSE
Linear regression	15.50	14.11	37.49	0.53	395.29
GPR (primary)	15.88	11.09	31.39	0.51	260.82
GPR (two-step)	9.43	7.48	26.59	0.83	93.37
RF (primary)	15.65	15.01	49.86	0.52	354.31
RF (two-step)	12.52	10.29	27.42	0.69	177.17
SVR (primary)	15.66	14.66	48.57	0.52	346.48
SVR (two-step)	15.36	14.73	50.20	0.54	320.32
XGBoost (primary)	12.65	11.20	32.68	0.69	194.37
XGBoost (two-step)	11.06	8.79	24.79	0.76	139.56

simpler linear model. The XGBoost (Primary) model was an exception, showing a marked improvement with a weighted MSE of 194.37 and an R^2 of 0.69; thus, when no prior domain knowledge is available, it should be the first-line choice, providing both competitive accuracy and an initial, data-driven ranking of influential variables. This general difficulty of the primary models to substantially advance beyond the linear regression baseline underscores the dominant influence of the 'time' variable, which, when not explicitly addressed, appears to overshadow the contributions of other features in these conventional modeling approaches.

Thus, a clear and consistent finding from these results is a significant benefit of the two-step modeling approach for several algorithms. The GPR two-step model stands out, achieving the lowest weighted MSE (93.37), MAE (7.48), and RMSE (9.43), alongside the highest R^2 value (0.83) among all models tested. Beyond its strong predictive accuracy, GPR can provide uncertainty estimates (confidence intervals) for its

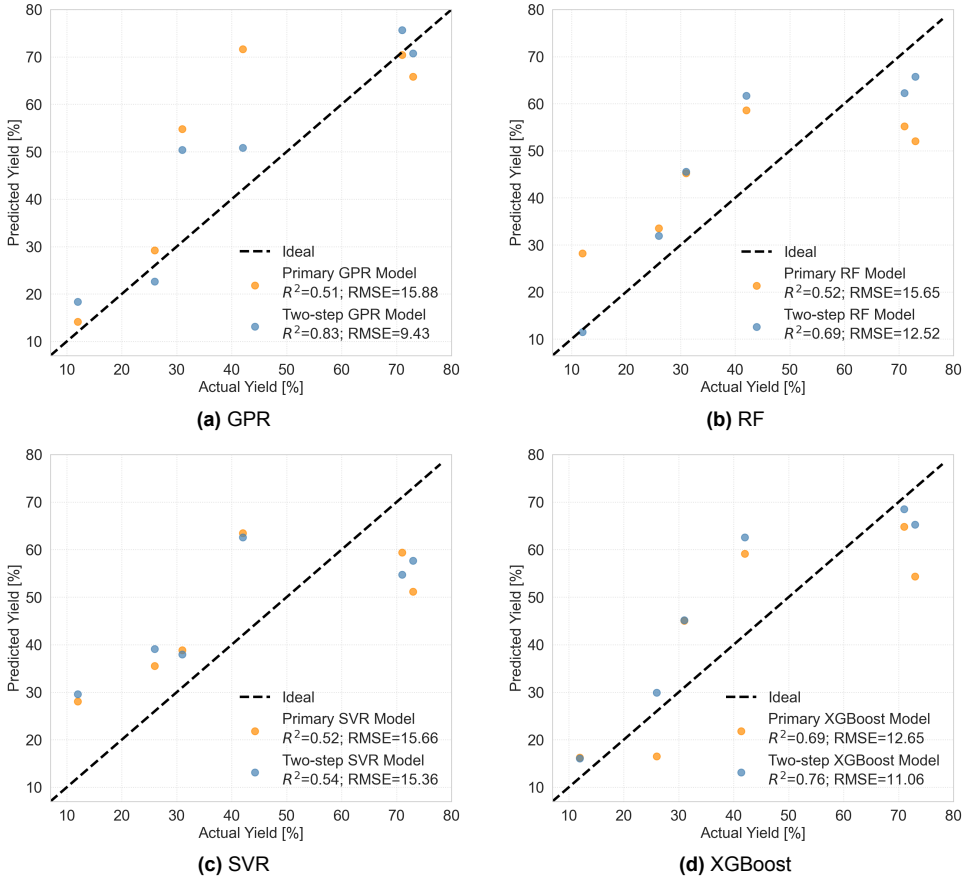


Figure 5.2: Predictions on the test set from primary and two-step variants of each regressor: (a) GPR, (b) RF, (c) SVR and (d) XGBoost.

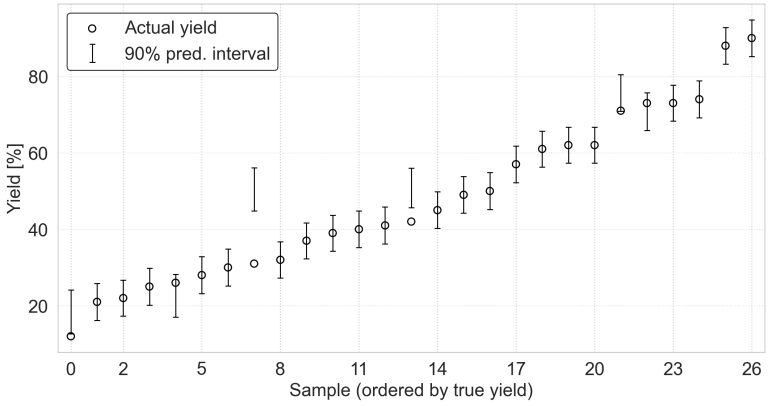
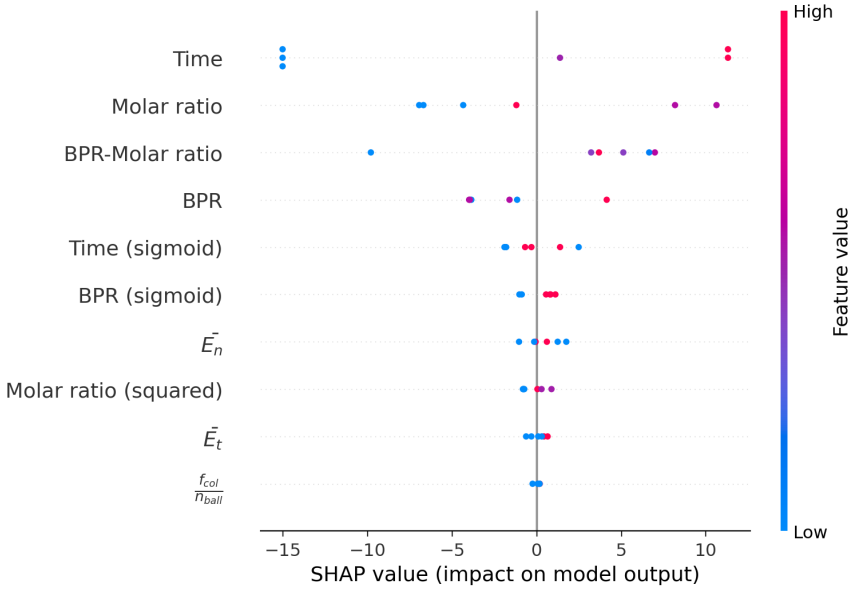


Figure 5.3: Actual yield versus predicted confidence intervals for the entire dataset.

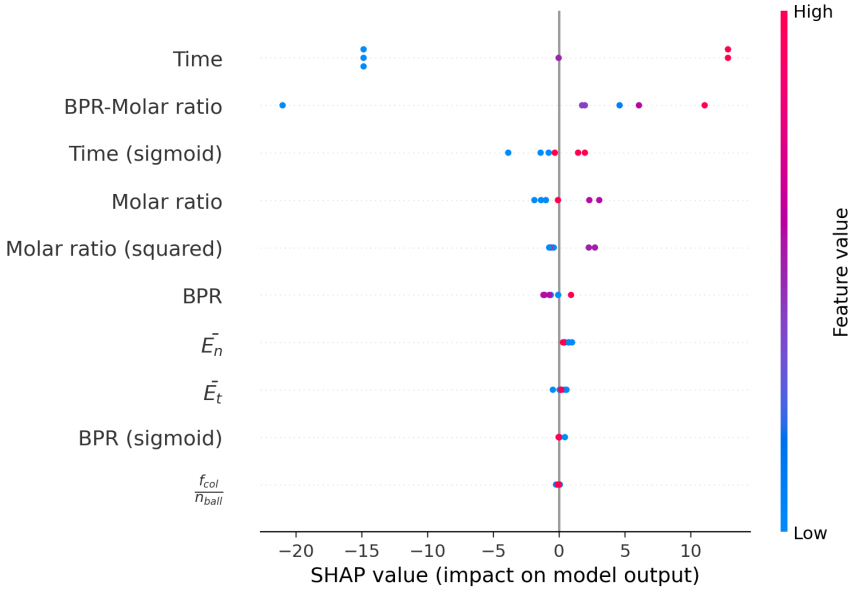
predictions (see Figure 5.3). Here, the effect of the weighted objective is evident; cases above 70% yield exhibit closer agreement with the model, indicating that errors at high yield were effectively down-weighted during training. This error scale is invaluable for guiding future experiments and assessing prediction reliability, especially when dealing with limited or costly experimental data, a common scenario in developing fields like mechanochemistry. The adaptability of its kernel functions also allows for encoding prior knowledge about the process, if available. The superior performance of the two-step GPR suggests that by first isolating the primary time trend, the GPR framework could more effectively model the subtle, potentially non-linear interactions among the remaining process parameters through its covariance structure.

To further delve into the interpretability of the more complex non-linear models, particularly the tree-based ensembles, SHAP (SHapley Additive exPlanations) analysis was employed for the two-step variants of RF and XGBoost. SHAP values provide a unified measure of feature importance by attributing to each feature the change in the expected model prediction when conditioning on that feature. A SHAP summary plot visualizes these attributions: each point represents a SHAP value for a feature and an instance, where the position on the x-axis indicates the impact on the model output (positive or negative), and the color represents the feature's value (high or low). Features are ranked by the sum of absolute SHAP values across all samples.

The XGBoost two-step model also demonstrated considerable improvements with the second strategy, emerging as the second-best performing model with a Weighted MSE of 139.56 and an R^2 of 0.76. Tree-based ensemble methods like XGBoost are inherently capable of capturing complex nonlinear relationships and variable interactions. The SHAP summary plot for this model (Figure 5.4 (a)) reveals that 'time' remains, as expected, the most influential feature for the overall two-step prediction, with higher time values generally pushing the prediction higher (positive SHAP values). Following 'time', features such as 'molar ratio' and the interaction 'BPR-molar ratio' exhibit significant



(a) XGBoost (two-step)



(b) Random forest (two-step)

Figure 5.4: SHAP summary plots for (a) the two-step XGBoost model and (b) the two-step random-forest model. See main text for interpretation.

importance, where higher values of these ratios tend to positively influence the predicted yield. These results align with our previous findings [3]. 'BPR' and 'time (sigmoid)' also show discernible impacts. It should be highlighted that, while the use of engineered features can be beneficial for capturing complex relationships and potentially improving model accuracy, it underscores a common trade-off in machine learning: a balance must often be struck between the enhanced predictive power gained from feature engineering and the goal of maintaining straightforward interpretability in terms of original process parameters. Despite this consideration, the two-step approach has effectively enabled the two-step XGBoost model to focus its learning capabilities on variance not explained by time, leading to more accurate predictions.

The Random Forest (RF) two-step model also benefited from the residual strategy, with a Weighted MSE of 177.17 and an R^2 of 0.69. Similar to XGBoost, RF models can effectively map non-linearities and interactions. The SHAP summary plot for the RF two-step model (Figure 5.4 (b)) shows a similar pattern of feature importance. 'Time' is again paramount, with a wide spread of SHAP values. 'BPR-molar ratio' and 'molar ratio' are the next most impactful features, with higher values generally increasing the predicted yield. Other engineered features like 'molar ratio (squared)' and 'time (sigmoid)' also contribute, though to a lesser extent than the top three. For both tree-based models, the features related to energy input ('$\bar{E}_n$', '$\bar{E}_t$') and collision frequency ('$\frac{f_{col}}{n_{ball}}$') show relatively lower overall SHAP values, suggesting a smaller impact on the output of these two-step models compared to the primary chemical and time-related parameters. However, it is crucial to reiterate that these mechanical milling properties were obtained under a constant fill ratio, and their influence is subject to change as this parameter is varied. We intend to investigate this in a future study, and it highlights the need for more experimental data that explores a wide range of operational parameters. Furthermore, this aligns with our previous finding that rotational speed, which has been abstracted into these variables, has a relatively lower relevance compared to the rest of the operational variables when maintaining a constant fill ratio [3].

For Support Vector Regression (SVR), the two-step approach provided a minimal improvement in Weighted MSE (320.32 for two-step vs 346.48 for primary) and R^2 (0.540 vs 0.522 for primary). SVR models, particularly with non-linear kernels like RBF, can be effective in high-dimensional spaces and are less sensitive to the dimensionality of the feature space. Their reliance on support vectors (a subset of training data) can make them memory efficient. However, in this case, the gains from the two-step strategy were less pronounced compared to GPR and the tree-based ensembles, suggesting that the primary SVR model might have already captured much of the structure the two-step approach aimed to resolve, or that its specific way of defining the decision boundary was less amenable to this sequential decomposition. As such, the SVR should not be investigated further until more data can be collected.

In summary, the GPR two-step model is the top-performing model across most key metrics based on the current dataset and evaluation criteria. The two-step modeling approach proved to be highly advantageous, particularly for GPR, XGBoost,

and RF, significantly enhancing their predictive accuracy, especially when considering the weighted error. These results emphasize the importance of considering tailored modeling strategies. The distinct characteristics of each algorithm (i.e., GPR's probabilistic outputs, the non-linear mapping capabilities and interpretability via SHAP of tree-based ensembles, and SVR's margin-based optimization) offer different strengths beyond raw predictive power. Therefore, the ultimate choice of predictive model in mechanochemical studies, or indeed any application, should not be solely dictated by a narrow focus on performance metrics. For instance, a model that performs slightly worse on a specific metric might be preferred if its intrinsic properties, such as superior interpretability, the ability to quantify uncertainty (as with GPR), or robustness to certain data characteristics, align more closely with the specific goals of the investigation or the practical constraints of its application. Factors such as data availability, the cost of acquiring more data, the need for uncertainty quantification for decision-making, and the desired level of insight into the underlying process mechanisms must be weighed alongside predictive accuracy.

5.3.3. Model generalization under mechanical regime change

To probe prospective generalization beyond the training distribution, we evaluated the two best-performing models on two new milling conditions (Table 5.7). In the original dataset (Table 5.1), the distribution of mechanical stressing conditions was effectively held constant, with a dissipation ratio of $\bar{E}_t/\bar{E}_n \approx 2.27$. Therefore, we expose the model to an unseen mechanical regime where the dissipation ratio is tuned to increase the dominance of tangential dissipation $\bar{E}_t/\bar{E}_n = 3.26$. This is practically achieved by reducing the fill ratio in the Emax ball mill.

We scaled the features with the training scaler and obtained predictions from the two-step GPR and XGBoost without refitting. Table 5.8 reports point predictions and absolute errors relative to the measured yields.
In case 27, the fill ratio (6%) departs from the training domain. The two-step models under-predict by ≈ 11 percentage points, which is broadly consistent with their held-out

Table 5.7: Out-of-sample milling conditions used for the generalization check. $\bar{E}_n$ and $\bar{E}_t$ are the mean normal and tangential energy dissipated per collision from DEM; $f_{\text{col}}/n_{\text{ball}}$ is the specific collision-frequency per ball. For these two cases, the fill ratio was 6%, and the dissipation ratio increased from $\bar{E}_t/\bar{E}_n = 2.27$ in Table 5.1 to 3.26 here, i.e., a more tangential-dominated stressing regime. Rotational speeds were 600 rpm (Case 27) and 788 rpm (Case 28).

Case	BPR	Mol. ratio	Time (h)	$\bar{E}_n$ (μJ)	$\bar{E}_t$ (μJ)	$f_{\text{col}}/n_{\text{ball}}$ (s^{-1})	Yield (%)
27	30	10	12.5	73	238	700	84
28	30	10	12.5	126	411	920	94

Table 5.8: Predictions on the two out-of-sample cases. Absolute errors are in percentage points of yield.

Case	$Yield_{true}$ (%)	$GPR_{two\text{-}step}$ (%)	$XGB_{two\text{-}step}$ (%)	$\lvert GPR_{two\text{-}step} - true \rvert$ (% points)	$\lvert XGB_{two\text{-}step} - true \rvert$ (% points)
27	84.0	73.14	72.42	10.86	11.58
28	94.0	74.09	72.42	19.91	21.58

RMSE (≈ 9.43) and indicates that a modest shift toward a more tangentially dominated stressing state can be tolerated when other operating factors remain consistent. However, in case 28, the fill ratio and rotational speed are shifted simultaneously, which further increases tangential stressing and the specific collision frequency. Errors rise to 20-22 percentage points, showing that the tangential mechanical regime alters the influence of the rotational speed on yield in a way that the models have not yet learned, leading to systematic under-prediction. As more variables move outside the training distribution, errors compound because of unseen nonlinear interactions between operational variables. This underscores the value of an expandable dataset design: targeted additions will expose these interactions and enable refitting for reliable use under regime changes. Because the DEM descriptors are mill-agnostic, different groups can explore the variables and ranges most relevant to them, and the pooled data will steadily improve accuracy and generalization.

5.4. Conclusions

In this chapter, we have demonstrated the effectiveness and applicability of various machine learning algorithms to predict conversion yield in the regeneration of $NaBH_4$ from a system of $NaBO_2 \cdot 4\,H_2O$ and MgH_2. More fundamentally, we established a practical methodology that couples mechanics-based descriptors with reaction stoichiometry into data-efficient predictors of yield. We have evaluated two distinct modeling strategies designed to account for the scarcity of data and the dominant influence of milling time in the process. Our findings indicate that carefully selected and configured ML models can provide valuable predictive capabilities, offering a pathway to optimize experimental efforts at a fraction of the time compared to the classic 'trial and error' approach, and gain deeper insights into the complex interplay of parameters in mechanochemistry.

The most compelling predictive performance was achieved by the Gaussian Process Regression (GPR) two-step model, which consistently outperformed all other evaluated algorithms across key metrics, including the lowest weighted MSE (93.37) and the highest R^2 (0.83). Following GPR, the two-step XGBoost and Random Forest models also delivered strong results. Beyond mere predictive accuracy, the choice of an appropriate ML model should also be guided by the specific objectives of the research and the intrinsic characteristics of the algorithms.

The practical implications of this work are significant. By developing reliable predictive models, researchers can substantially reduce the number of exploratory experiments, leading to considerable savings in time, materials, and energy. This is particularly pertinent given the current state of mechanochemistry, where experiments can be resource-intensive and exploratory. While the current data set provided a strong starting point, further exploration with more variability in mechanical conditions will enhance this linkage. Currently, performance degrades when multiple variables move outside the training domain, showcasing regime-dependent, nonlinear interactions. To facilitate this, the invariant mechanical characterization used in the dataset makes it readily expandable by independent experiments. Future work should focus on incorporating such expanded datasets, potentially exploring additional feature engineering techniques and advanced deep learning architectures once data volume permits.

While prediction models can be built from different inputs, the mechanical descriptors used here provide a common parameter that enables transferability across mills and scales. Generating those descriptors with full DEM simulations remains time-intensive, especially for larger simulations as processes are scaled up, and it is not yet clear whether the master curve logic generalizes to all ball-mill geometries. As a first approach to lowering this barrier, the next chapter develops a surrogate that approximates kinematics and dissipation at much lower computational cost, providing a practical path to faster iteration and adapting the framework across different mills.

References

[1] J. Batteas and T. Friščić. 'Shaken not stirred: procedures in mechanochemical syntheses and how to define them'. In: *RSC Mechanochemistry* 2 (2025), pp. 175–177. DOI: 10.1039/D5MR90005K. URL: https://pubs.rsc.org/en/content/articlehtml/2025/mr/d5mr90005k.

[2] V. Martinez, T. Stolar, B. Karadeniz, I. Brekalo and K. Užarević. 'Advancing mechanochemical synthesis by combining milling with different energy sources'. In: *Nature Reviews Chemistry* 7.1 (Jan. 2023), pp. 51–65. ISSN: 2397-3358. DOI: 10.1038/s41570-022-00442-1. URL: https://doi.org/10.1038/s41570-022-00442-1.

[3] S. Garrido Nuñez, D. L. Schott and J. T. Padding. 'Optimization of operational parameters in the mechanochemical regeneration of sodium borohydride (NaBH4)'. In: *International Journal of Hydrogen Energy* 97 (2025), pp. 640–648. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2024.11.360. URL: https://www.sciencedirect.com/science/article/pii/S0360319924050511.

[4] A. H. Hergesell, C. L. Seitzinger, J. Burg, R. J. Baarslag and I. Vollmer. 'Influence of ball milling parameters on the mechano-chemical conversion of polyolefins'. In: *RSC Mechanochem.* 2 (2 2025), pp. 263–272. DOI: 10.1039/D4MR00098F. URL: http://dx.doi.org/10.1039/D4MR00098F.

[5] F. Basoccu, P. Caboni and A. Porcheddu. 'Mechanochemistry: Unravelling the Impact of Metal Leaching in Organic Synthesis'. In: *ChemSusChem* n/a.n/a (), e202402547. DOI: https://doi.org/10.1002/cssc.202402547. eprint: https://chemistry-europe.onlinelibrary.wiley.com/doi/pdf/10.1002/cssc.202402547. URL: https://chemistry-europe.onlinelibrary.wiley.com/doi/abs/10.1002/cssc.202402547.

[6] A. Y. Ibrahim, R. T. Forbes and N. Blagden. 'Spontaneous crystal growth of co-crystals: the contribution of particle size reduction and convection mixing of the co-formers'. In: *CrystEngComm* 13 (4 2011), pp. 1141–1152. DOI: 10.1039/C004371K. URL: http://dx.doi.org/10.1039/C004371K.

[7] A. Stolle, R. Schmidt and K. Jacob. 'Scale-up of organic reactions in ball mills: process intensification with regard to energy efficiency and economy of scale'. In: *Faraday Discuss.* 170 (0 2014), pp. 267–286. DOI: 10.1039/C3FD00144J. URL: http://dx.doi.org/10.1039/C3FD00144J.

[8] S. Garrido Nuñez, D. L. Schott and J. T. Padding. 'Predictive models for energy dissipation in mechanochemical ball milling'. In: *Powder Technology* 457, 120919 (2025). ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2025.120919. URL: https://www.sciencedirect.com/science/article/pii/S0032591025003146.

[9] O. Jafter, S. Lee, J. Park, C. Cabanetos and D. Lungerich. 'Navigating Ball Mill Specifications for Theory-to-Practice Reproducibility in Mechanochemistry'. In: *Angewandte Chemie International Edition* 63 (Oct. 2024), e202409731. DOI: 10.1002/anie.202409731.

[10] J. F. Reynes, V. Isoni and F. García. 'Tinkering with Mechanochemical Tools for Scale Up'. In: *Angewandte Chemie International Edition* 62.44 (2023), e202300819. DOI: https://doi.org/10.1002/anie.202300819. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.202300819. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.202300819.

[11] M. Senna. 'The Optimization of Mechanochemical Processes toward Functional Nanocomposite Materials'. In: *Powders* 2.3 (2023), pp. 659–677. ISSN: 2674-0516. DOI: 10.3390/powders2030041. URL: https://www.mdpi.com/2674-0516/2/3/41.

[12] J. Batteas, K. G. Blank, E. Colacino, F. Emmerling, T. Friščić, J. Mack, J. Moore, M. E. Rivas and W. Tysoe. 'Moving mechanochemistry forward'. In: *RSC Mechanochem.* 2 (1 2025), pp. 10–19. DOI: 10.1039/D4MR90021A. URL: http://dx.doi.org/10.1039/D4MR90021A.

[13] E. Angloua, Y. Chang, W. Bradley, C. Sievers and F. Boukouvala. 'Modeling Mechanochemical Depolymerization of PET in Ball-Mill Reactors Using DEM Simulations'. In: *ACS Sustainable Chemistry & Engineering* 12.24 (2024), pp. 9003–9017. DOI: 10.1021/acssuschemeng.3c06081. URL: https://pubmed.ncbi.nlm.nih.gov/38903749/.

[14] J. Yu, K. Raju, S.-H. Jin, Y. Lee and H.-K. Lee. 'A machine learning approach for ball milling of alumina ceramics'. In: *The International Journal of Advanced Manufacturing Technology* 123.11-12 (2022), pp. 3451–3462. DOI: 10.1007/s00170-022-10999-2. URL: https://www.springerprofessional.de/en/a-machine-learning-approach-for-ball-milling-of-alumina-ceramics/23732142.

[15] Y. Li, J. Bao, T. Chen, A. Yu and R. Yang. 'Prediction of ball milling performance by a convolutional neural network model and transfer learning'. In: *Powder Technology* 403 (2022), p. 117409. ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2022.117409. URL: https://www.sciencedirect.com/science/article/pii/S0032591022003035.

[16] M. Boyer, D. Tabor, T.-H. Chao and D. Williams. *Center for the Mechanical Control of Chemistry Database: Interface*. Version v0.0.1.1. 2025. DOI: 10.5281/zenodo.14827611. URL: https://doi.org/10.5281/zenodo.14827611.

[17] D. Tan and F. García. 'Main group mechanochemistry: from curiosity to established protocols'. In: *Chemical Society Reviews* 48.9 (2019), pp. 2274–2292. DOI: 10.1039/C7CS00813A. URL: https://pubs.rsc.org/en/content/articlehtml/2019/cs/c7cs00813a.

[18] A. A. L. Michalchuk, I. A. Tumanov, S. Konar, S. A. J. Kimber, C. R. Pulham and E. V. Boldyreva. 'Challenges of Mechanochemistry: Is In Situ Real-Time Quantitative Phase Analysis Always Reliable? A Case Study of Organic Salt Formation'. In: *Advanced Science* 4.9 (2017), p. 1700132. DOI: https://doi.org/10.1002/advs.201700132. eprint: https://advanced.

onlinelibrary.wiley.com/doi/pdf/10.1002/advs.201700132. URL: https://advanced.onlinelibrary.wiley.com/doi/abs/10.1002/advs.201700132.

[19] T. Jarg, J. Tamm, E. Suut-Tuule, K.-M. Lootus, D. Kananovich and R. Aav. 'How reliable is internal standard method in monitoring mechanochemical synthesis? A case study of triphenylmethane in HPLC-UV-MS analysis of hemicucurbit[n]urils'. In: *RSC Mechanochem.* (2025), pp. -. DOI: 10.1039/D4MR00145A. URL: http://dx.doi.org/10.1039/D4MR00145A.

[20] C. Weidenthaler. 'In Situ Analytical Methods for the Characterization of Mechanochemical Reactions'. In: *Crystals* 12.3 (2022). ISSN: 2073-4352. DOI: 10.3390/cryst12030345. URL: https://www.mdpi.com/2073-4352/12/3/345.

[21] S. M. Mennen, C. Alhambra, C. L. Allen, M. Barberis, S. Berritt, T. A. Brandt, A. D. Campbell, J. Castañón, A. H. Cherney, M. Christensen, D. B. Damon, J. Eugenio de Diego, S. García-Cerrada, P. García-Losada, R. Haro, J. Janey, D. C. Leitch, L. Li, F. Liu, P. C. Lobben, D. W. C. MacMillan, J. Magano, E. McInturff, S. Monfette, R. J. Post, D. Schultz, B. J. Sitter, J. M. Stevens, I. I. Strambeanu, J. Twilton, K. Wang and M. A. Zajac. 'The Evolution of High-Throughput Experimentation in Pharmaceutical Development and Perspectives on the Future'. In: *Organic Process Research & Development* 23.6 (2019), pp. 1213–1242. DOI: 10.1021/acs.oprd.9b00140. URL: https://doi.org/10.1021/acs.oprd.9b00140.

[22] X. Caldentey and E. Romero. 'High-Throughput Experimentation as an Accessible Technology for Academic Organic Chemists in Europe and Beyond'. In: *Chemistry–Methods* 3.5 (2023), e202200059. DOI: https://doi.org/10.1002/cmtd.202200059. eprint: https://chemistry-europe.onlinelibrary.wiley.com/doi/pdf/10.1002/cmtd.202200059. URL: https://chemistry-europe.onlinelibrary.wiley.com/doi/abs/10.1002/cmtd.202200059.

[23] C. B. Santiago, J.-Y. Guo and M. S. Sigman. 'Predictive and mechanistic multivariate linear regression models for reaction development'. In: *Chem. Sci.* 9 (9 2018), pp. 2398–2412. DOI: 10.1039/C7SC04679K. URL: http://dx.doi.org/10.1039/C7SC04679K.

[24] L. M. Roch, F. Häse, C. Kreisbeck, T. Tamayo-Mendoza, L. P. E. Yunker, J. E. Hein and A. Aspuru-Guzik. 'ChemOS: An orchestration software to democratize autonomous discovery'. In: *PLOS ONE* 15.4 (Apr. 2020), pp. 1–18. DOI: 10.1371/journal.pone.0229862. URL: https://doi.org/10.1371/journal.pone.0229862.

[25] D. Reker, E. A. Hoyt, G. J. Bernardes and T. Rodrigues. 'Adaptive Optimization of Chemical Reactions with Minimal Experimental Information'. In: *Cell Reports Physical Science* 1.11 (2020), p. 100247. ISSN: 2666-3864. DOI: https://doi.org/10.1016/j.xcrp.2020.100247. URL: https://www.sciencedirect.com/science/article/pii/S2666386420302654.

[26] N. H. Angello, V. Rathore, W. Beker, A. Wołos, E. R. Jira, R. Roszak, T. C. Wu, C. M. Schroeder, A. Aspuru-Guzik, B. A. Grzybowski and M. D. Burke. 'Closed-loop optimization of general reaction conditions for heteroaryl Suzuki-Miyaura coupling'. In: *Science* 378.6618 (2022), pp. 399–405. DOI: 10.1126/science.adc8743. eprint: https://www.science.org/doi/pdf/10.1126/science.adc8743. URL: https://www.science.org/doi/abs/10.1126/science.adc8743.

[27] A. I. Leonov, A. J. S. Hammer, S. Lach, S. H. M. Mehr, D. Caramelli, D. Angelone, A. Khan, S. O'Sullivan, M. Craven, L. Wilbraham and L. Cronin. 'An integrated self-optimizing programmable chemical synthesis and reaction engine'. In: *Nature Communications* 15.1 (2024), p. 1240. DOI: 10.1038/s41467-024-45444-3. URL: https://doi.org/10.1038/s41467-024-45444-3.

[28] N. H. Angello, D. M. Friday, C. Hwang, S. Yi, A. H. Cheng, T. C. Torres-Flores, E. R. Jira, W. Wang, A. Aspuru-Guzik, M. D. Burke, C. M. Schroeder, Y. Diao and N. E. Jackson. 'Closed-loop transfer enables artificial intelligence to yield chemical knowledge'. In: *Nature* 633.8029 (2024), pp. 351–358. DOI: 10.1038/s41586-024-07892-1. URL: https://doi.org/10.1038/s41586-024-07892-1.

[29] M. Kessler and R. Rinaldi. 'Kinetic Energy Dose as a Unified Metric for Comparing Ball Mills in the Mechanocatalytic Depolymerization of Lignocellulose'. In: *Frontiers in Chemistry* 9 (2022), p. 816553. DOI: 10.3389/fchem.2021.816553. URL: https://doi.org/10.3389/fchem.2021.816553.

[30] scikit-learn developers. *Scikit-learn User Guide*. Release 1.7.1. scikit-learn. 2025. URL: https://scikit-learn.org/stable/user_guide.html (visited on 28/07/2025).

[31] R. Mantovani, A. Rossi, J. Vanschoren, B. Bischl and A. de Carvalho. 'Effectiveness of Random Search in SVM hyper-parameter tuning'. In: July 2015. DOI: 10.1109/IJCNN.2015.7280664.

[32] M. S. Sigman, K. C. Harper, E. N. Bess and A. Milo. 'The Development of Multidimensional Analysis Tools for Asymmetric Catalysis and Beyond'. In: *Accounts of Chemical Research* 49.6 (2016), pp. 1292–1301. DOI: 10.1021/acs.accounts.6b00194. URL: https://doi.org/10.1021/acs.accounts.6b00194.

[33] T. T. Cai and P. Hall. 'Prediction in functional linear regression'. In: *The Annals of Statistics* 34.5 (2006), pp. 2159–2179. DOI: 10.1214/009053606000000830. URL: https://doi.org/10.1214/009053606000000830.

[34] M. Hashemitaheri, S. M. R. Mekarthy and H. Cherukuri. 'Prediction of specific cutting forces and maximum tool temperatures in orthogonal machining by Support Vector and Gaussian Process Regression Methods'. In: *Procedia Manufacturing* 48 (2020). 48th SME North American Manufacturing Research Conference, NAMRC 48, pp. 1000–1008. ISSN: 2351-9789. DOI: https://doi.org/10.1016/j.promfg.2020.05.139. URL: https://www.sciencedirect.com/science/article/pii/S2351978920315912.

[35] V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti and G. Csányi. 'Gaussian Process Regression for Materials and Molecules'. In: *Chemical Reviews* 121.16 (2021), pp. 10073–10141. DOI: 10.1021/acs.chemrev.1c00022. URL: https://doi.org/10.1021/acs.chemrev.1c00022.

[36] B. M. Abraham, M. V. Jyothirmai, P. Sinha, F. Viñes, J. K. Singh and F. Illas. 'Catalysis in the digital age: Unlocking the power of data with machine learning'. In: *WIREs Computational Molecular Science* 14.5 (2024), e1730. DOI: https://doi.org/10.1002/wcms.1730. eprint: https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1730. URL: https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wcms.1730.

[37] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. Springer, 2006. ISBN: 978-1-4612-1494-6. DOI: 10.1007/978-1-4612-1494-6.

[38] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press, 2006. ISBN: 978-0-262-18253-9. DOI: 10.7551/mitpress/2320.

[39] L. Breiman. 'Random forests'. In: *Machine Learning* 45.1 (2001), pp. 5–32. DOI: 10.1023/A:1010933404324.

[40] Y.-C. Lo, S. E. Rensi, W. Torng and R. B. Altman. 'Machine learning in chemoinformatics and drug discovery'. In: *Drug Discovery Today* 23.8 (2018), pp. 1538–1546. DOI: 10.1016/j.drudis.2018.05.010. URL: https://doi.org/10.1016/j.drudis.2018.05.010.

[41] S. K. Niazi and Z. Mariam. 'Recent Advances in Machine-Learning-Based Chemoinformatics: A Comprehensive Review'. In: *International Journal of Molecular Sciences* 24.14 (2023). ISSN: 1422-0067. DOI: 10.3390/ijms241411488. URL: https://www.mdpi.com/1422-0067/24/14/11488.

[42] G. Biau and E. Scornet. 'A Random Forest Guided Tour'. In: *TEST* 25.2 (2016), pp. 197–227. DOI: 10.1007/s11749-016-0481-7.

[43] A. J. Smola and B. Schölkopf. 'A Tutorial on Support Vector Regression'. In: *Statistics and Computing* 14.3 (1st Aug. 2004), pp. 199–222. ISSN: 1573-1375. DOI: 10.1023/B:STCO.0000035301.49549.88. URL: https://doi.org/10.1023/B:STCO.0000035301.49549.88.

[44] H. Li, Y. Liang and Q. Xu. 'Support vector machines and its applications in chemistry'. In: *Chemometrics and Intelligent Laboratory Systems* 95.2 (2009), pp. 188–198. ISSN: 0169-7439. DOI: https://doi.org/10.1016/j.chemolab.2008.10.007. URL: https://www.sciencedirect.com/science/article/pii/S0169743908001998.

[45] R. G. Brereton and G. R. Lloyd. 'Support Vector Machines for classification and regression'. In: *Analyst* 135 (2 2010), pp. 230–267. DOI: 10.1039/B918972F. URL: http://dx.doi.org/10.1039/B918972F.

[46] T. Chen and C. Guestrin. ‘XGBoost: A Scalable Tree Boosting System’. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. 2016, pp. 785–794. DOI: 10.1145/2939672.2939785.

[47] J. H. Friedman. ‘Greedy Function Approximation: A Gradient Boosting Machine’. In: *Annals of Statistics* 29.5 (2001), pp. 1189–1232.

[48] D. Boldini, F. Grisoni, D. Kuhn, L. Friedrich and S. A. Sieber. ‘Practical guidelines for the use of gradient boosting for molecular property prediction’. In: *Journal of Cheminformatics* 15.1 (2023), p. 73. DOI: 10.1186/s13321-023-00743-7. URL: https://doi.org/10.1186/s13321-023-00743-7.

6

Accelerating granular dynamics simulations: a graph neural network surrogate

Thus far, we have leveraged mechanics-based descriptors because of their transferability across mills and scales, but generating them with full Discrete Element Method (DEM) simulations is computationally costly, especially for larger, scale-up simulations. Moreover, it remains uncertain whether the master curve logic generalizes across different types of mills. As a first approach, this chapter develops a physics-informed surrogate: a signed-distance-function graph neural network (SGN) trained on high-resolution data of the Retsch Emax. The model predicts ball kinematics and cumulative dissipation across the mill's contact network at a much looser timestep than DEM. This chapter represents a proof-of-concept built on a limited training set and a single baseline geometry; substantially more data across speeds, fill ratios, ball sizes, and geometries will be needed to calibrate magnitudes and generalize with confidence. With this chapter, we provide a practical approach that lowers compute cost, potentially eases adoption for non-specialists, and creates a testbed to probe transferability as datasets grow.

Parts of this chapter are adapted from: Garrido Nuñez, S., Schott, D. L., and Padding, J. T. 'Accelerating granular dynamics simulations: A graph neural network surrogate for complex high-energy ball milling'. In: Powder Technology 468, 121653 (2026).

6.1. Introduction

The rapid advancement of artificial intelligence can enable significant progress in scientific computing. In particular, surrogate models can approximate complex physical phenomena, such as particle interactions in granular systems or particle-based fluid representations, at much lower computational cost than traditional methods like the discrete element method (DEM) or smoothed particle hydrodynamics (SPH) [1–3].

Such surrogates are especially appealing in high-energy ball milling, where DEM is a well-established method for a wide range of applications, such as mechanochemistry [4, 5], mechanical alloying [6], ultra-fine milling, and particle breakage [7, 8]. Although each of these application fields faces its own set of distinct challenges for producing a valuable model, they share some critical commonalities, namely, they require a combination of a large number of particles (discrete elements), a very small timestep to accurately numerically integrate the underlying equations of motion, or possibly a combination of both [9, 10]. These conditions may be manageable when reproducing well-defined granular flows on a lab scale. However, as attention shifts towards industrial-scale applications, the computational cost needed to resolve numerous particle-particle and particle-wall contacts at small timesteps can become prohibitively expensive. This is especially problematic for emerging technologies such as mechanochemistry, where there are no clear connections to larger-scale machinery [11]. The intensive iterative design and implementation required slows the adoption of mechanochemistry, delaying its potential contributions to sustainability goals as defined by green chemistry principles [12]. Consequently, surrogate models are an appealing alternative in this context, offering a potential pathway to accelerate progress.

Among data-driven surrogates, graph neural networks (GNNs) have rapidly become a compelling paradigm for learned particle simulators, representing particles as nodes and interactions as edges, and rolling out dynamics via message passing [13]. The seminal Graph Network–based Simulator (GNS) demonstrated accurate, long-horizon roll-outs over fluids, rigid bodies, and simple granular settings, with strong generalization in particle count and initial conditions. [1]. Building on this, Choi & Kumar developed GNN surrogates specifically for granular flows (e.g., column collapse), reporting hundreds-fold speedups relative to high-fidelity solvers while preserving key flow features and scaling to larger domains than seen in training [14]. There is also emerging work coupling GNS with inverse design/optimization to tune DEM parameters or device settings efficiently [15].

A persistent challenge in learned granular simulators is boundary handling in complex geometry while preserving accurate physics. Early approaches either encoded distances to simple box-like boundaries or introduced virtual/ghost entities in so-called boundary GNNs to approximate walls. While these approaches have improved model generality, the introduction of virtual entities brings additional complexity and potential inaccuracies in wall-interaction physics [16–23]. To address these limitations, Li and Sakai proposed a signed distance function-based GNN (SGN) that encodes arbitrarily shaped boundaries as continuous distance fields, allowing the network to handle

complex geometries without virtual particles [2]. In parallel, physics-informed GNN architectures have been introduced to embed hard constraints from mechanics. For instance, Sharma and Fink enforce conservation of linear and angular momentum at individual collisions via a special message-passing scheme, yielding stable long-term predictions for 3D granular systems with inelastic impacts [24].

Despite these advances, no prior work has applied GNN-based surrogates to the extreme dynamic regime of high-energy ball milling. In such systems, particles collide violently under rapidly evolving boundary conditions (e.g., tumbling jars and moving reactor walls). Existing surrogate models have yet to demonstrate they can capture this highly dynamic, dissipative environment. In fact, prior data-driven studies have focused on macro-scale performance metrics of low-speed mills, such as predicting particle size distribution, bed height, or mixing quality, rather than simulating the detailed collision dynamics [16, 23, 25]. Related SDF-GNN work (e.g., Li & Sakai, 2024 [2]) has also targeted static or low-speed boundaries. Thus, this gap motivates the present work to develop a surrogate approach that can faithfully emulate the physics of a high-energy milling process.

To achieve this, we develop a surrogate model capable of handling complex dynamic boundaries, involving oscillatory and translational motion, characteristic of high-energy ball milling equipment used in powder processing. Specifically, we adapt an SGN to accurately capture complex energy input mechanisms and intense, high-energy dynamics and collisions, beyond the low-velocity regimes of prior studies. We extend the model by introducing moving boundaries and a secondary output for cumulative energy dissipation alongside local particle kinematics, enabling a dual prediction that is critical for mechanochemistry. Furthermore, our approach shows promising generalization to unseen motions and modified geometries, facilitating systematic study and design. Although our primary focus is on mechanochemical applications, this methodology can be adapted to virtually any high-energy ball milling scenario. Because the boundary kinematics and regimes differ compared to previous studies, we do not pursue a numerical head-to-head comparison; our contribution is complementary, extending SDF-based surrogates to time-varying, fast-moving wall regimes.

All associated code is freely available, including scripts for data extraction, transformation, and loading, as well as those used for model construction, training, and generative simulations, at the following location: `github.com/sgarridonunez/SGN_ball_milling`. In the subsequent sections, we detail the specific SGN architecture and training procedure using DEM simulation data, present validation results comparing surrogate predictions against ground truth values for particle dynamics and energy dissipation, and discuss the model's performance and generalization capabilities.

6.2. Methodology

6.2.1. Discrete Element Method (DEM) and simulation setup

The Discrete Element Method (DEM) is used to generate the data needed for training the surrogate model and serves as ground truth. In this study, Altair® EDEM™ 2021.2 was used as the DEM solver, and Python 3.9.12 was used for data post-processing. EDEM™ follows a soft-sphere approach by calculating the contact forces for each particle interaction using Hertz and Mindlin's contact model, which can capture the non-linear interactions that arise when particle–particle or particle–wall collisions occur [26, 27].

Newton's equations of motion are solved numerically to predict the evolution of the (angular) velocity of each particle:

$$m_i \frac{d\mathbf{V}_i}{dt} = \mathbf{F}_{c,i} + m_i \mathbf{g} \tag{6.1}$$

$$I_i \frac{d\boldsymbol{\omega}_i}{dt} = \boldsymbol{\tau}_i \tag{6.2}$$

where m_i, I_i, $\mathbf{V}_i$, and $\boldsymbol{\omega}_i$ are the mass, moment of inertia, velocity, and angular velocity, respectively, of particle i. $\mathbf{F}_{c,i}$ and $\boldsymbol{\tau}_i$ represent the total contact force and total contact torque (relative to the particle's center of mass), respectively. The total force and torque are determined by summing over all neighbors in contact with particle i. Each discrete element has its own radius R, mass m, Young's modulus Y, shear modulus G, coefficient of restitution e, and Poisson ratio v. The contact force $\mathbf{F}_{c,ij}$ on a particle i due to its interaction with another particle j (or wall) is the vector sum of a normal force $\mathbf{F}_{n,ij}$ and tangential force $\mathbf{F}_{t,ij}$:

$$\mathbf{F}_{c,ij} = \mathbf{F}_{n,ij} + \mathbf{F}_{t,ij} = \left(K_n \boldsymbol{\delta}_{n,ij} - \gamma_n \mathbf{V}_{n,ij}\right) + \left(K_t \boldsymbol{\delta}_{t,ij} - \gamma_t \mathbf{V}_{t,ij}\right) \tag{6.3}$$

with:

$$K_n = \frac{4}{3} Y^* \sqrt{R^* \delta_n} \tag{6.4}$$

$$\gamma_n = -2\sqrt{\frac{5}{6}}\, \beta \sqrt{S_n m^*} \geq 0 \tag{6.5}$$

$$K_t = 8G^* \sqrt{R^* \delta_n} \tag{6.6}$$

$$\gamma_t = -2\sqrt{\frac{5}{6}}\, \beta \sqrt{S_t m^*} \geq 0 \tag{6.7}$$

$$S_n = 2Y^* \sqrt{R^* \delta_n} \tag{6.8}$$

$$S_t = 8G^* \sqrt{R^* \delta_n} \tag{6.9}$$

$$\beta = \frac{\ln(e)}{\sqrt{\ln^2(e) + \pi^2}} \tag{6.10}$$

$$\frac{1}{Y^*} = \frac{1 - \nu_1^2}{Y_1} + \frac{1 - \nu_2^2}{Y_2} \tag{6.11}$$

$$\frac{1}{G^*} = \frac{2(2 - \nu_1)(1 + \nu_1)}{Y_1} + \frac{2(2 - \nu_2)(1 + \nu_2)}{Y_2} \tag{6.12}$$

$$\frac{1}{R^*} = \frac{1}{R_1} + \frac{1}{R_2} \tag{6.13}$$

$$\frac{1}{m^*} = \frac{1}{m_1} + \frac{1}{m_2} \tag{6.14}$$

Here, $\mathbf{V}_{n,ij}$ and $\mathbf{V}_{t,ij}$ denote the relative velocities in the normal and tangential directions between particles i and j at the point of contact. The vectors $\boldsymbol{\delta}_{n,ij}$ and $\boldsymbol{\delta}_{t,ij}$ represent the normal and tangential overlaps between the particles, with the tangential overlap obtained by integrating the relative tangential velocity over time and projecting it onto the current tangential direction. The constants K_n and K_t are the elastic coefficients for normal and tangential contacts, respectively, while γ_n and γ_t correspond to the viscoelastic damping coefficients for these contacts.

On the right-hand side of Equation 6.3, the expression within the first set of parentheses represents the normal force, and the expression in the second set corresponds to the tangential force. Specifically, the normal force comprises two components: a spring force and a normal damping force $\mathbf{F}_{n,d}$, while the tangential force is made up of a shear force and a tangential damping force $\mathbf{F}_{t,d}$. The magnitude of the tangential force F_t is limited according to the Coulomb friction law: if $F_t \geq \mu_f F_n$, where μ_f is the friction coefficient and F_n the magnitude of the normal force, then F_t is set equal to F_n (while still oriented in the tangential direction).

Additionally, the contact torque $\boldsymbol{\tau}_{ij}$ acting on particle i as a result of its interaction with particle (or wall element) j is determined by the cross product of the vector

$\mathbf{R}_{ij}$, which extends from the center of mass of particle i to the contact point with particle j, and the tangential contact force $\mathbf{F}_{t,ij}$. Given that the particles experience continuous rolling motion, particularly in interactions with a wall, it is essential to account for any slight deviations from perfect sphericity. This is achieved by introducing a rolling torque, $\boldsymbol{\tau}_{r,ij}$, which is computed using the coefficient of rolling friction μ_r, the magnitude of the normal contact force $F_{n,ij}$, the distance R_{ij} from the center of mass to the contact point, and the orientation of the particle's relative angular velocity, $\boldsymbol{\omega}_{rel}$. These relationships are described by:

$$\boldsymbol{\tau}_{ij} = \mathbf{R}_{ij} \times \mathbf{F}_{t,ij} + \boldsymbol{\tau}_{r,ij} \tag{6.15}$$

$$\boldsymbol{\tau}_{r,ij} = -\mu_r F_{n,ij} R_{ij} \frac{\boldsymbol{\omega}_{rel}}{\omega_{rel}} \tag{6.16}$$

Finally, the energy dissipated over the time interval from t_1 to t_2, due to the damping effects characterized by γ_n and γ_t, is calculated as follows:

$$E_n = \int_{t_1}^{t_2} \mathbf{F}_{n,d} \cdot \mathbf{V}_{n,ij} \, \mathrm{d}t \tag{6.17}$$

$$E_t = \int_{t_1}^{t_2} \mathbf{F}_{t,d} \cdot \mathbf{V}_{t,ij} \, \mathrm{d}t \tag{6.18}$$

The DEM model has been calibrated and validated for the mechanochemical regeneration of sodium borohydride $NaBH_4$ in the *Emax* high-energy ball mill produced by the German company Retsch. We refer to chapter 3 for more details [4, 28]. The crucial properties relevant to this work can be found in Table 3.2.

We model the commercially available high-energy ball mill *Emax* produced by the company Retsch. The machine can allocate proprietary grinding jars with 125 ml of

Table 6.1: Material and numerical parameters used for the milling balls and jar walls (steel X46Cr13) while accounting for $NaBO_2 \cdot 4\,H_2O$ and MgH_2.

Parameter	Value
Particle diameter	0.01 m
Coefficient of restitution	0.30
Coefficient of friction	0.30
Coefficient of rolling friction	0.045
Density	7700 kg m^{-3} [29]
Young's modulus	2.05 GPa [29]
Poisson's ratio	0.235 [29]
Simulation time step	9.5×10^{-7} s
Total simulation time	15 s
Time-integration method	Euler

volume that follow a circular motion with a rotational speed *n* up to 2000 revolutions per minute with an amplitude (radius) *A* of 1.7 centimeters, see Fig. 6.1. The movement of the jar has been replicated in our simulations. An STL file was built and imported into EDEM™ to represent the geometry of the milling jar accurately.

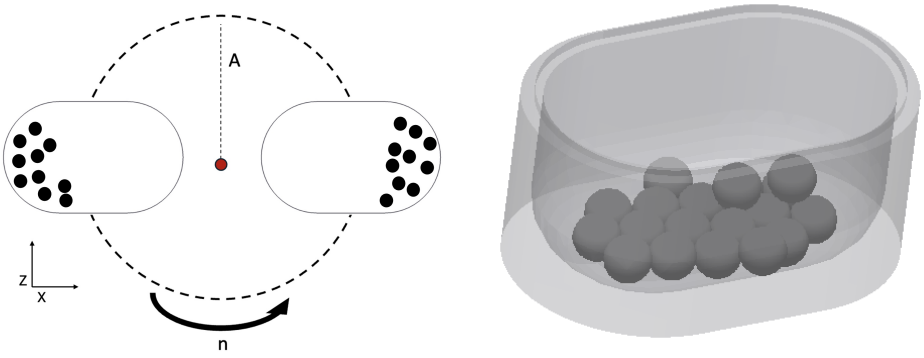
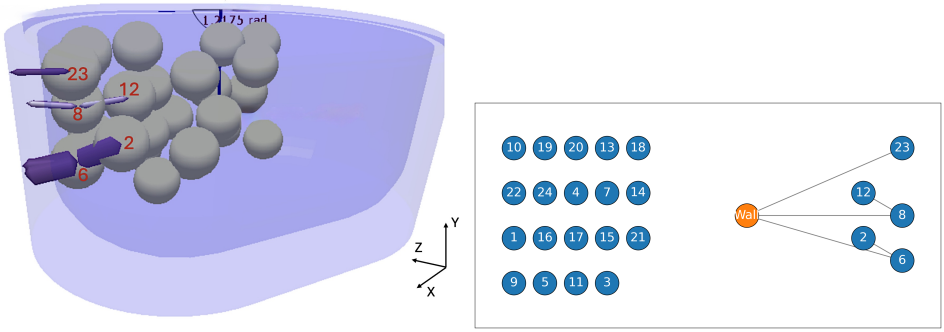


Figure 6.1: (a) Schematic of jar movement (A = 1.7cm, n = 300rpm) (b) 3D model of the milling jar [4].

The system is initially set up by generating all the discrete media over a five-second interval, which allows them to settle into their resting positions within the jar before any motion begins. After this initialization phase, the simulation runs for an additional 10 seconds with a rotational speed of 300 rpm to capture the dynamic behavior of the system. A fill ratio of 10% is used, corresponding to a total of 24 milling balls. To reduce computational complexity, the model is simplified by including only milling balls as discrete elements. This simplification is justified because the influence of the processed material can be effectively represented by calibrating the friction and restitution coefficients [5, 30, 31].

6.2.2. Graph Neural Network and Surrogate Model

A graph is a representation composed of a set of nodes and a set of edges that connect pairs of these nodes [32]. This structure models relationships or interactions between objects in various domains, such as computer networks, social networks, biological systems, and, in the context of this paper, granular systems (see Figure 6.2). Moreover, graphs provide a natural framework for message passing, where nodes exchange information with their neighbors along the edges [33]. This capability is fundamental in graph neural network architectures, enabling iterative aggregation of local information to capture complex, global patterns within the graph.



(a) (b)

Figure 6.2: (a) DEM connectivity example, (b) Graph representation of DEM timestep. Note that each particle has an individual ID used to define connectivity at any given timestep.

Architecture

In this study, node and edge features encode the dynamic state and geometric context of the system. Specifically, the dynamic state is encoded through particle velocity, and the geometric context is dictated by a Signed Distance Function (SDF), leading to the definition of an SDF-based graph neural network (SGN) as proposed by Li et al. [2]. As such, we follow the same terminology for consistency. The SDF is a fixed field computed from the watertight, triangulated jar STL (see Figure 6.3). It provides a per-particle distance channel for node features and a proximity vector on particle–wall edges. The mill's boundary motion is applied to the wall: for each time step the STL is rigidly translated in space by the time-dependent center of mass (CoM), and the SDF is evaluated relative to the moved wall. Particle states are not directly forced by this motion; they are updated by integrating predicted smooth accelerations. Conditioning on the moving boundary allows inference under unseen motion trajectories, provided the training data spans comparable kinematic/energy regimes.

For closed (watertight) surfaces, the inside/outside sign is unambiguous and internal cavities are supported; open holes or non-manifold seams can introduce sign ambiguity and should be repaired before SDF generation. For new jar shapes, we simply recompute the SDF field for the new STL; no retraining is required for inference, whereas new training datasets require regenerating the SDF-derived features.

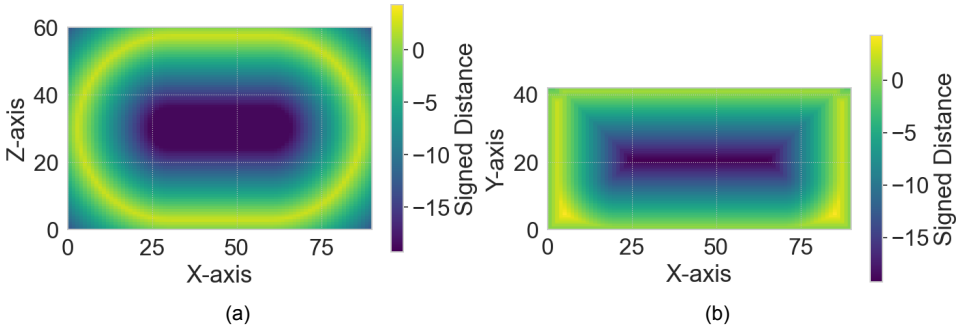


Figure 6.3: Geometry's SDF field slices (a) XZ plane, (b) XY plane. In our system, negative values indicate positions inside the jar. Thus, the most negative values represent positions furthest away from the wall.

The granular system at a given time is represented as a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The set of nodes $\mathcal{V}$ includes nodes v_p representing each individual particle ($p = 1, \ldots, N_p$) and a single, dedicated node v_w representing properties of the wall boundary. Note that the node v_w does not dictate any spatial context information for the particles; it is simply established to define variables relevant to the geometry, such as its center of mass, and to be able to keep track of collisions between particles and the wall, which are crucial in ball milling.

The set of edges $\mathcal{E}$ comprises two subsets: particle-particle edges $\mathcal{E}_{pp} = \{e_{ij}\}$ and particle-wall edges $\mathcal{E}_{pw} = \{e_{iw}\}$. For generating the graphs used during offline training, the edge sets $\mathcal{E}_{pp}$ and $\mathcal{E}_{pw}$ are constructed directly from the contact pairs reported by the source high-fidelity DEM simulation (i.e. from Altair® EDEM™) at each corresponding time step. This implicitly defines the interaction range learned by the model from the training data.

To capture temporal dependencies, features for particle and wall nodes incorporate information over a history window covering the current and τ preceding time steps (total window size of $\tau + 1$). For a particle node v_i at time step t_n, the input node feature vector $\boldsymbol{\epsilon}_{v_i}(t_n)$ is constructed by concatenating features related to particle kinematics and boundary interactions:

$$\boldsymbol{\epsilon}_{v_i}(t_n) = \text{concat}\Big(\{\mathbf{V}_p(v_i, t_j)\}_{j=n-\tau}^{n}, \{\phi_{SDF}(v_i, t_j), \nabla\phi_{SDF}(v_i, t_j)\}_{j=n-\tau}^{n}\Big) \tag{6.19}$$

where $\mathbf{V}_p(v_i, t_j)$ is the translational velocity of particle i at time t_j. Crucially, the SDF value $\phi_{SDF}(v_i, t_j)$ and its gradient $\nabla\phi_{SDF}(v_i, t_j)$ are calculated for all particle positions relative to the time-dependent boundary geometry (since the geometry is in oscillatory motion) at each time step t_j in the history window. This provides a continuous geometric and contact context to every particle node. The feature vector for the wall node, $\boldsymbol{\epsilon}_{v_w}(t_n)$, includes information about the wall's state over the time window, such as its Center of Mass (CoM) position and rotational velocity, padded with zeros to match the

dimensionality of particle node features for each snapshot in the window.

Edge features encode the relative spatial configuration or interaction properties between connected nodes identified by the DEM simulation during data generation. For particle-particle edges $e_{ij} \in \mathscr{E}_{pp}$, the features $\boldsymbol{\epsilon}_{e_{ij}}(t_n)$ include the relative distance vector:

$$\boldsymbol{\epsilon}_{e_{ij}}(t_n) = \{\mathbf{r}_{ij}(t_n)\} \quad \text{where} \quad \mathbf{r}_{ij}(t_n) = \mathbf{X}_j(t_n) - \mathbf{X}_i(t_n) \tag{6.20}$$

For particle-wall edges $e_{iw} \in \mathscr{E}_{pw}$, the features $\boldsymbol{\epsilon}_{e_{iw}}(t_n)$ represent the particle-wall interaction, using the SDF distance vector derived from the particle's SDF features:

$$\boldsymbol{\epsilon}_{e_{iw}}(t_n) = \phi_{SDF}(v_i, t_n) \frac{\nabla\phi_{SDF}(v_i, t_n)}{\|\nabla\phi_{SDF}(v_i, t_n)\|} \tag{6.21}$$

These specific input features (velocity history, SDF history, relative positions) are chosen because they can be readily updated or recalculated during the recursive simulation phase using only the model's outputs (acceleration integrated to velocity and position) and the known boundary motion, enabling a closed-loop prediction while sliding the history window (W). This contrasts with features like contact forces or tangential overlaps, which are outputs of the DEM simulation but cannot be directly calculated during the surrogate's recursive loop without making further assumptions or predictions of unknown future contacts. All input node (particle and wall) and edge (particle-particle and particle-wall) features are normalized using the mean and standard deviation derived from the training dataset prior to being processed by the network. Separate normalization statistics are maintained for particle nodes, wall nodes, particle-particle edges, and particle-wall edges. Additionally, Gaussian noise with a standard deviation of 0.005 ($\sigma = 0.005$) is added to the normalized velocity features during training to enhance robustness during recursive inference.

The SGN architecture follows the established Encoder–Processor–Decoder paradigm. First, the Encoder employs independent Multi-Layer Perceptrons (MLPs), MLP_v and MLP_e, with ReLU activations to map the input node and edge features to initial latent embeddings, $\mathbf{h}_v^0$ and $\mathbf{h}_e^0$, respectively:

$$\mathbf{h}_v^0 = \text{MLP}_v(\boldsymbol{\epsilon}_v) \quad \forall v \in \mathscr{V} \tag{6.22}$$

$$\mathbf{h}_e^0 = \text{MLP}_e(\boldsymbol{\epsilon}_e) \quad \forall e \in \mathscr{E} \tag{6.23}$$

Note that while a single MLP_v is shown, distinct initial layers or feature handling could be applied to particle vs. wall nodes if necessary. Similarly, MLP_e processes both edge types. Second, the Processor consists of L_p interaction layers, performing iterative message passing to refine node representations by propagating information through the graph. Within each layer l:

1. An edge-update MLP, ψ^e, computes messages based on the embeddings of connected nodes and the edge itself:

$$\mathbf{m}_{e_{ij}}^l = \psi^e\left(\mathbf{h}_{v_i}^l, \mathbf{h}_{v_j}^l, \mathbf{h}_{e_{ij}}^l\right) \tag{6.24}$$

2. An aggregation function, (element-wise mean in our implementation), pools incoming messages for each node v_i (including the wall node v_w) from its neighborhood $\mathcal{N}(i)$:

$$\bar{\mathbf{m}}_{v_i}^{l} = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \mathbf{m}_{e_{ij}}^{l} \tag{6.25}$$

3. A node-update MLP, ψ^{v}, updates the node embedding using its previous state and the aggregated message:

$$\mathbf{h}_{v_i}^{l+1} = \psi^{v}\left(\mathbf{h}_{v_i}^{l}, \bar{\mathbf{m}}_{v_i}^{l}\right) \tag{6.26}$$

Third, the Decoder utilizes an MLP, MLP_d, to map the final node embeddings from the processor, $\mathbf{h}_v^{L_p}$, to the target outputs. For particle nodes v_p, the primary target is the normalized particle acceleration $\mathbf{A}_{\text{norm}}(v_p, t_n)$, as this allows the model to drive the system's dynamics through integration. The output for the wall node v_w is also computed, but disregarded for the primary task.

$$\mathbf{A}_{\text{norm}}(v_p, t_n) = \text{MLP}_d\left(\mathbf{h}_{v_p}^{L_p}\right) \tag{6.27}$$

Additionally, a Global Readout branch aggregates the final node embeddings $\mathbf{h}_v^{L_p}$ (via mean pooling across all particle nodes with an edge, i.e. particles undergoing a collision) and passes the result through a separate MLP to predict a normalized global system property, specifically for this work, the incremental energy dissipation $\Delta E_{\text{norm}}(t_n)$ (result from adding Equation 6.17 and Equation 6.18). Because ΔE_{norm} is learned from DEM targets, calibration across operating regimes depends on training coverage of dissipation magnitudes and collision statistics (e.g., RPM, amplitude, fill ratio, materials). This global variable was selected due to its established relevance in characterizing the potential outcomes of mechanochemical processes [4]. Table 6.2 summarizes the graph inputs and targets used in this work.

Table 6.2: Summary of SGN graph inputs and targets. Per-snapshot node features are concatenated over a window $W = \tau + 1$.

Component	Value
Node Features	
Particle (per snapshot)	$[\mathbf{V},\ \phi_{\mathrm{SDF}},\ \nabla\phi_{\mathrm{SDF}}]$ (Dim. 7)
Wall (per snapshot)	[**CoM**, RPM] (zero-padded; Dim. 7)
Edge Features	
PP edge (particle–particle)	r (Dim. 3)
PW edge (particle–wall)	$\phi_{\mathrm{SDF}} \cdot \hat{\nabla}\phi_{\mathrm{SDF}}$ (Dim. 3)
Targets	
Node (per particle)	$\mathbf{A}_{norm}$ (Dim. 3)
Global (per graph)	ΔE_{norm} (Dim. 1)

Symbols: $\mathbf{V} = [v_x, v_y, v_z]$ (velocity); ϕ_{SDF} (signed distance to wall surface); $\nabla\phi_{\mathrm{SDF}} = [\partial_x\phi_{\mathrm{SDF}}, \partial_y\phi_{\mathrm{SDF}}, \partial_z\phi_{\mathrm{SDF}}]$ (SDF gradient); $\hat{\nabla}\phi_{\mathrm{SDF}} = \nabla\phi_{\mathrm{SDF}} / \|\nabla\phi_{\mathrm{SDF}}\|$ (unit normal); $\mathbf{CoM} = [\mathrm{CoM}_x, \mathrm{CoM}_y, \mathrm{CoM}_z]$ (jar center-of-mass position); $r = X_j - X_i$ (PP separation vector); $\mathbf{A}_{norm} = [a_x, a_y, a_z]$ (normalized acceleration); ΔE_{norm} (per-step normalized dissipated energy).

Training

The network's learnable parameters θ are trained offline using supervised learning on data generated by high-fidelity DEM simulations, which are sampled at a fine time resolution (Δt_{DEM}). Accurately defining Δt_{DEM} is essential for enabling the model to capture the dynamics effectively. In this work, because the median collision duration is approximately 0.0005 seconds, we selected a Δt_{DEM} of 0.0001 seconds to ensure that the model can accurately learn the evolution of collisions. Although this parameter can be fine-tuned depending on the application, in systems where individual collisions are critical for realistic granular flow, it is advisable not to exceed the median collision duration. It is also important to note that the training data utilized comprises snapshots taken only after the initial particle generation phase is complete and the motion of the milling jar has commenced, focusing the model on the relevant dynamic interactions. Then, the first 4.5 seconds are used for training, resulting in a total of 45,000 snapshots.

The model was trained on one Nvidia A100 [34] with a batch size of 2 and a learning rate initially set to 1e-4, managed by an Adam optimizer and an exponential scheduler targeting a final rate of 1e-6 with a maximum of 2000 training epochs. The objective is to minimize a suitable loss function between the SGN's predictions ($\mathbf{A}_{\mathrm{norm}}, \Delta E_{\mathrm{norm}}$) and the corresponding ground truth values derived from the DEM data. Specifically, the total loss function $\mathscr{L}_{total}$ is a weighted sum of the loss calculated for the primary task (node acceleration prediction, $\mathscr{L}_{node}$) and the loss for the auxiliary global

prediction task (energy dissipation, $\mathscr{L}_{global}$):

$$\mathscr{L}_{total} = \alpha \mathscr{L}_{node} + \mathscr{L}_{global} \tag{6.28}$$

with

$$\mathscr{L}_{\text{node}} = \frac{1}{N} \sum_{k=1}^{N} H_{\delta}\left(\mathbf{A}^{\text{pred}}_{\text{norm},k} - \mathbf{A}^{\text{gt}}_{\text{norm},k}\right) \tag{6.29a}$$

$$\mathscr{L}_{\text{global}} = \frac{1}{B} \sum_{b=1}^{B} H_{\delta}\left(\Delta E^{\text{pred}}_{\text{norm},b} - \Delta E^{\text{gt}}_{\text{norm},b}\right) \tag{6.29b}$$

where B is the number of graphs in the mini-batch, N is the total number of particle-node acceleration components in the batch (wall node excluded), $\mathbf{A}^{\text{pred}}_{\text{norm}}$ and $\mathbf{A}^{\text{gt}}_{\text{norm}}$ are the predicted and ground-truth normalized per-particle accelerations (the index k runs over all particle components), and $\Delta E^{\text{pred}}_{\text{norm}}$ and $\Delta E^{\text{gt}}_{\text{norm}}$ are the predicted and ground-truth normalized per-step dissipated-energy increments (the former produced by the global head from pooled node embeddings over particle nodes with at least one incident edge). We use the Huber penalty (Equation 6.30) with threshold $\delta = 2$:

$$H_{\delta}(e) = \begin{cases} \dfrac{1}{2\delta}\, e^2, & |e| < \delta, \\ |e| - \dfrac{\delta}{2}, & |e| \geq \delta. \end{cases} \tag{6.30}$$

The weighting factor α (set to 3.0 in our implementation) allows for prioritizing the accuracy of the particle dynamics prediction during training relative to the global energy prediction. In this work, both $\mathscr{L}_{node}$ and $\mathscr{L}_{global}$ utilize the Huber loss function. Due to the high-energy collisions inherent in ball milling, the ground truth acceleration distribution can exhibit large spikes relative to median values. This occurs not only because of the large contact forces during impacts but also because sampling the DEM simulation at Δt_{DEM} can alias high-frequency events. Collision dynamics occurring at the DEM's finer internal timestep might not be fully resolved between samples, leading to apparent discontinuities or spikes in the calculated acceleration used for training. Consequently, the Huber loss is employed for its robustness to such outliers, combining the benefits of L2 loss (mean squared error, MSE) for small errors and L1 loss (mean absolute error, MAE) for large deviations. Specifically, for errors below a predefined threshold δ, it penalizes deviations quadratically, ensuring smooth convergence, while for errors above the threshold, it applies a linear penalty, thereby reducing the influence of extreme values on the overall training process.

Recursive stage

Once trained, the SGN model enables efficient online recursive simulation. The process (Figure 6.5) starts by initializing a state window with $\tau + 1$ snapshots from DEM data. Then, for each subsequent time step t_{n+1}:

1. Input features are constructed from the current window ending at t_n.

2. The SGN predicts normalized acceleration $\mathbf{A}_{\text{norm}}(t_n)$ and energy increment $\Delta E_{\text{norm}}(t_n)$.

3. Predictions are denormalized to physical units $\mathbf{A}(t_n)$ and $\Delta E(t_n)$.

4. Particle states are advanced using a numerical integrator (e.g., Euler) with the surrogate simulation time step Δt:

$$\mathbf{V}(t_{n+1}) = \mathbf{V}(t_n) + \mathbf{A}(t_n)\Delta t \tag{6.31}$$
$$\mathbf{X}(t_{n+1}) = \mathbf{X}(t_n) + \mathbf{V}(t_n)\Delta t \tag{6.32}$$

   The definition of Δt is critical for ensuring stability during the recursive simulation stage. We adopt a timestep equal to the training dataset's sample frequency (0.0001 s). Using larger timesteps leads to stability issues, as collisions may be missed or excessive penetration between particles and the wall may occur, resulting in exponential error accumulation. This limitation primarily arises from the high-energy dynamics, which inherently involve high velocities and accelerations that make spatial definition overly sensitive to small changes. Nonetheless, this surrogate solving timestep represents a 10,426% relaxation compared to the DEM solving timestep used in Altair® EDEM™ (see Table 3.2).

5. The wall boundary CoM for t_{n+1} is obtained via extrapolation (using pre-calculated periodic splines).

6. New SDF values are recomputed from the updated particle positions $X(t_{n+1})$ and the updated wall geometry. The graph connectivity is then rebuilt: particle–particle edges $e_{ij} \in \mathscr{E}_{pp}$ are added if $\|X_j(t_{n+1}) - X_i(t_{n+1})\| \le r_c$ (we recommend $r_c \le 1.5\,R$), and particle–wall edges $e_{iw} \in \mathscr{E}_{pw}$ are added if $\phi_{\text{SDF}}(v_i, t_{n+1}) \ge \phi_{\text{pw}}$ with $\phi_{\text{SDF}} < 0$ representing zones inside the surface. The choice of ϕ_{pw} is guided by the DEM contact distribution (e.g., capturing ≥ 90% of ground-truth PW contacts; see Figure 6.4). In this work, we set the value at -0.0052 m, but it will vary according to the ball's kinetic energy and physical properties, as they will dictate the depth of penetration. At sharp edges/vertices, the nearest-point direction can be ambiguous; since edge creation uses only ϕ_{SDF}, this does not rely on normals, which are used to compute edge features. Residual near-wall noise is handled in Step 7.

7. Optional correction (snap-back) is applied only for shallow, near-wall penetrations: if $\phi_{\text{SDF}}(v_i, t_{n+1}) \ge \phi_{\text{sb}}$ with $\phi_{\text{pw}} < \phi_{\text{sb}} < 0$, we project to the threshold value:

$$\mathbf{x}_i(t_{n+1}) \leftarrow \mathbf{x}_i(t_{n+1}) - \left(\phi_i - \phi_{\text{sb}}\right)\mathbf{n}_i, \quad \text{where } \phi_i = \phi_{\text{SDF}}(v_i, t_{n+1}),\ \mathbf{n}_i = \nabla\phi_i / \|\nabla\phi_i\|.$$

   For deeper overlaps ($\phi_{\text{SDF}} < \phi_{\text{sb}}$) no snap-back is used; the contact dynamics resolve the interaction (parameters ϕ_{pw} and ϕ_{sb} are listed in Table 3.3). The optional correction's sole purpose is to prevent nonphysical interpenetration from accumulating due to prediction error or integration overshoot.

8. A new snapshot dictionary for t_{n+1} is assembled using the updated states and recalculated geometric features (including the re-determined contacts/edges).

9. The time window is advanced by removing the oldest snapshot and adding the new one.

This iterative process can be visualized in Figure 6.5 and allows the surrogate model to generate the system's stable evolution over extended periods, driven solely by its own predictions after initialization. A summary of the parameters used for the architecture of the SGN and the recursive stage is presented in Table 3.3. The SGN was implemented using Pytorch 2.1.

Table 6.3: Key parameters for the structured-graph-network (SGN) model and the recursive simulation loop used during roll-outs.

Parameter	Value	Unit / description
Model architecture		
History window size ($\tau + 1$)	7	time steps
Hidden dimension	256	–
MLP layers	4	–
Interaction layers (L_p)	1	–
Huber-loss threshold (δ)	2	–
Loss-weighting factor (α)	3	–
Recursive simulation loop		
Time step (Δt)	1×10^{-4}	s
Integration type	Euler	–
PP contact threshold (r_c)	0.0015	m
PW contact threshold (ϕ_{pw})	−0.0052	m
Snap-back threshold (ϕ_{sb})	−0.0049	m

Contact rules: We adopt $\phi_{\text{SDF}} < 0$ inside the geometry. PP edges are added when $\|\Delta \mathbf{x}\| \le r_c$. PW edges are added when $\phi_{\text{SDF}} \ge \phi_{pw}$. Snap-back is applied only when $\phi_{\text{SDF}} \ge \phi_{sb}$, with $\phi_{pw} < \phi_{sb} < 0$.

6.3. Results and Discussion

In this section, the performance of the SGN surrogate model is evaluated using three distinct assessment methods. First, we measure how accurately the model predicts the bulk dynamics of the standard high-energy ball milling process in the *Emax* machine. Second, we assess performance using a mechanochemistry-specific variable: the global energy dissipation of the system, which continuously increases as collisions occur. Our previous work has shown that this variable can effectively characterize a mechanochemical process from a mechanical standpoint [4]. Third, we evaluate the model's generalization by testing its ability to handle unseen motions and modifications to the base geometry. These evaluations are crucial to demonstrate the potential of the

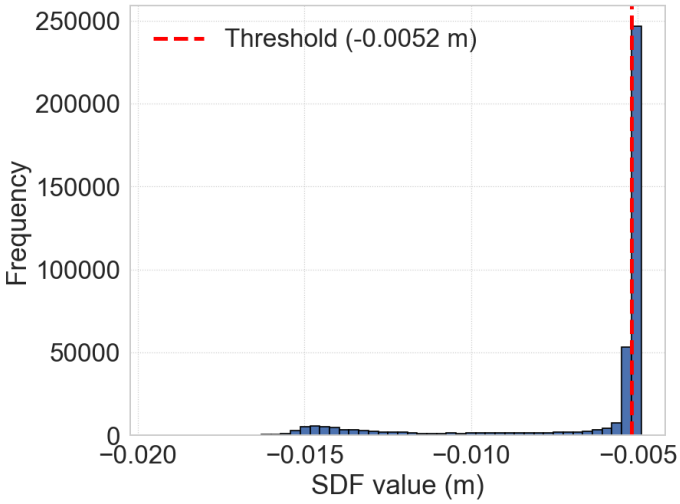


Figure 6.4: Distribution of SDF values for particle-wall contacts. Values closer to zero indicate deeper penetration. The defined threshold of -0.0052 covers over 90% of all particle-wall contacts.

method in the iterative design processes required to scale up and optimize. A sensitivity analysis of critical hyperparameters in Table 3.3, together with an ablation study on the relevance of the global loss $\mathscr{L}_{\text{global}}$, can be found in Appendix A, where we show how the predictive capacity and stability of the model are affected.

6.3.1. Standard high-energy milling process

The model was trained on 4.5 seconds of high-resolution data generated by a DEM simulation, which also serves as the initial benchmark (Figure 6.6). To evaluate the accuracy and stability of the model when predicting the bulk motion of particles in the system, it is run recursively (online) for 15 seconds. At each available time step t, we compare the predicted ball positions $\hat{\mathbf{X}}_t$ to the ground-truth reference $\mathbf{X}_t$ (see Figure 6.7), and compute spatial MSE according to Equation 6.33. We report the MSE because, while the Huber loss was used during training to reduce the influence of occasional high-frequency acceleration spikes, MSE provides a single, widely understood scalar that directly quantifies average squared deviations in particle positions for straightforward benchmarking of bulk-dynamics accuracy.

$$\text{MSE}_t = \frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\mathbf{x}}_t^{(i)} - \mathbf{x}_t^{(i)}\right\|^2 \tag{6.33}$$

Although this is the simplest objective of the model, it is crucial to ensure stability and accurate global predictions, which rely on accurate bulk dynamics and proper collision identification. Looking at Figure 6.7, it is possible to visualize the oscillatory motion of the system and its effect on the predictions. The peaks of these curves coincide with the

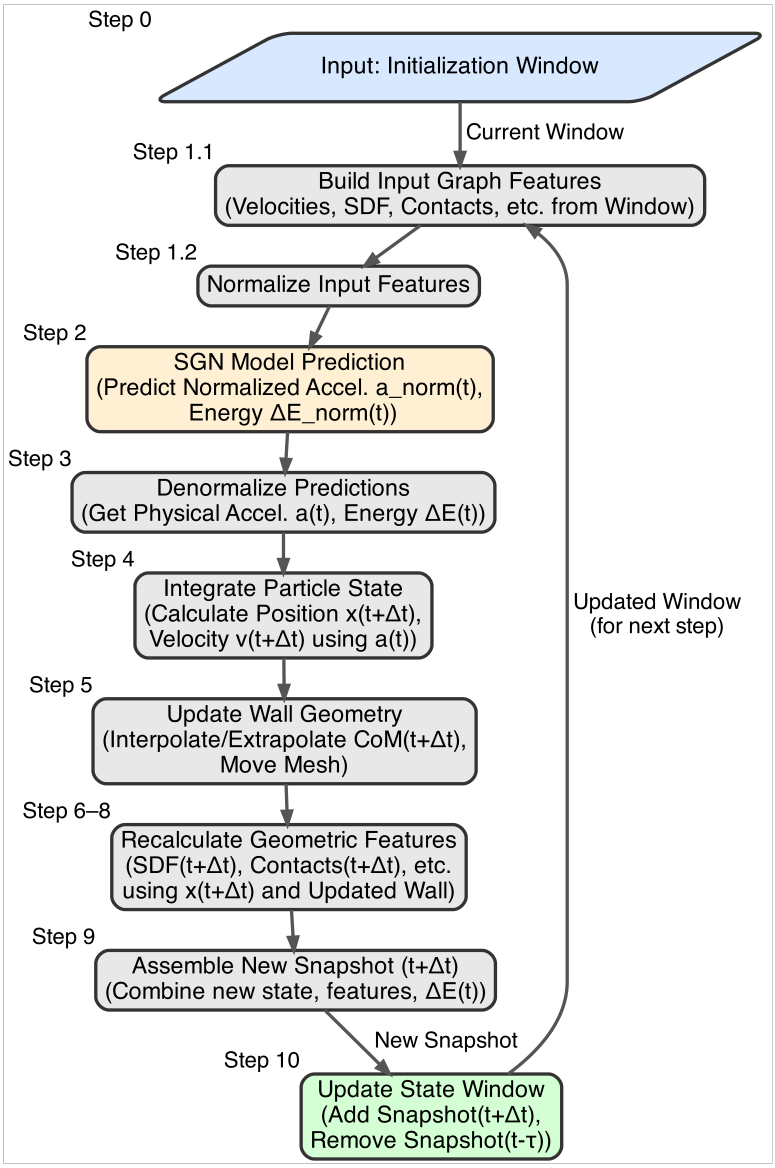


Figure 6.5: Recursive loop flowchart.

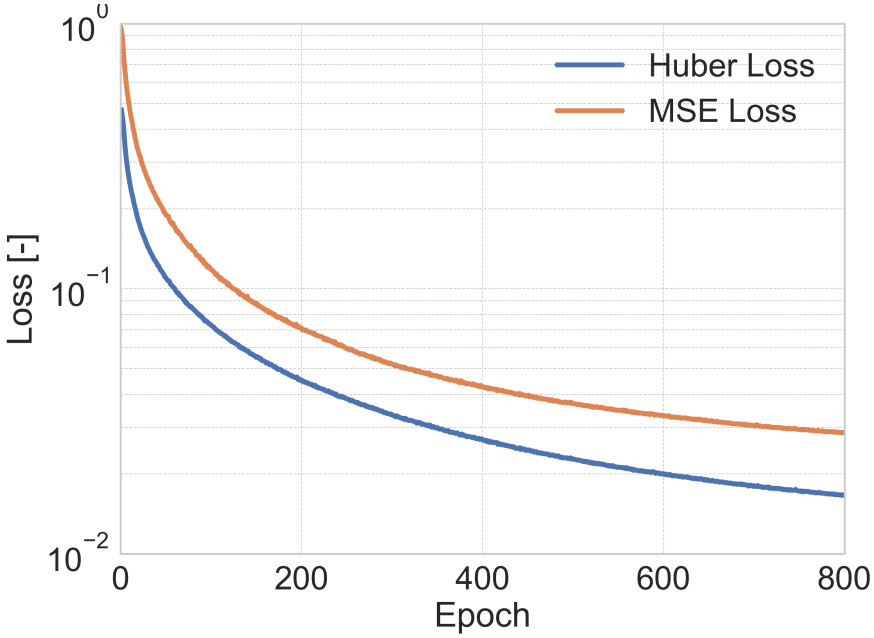


Figure 6.6: Standardized loss history for the SGN surrogate model. Note that the MSE loss is shown for comparison purposes.

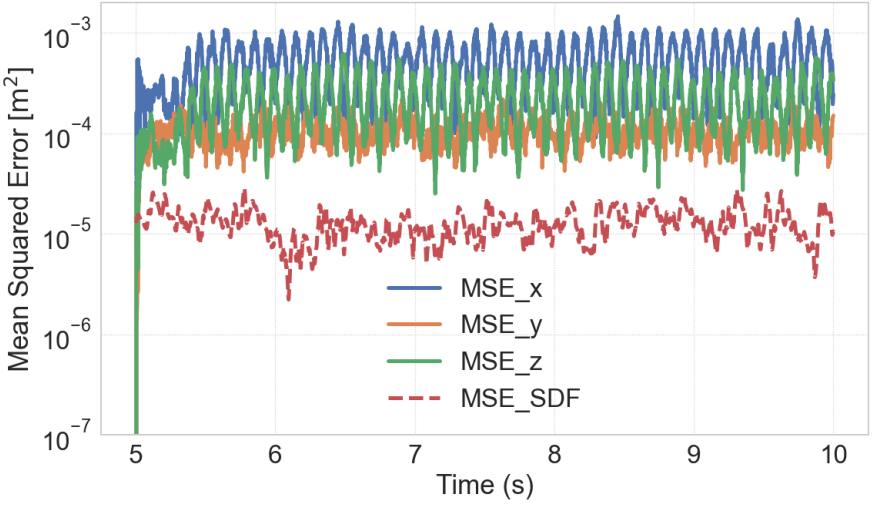


Figure 6.7: MSE loss history for the recursive (online) stage of the SGN. Here, time is measured from t = 5s, marking the start of motion in the DEM simulation, which provides reference data only up to t = 10s. Beyond this point, the model is let run recursively until t = 15s.

moments where the milling jar changes direction in the x-axis. Then, the predicted global energy dissipation can be compared in Figure 6.8.
Thus, the combination of Figure 6.7 and Figure 6.8 shows that the SGN model is capable of accurately representing the bulk motion of the high-energy system and maintaining an accurate track of the energy dissipation that occurs in the system with a stable relative error of 2.78% while using a solving timestep 10,426% larger than the original DEM simulation. To provide a more intuitive representation of the accuracy of the predictions, we used Blender 4.3.2 to reproduce the motion of the milling balls predicted by the SGN and compare them to the original DEM visualization. This can be seen in Figure 6.9.

6.3.2. Unseen motions

Upon establishing that the model can accurately represent the original bulk dynamics and energy dissipation, we now test the model with two new, unseen motions that have a direct effect on the dynamics of the milling balls. To define these arbitrary motions, we use two different Lissajous curves as they operate under the cyclic behavior that a typical high-energy ball mill experiences. The trajectories of the two tested motions are shown in Figure 6.10, and where implemented in Altair® EDEM™ according to Equation 6.34 and Equation 6.35.

Motion 1

$$\begin{aligned} x_1(t) &= 1.7\,\sin(10\pi t) - 0.85\,\cos(20\pi t), \\ z_1(t) &= 1.7\,\cos(10\pi t) - 0.85\,\sin(20\pi t). \end{aligned} \tag{6.34}$$

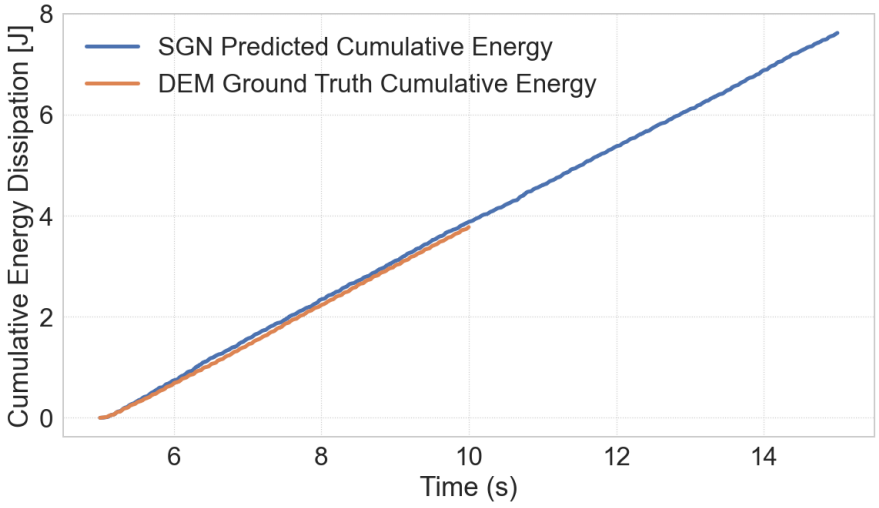


Figure 6.8: Cumulative energy dissipation: comparison of SGN-predicted versus DEM ground-truth energy dissipation. Note that DEM simulation stops at t = 10s.

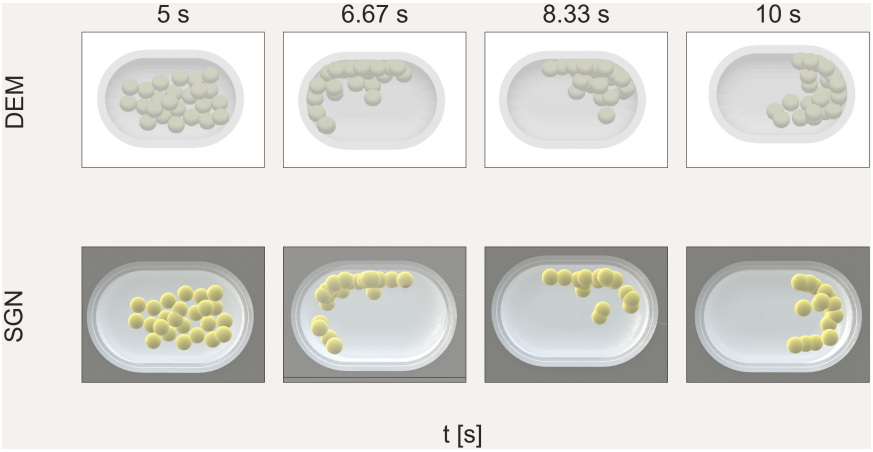


Figure 6.9: Comparison grid of DEM and SGN simulation results.

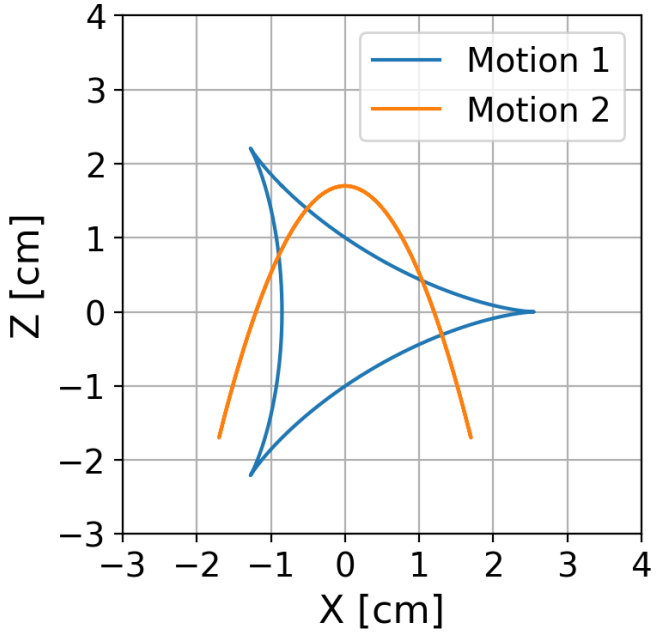


Figure 6.10: Lissajous motions used for testing the SGN.

Motion 2

$$x_2(t) = 1.7\ \sin(10\pi t),$$
$$z_2(t) = 1.7\ \cos(20\pi t). \tag{6.35}$$

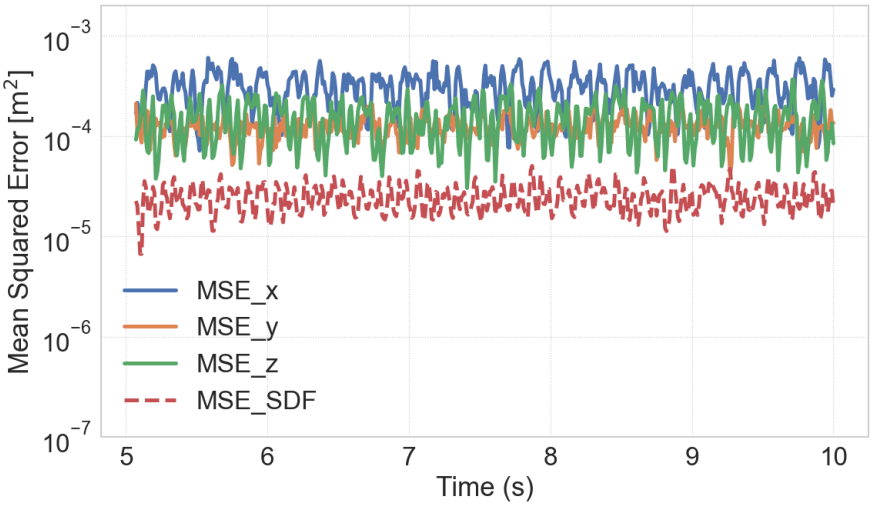
The model is capable of reproducing accurate and stable predictions of the bulk dynamics for both unseen motions, with MSE errors comparable to those of the standard case presented in the previous section, as per Figure 6.11 and can be visualized in Figure 6.12 and Figure 6.13.

The model's ability to predict energy dissipation is substantially weaker than its performance on bulk dynamics. For Motion 1, the predictions maintain a constant relative error of 23.84 %. For Motion 2, the error rises to a steady 46.62 %. Nonetheless, the shape of the time series is captured remarkably well. In both cases, the simulation follows not only the overall trend but also the short-time-scale wiggles (i.e., the small, rapid oscillations superimposed on the mean growth, so the predicted and reference curves rise and fall almost in lock step). This can be quantified by the Pearson correlation coefficient (r = 0.9998 for both cases). Pearson's r measures linear association between two variables: an r of 1 means every peak, dip, and inflection in one series occurs at exactly the same relative level in the other (perfect synchrony of the wiggles), while an r of 0 would indicate no consistent linear pattern. Because r is insensitive to uniform scaling or offsets, the coefficient can be close to 1 even when the absolute magnitudes are biased, as we see here. In other words, the model slightly underestimates the magnitude of each dissipation event, but it gets the timing and relative spacing of those events almost perfectly right. This can be visualized in Figure 6.14 and Figure 6.15. To address this scaling bias, one could expand the training set to cover a broader spectrum of dissipation magnitudes, introducing both lower and higher energy cases. This would encourage the network to learn appropriate scaling factors across broader operational ranges, reducing systematic bias and improving calibration of its outputs. Finally, by exposing the surrogate to diverse collision magnitudes and scenarios, its transferability to new systems should improve, potentially eliminating the need for manual post-processing adjustments. We intend to explore this in a future study.

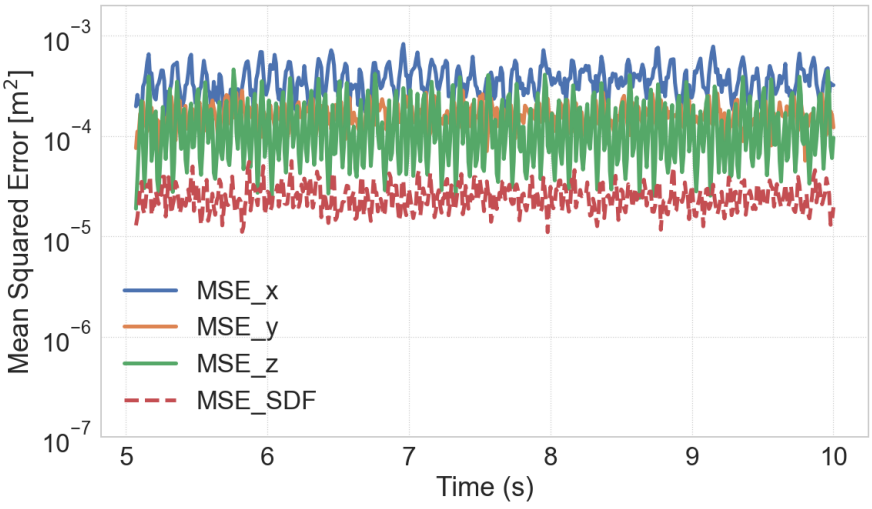
6.3.3. Modifications to original geometry

To test the hypothesis that the underestimation of energy dissipation magnitude in unseen motions arises from a mismatch in the distribution of known collision and velocity features, and at the same time probe the model's ability to handle unseen geometric features, we introduce a slightly tweaked design to boost the collision frequency. Specifically, we insert a cylindrical barrier at the center of the jar (see Figure 6.16) and drive the system with Motion 1 (see Figure 6.10).

By combining this geometry modification with Motion 1, we can also verify whether the surrogate still generates physically plausible particle trajectories and remains numerically stable when both jar shape and motion lie outside its training reference. To illustrate our hypothesis in action, we compare the model's predicted cumulative energy dissipation for the modified geometry + Motion 1 case against the original geometry



(a) Motion 1



(b) Motion 2

Figure 6.11: MSE-loss history comparison for (a) Motion 1 and (b) Motion 2.

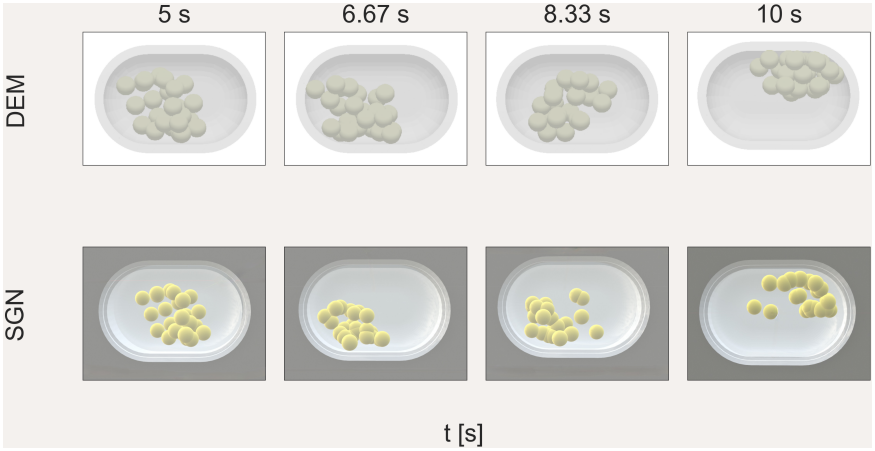


Figure 6.12: Comparison grid of DEM and SGN simulation results for unseen Motion 1.

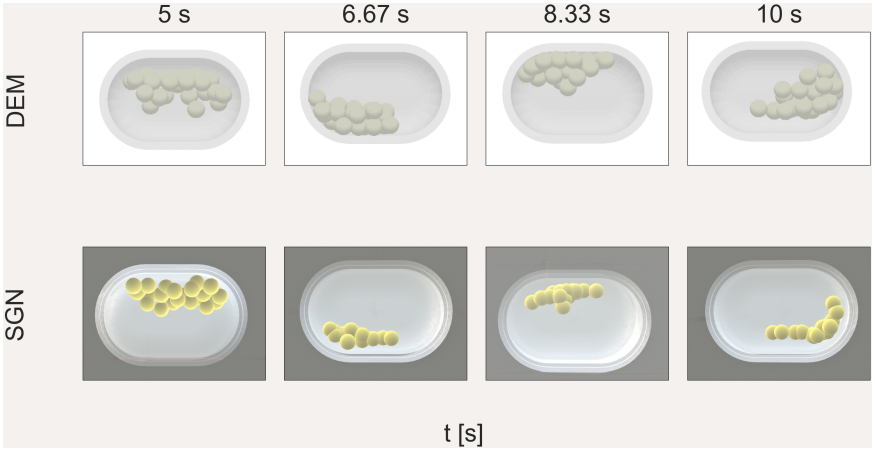


Figure 6.13: Comparison grid of DEM and SGN simulation results for unseen Motion 2.

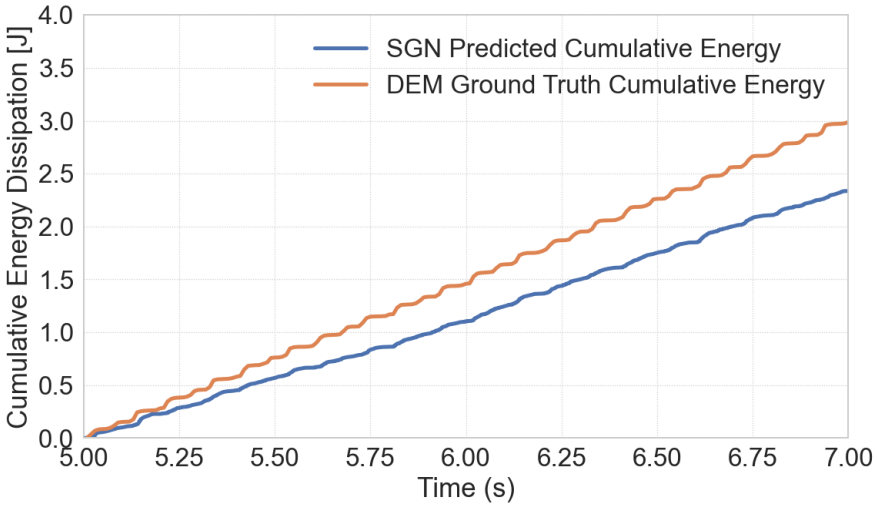


Figure 6.14: Cumulative energy dissipation for unseen Motion 1: comparison of SGN-predicted versus DEM ground-truth energy dissipation. Note that we plot until t = 7s to facilitate the observation of the dissipation's evolution.

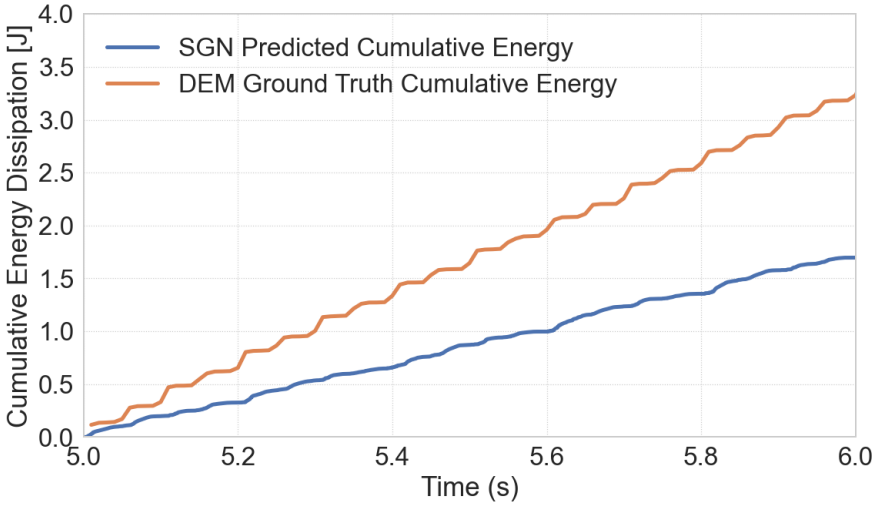


Figure 6.15: Cumulative energy dissipation for unseen Motion 2: comparison of SGN-predicted versus DEM ground-truth energy dissipation. Note that we plot until t = 6s to facilitate the observation of the dissipation's evolution.

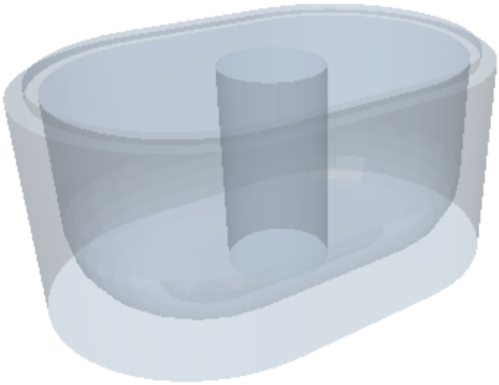


Figure 6.16: Modification to original geometry by adding a cylinder at its center.

+ Motion 1 ground-truth curve (see Figure 6.17). Although this reference no longer corresponds to the actual physics of the modified geometry, it serves as a controlled experiment: if the barrier boosts collision frequency, thus increasing the net impact statistics, then, when we compare its predictions to the original geometry baseline, the underestimation bias should shrink.

Indeed, we observe a substantial drop in relative error to 5.97%, confirming that the original dissipation bias stems from a mismatch in feature distributions. The numerical agreement of this comparison has no physical validity; it exists solely to validate our distribution-matching hypothesis. Furthermore, the snapshots in Figure 6.18 are presented solely for illustration; they demonstrate that the model accurately handles modifications to the original geometry while remaining stable over time.

6.3.4. Runtime

On an Apple M1 Max (10-core CPU; CPU-only to match the DEM run requirements), an SGN rollout of the Emax case (15 s physical time, $\Delta t = 10^{-4}$ s) completed in 3,443 s (~57 min), whereas the corresponding DEM simulation to solver completion took 22,734 s (~379 min). Thus, the SGN was about 6.6× faster with an 84.9% shorter runtime. This comparison excludes the subsequent post-processing of DEM data required to compute the dissipation target ΔE; including it would further increase the DEM wall clock as it requires manual processing, so we report the solver time only. We observe similar speedups across all our tested cases. In profiling, SDF re-evaluation dominates SGN runtime due to the rapidly moving boundary; this cost is hard to avoid because accurate particle-wall spatial information must be maintained each step for stability. Additionally, increasing Δt leads to missed contacts or boundary escapes given the high rotational speed.

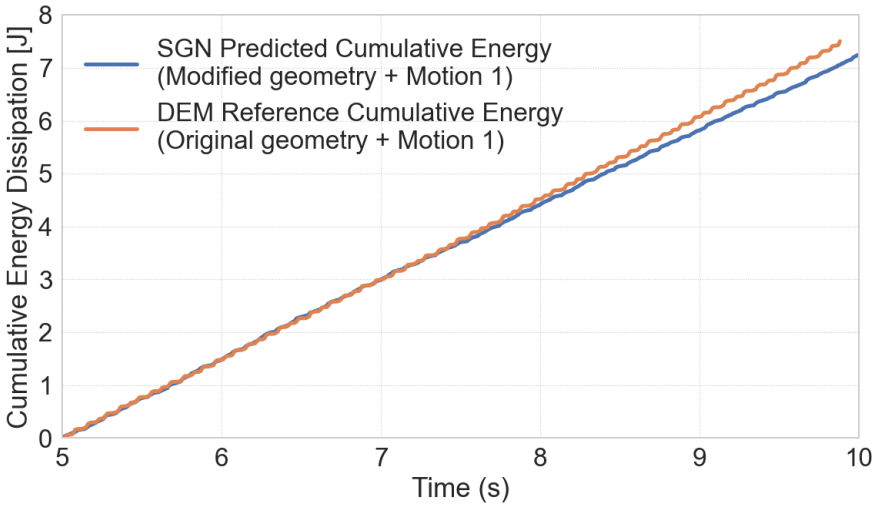


Figure 6.17: Cumulative energy dissipation distribution□matching experiment: comparison of SGN predicted energy versus the DEM baseline reference (original geometry + Motion 1). Note that this ground-truth comparison has no physical validity but serves to illustrate how slowing particle kinematics reduces the under-prediction bias.

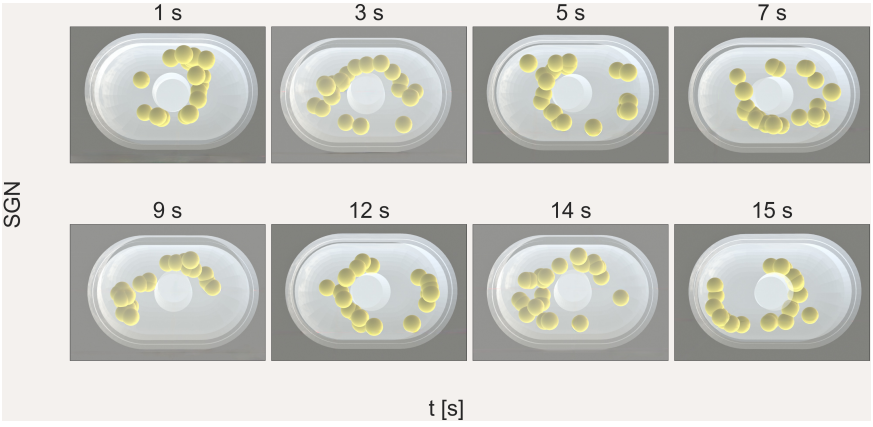


Figure 6.18: SGN simulation results for modified geometry and Motion 1.

6.4. Conclusions

In this chapter, we developed and validated a Signed-Distance-Function Graph Neural Network (SGN) that serves as a faithful surrogate for Discrete Element Method (DEM) simulations of high-energy ball milling. By embedding the jar geometry directly through an SDF field that supports dynamic translational motion, the model overcomes the static,

non-translational boundary limitations of previous surrogates. When coupled with a message-passing graph network, it captures both particle–particle and particle–wall interactions effectively. In contrast with previous surrogates, the SGN is also specifically designed to handle the high-impact velocities and collision frequencies characteristic of high-energy milling processes. Training on 45,000 high-resolution DEM snapshots, the SGN simultaneously learns local accelerations and a global energy dissipation metric, providing a physics-aware description that goes beyond purely kinematic fits. This secondary metric is especially informative for mechanochemistry because cumulative dissipated energy directly tracks the mechanical work that activates solid-state reactions, but it could be changed to accommodate other applications. For instance, in fluidized bed reactors, the model could instead output granular temperature to monitor mixing efficiency; in hopper or silo flows, it could report stress accumulation to predict clogging; and in continuous granulation processes, it could track particle residence time to optimize throughput.

Benchmarking against a reference DEM simulation of the *Emax* mill revealed that the surrogate reproduces bulk motion with a mean squared error plateau of ~ 2×10^{-4} m^2 and tracks cumulative energy dissipation with a stable 2.8% relative error. Importantly, these results are obtained with a time step of 1×10^{-4} s, equivalent to a 10,426 % relaxation over the DEM solver step (9.5×10^{-7} s). The ability to function with such relaxed temporal resolution while remaining numerically stable makes the SGN a practical drop-in replacement for exploratory studies, sensitivity scans, and digital-twin applications. On runtime, the CPU-only rollout was ~6.6× faster than the DEM solver, and this comparison excludes the additional DEM post-processing required to compute ΔE_{total}. Profiling shows that SDF evaluation dominates cost because fast boundary motion requires maintaining accurate per-particle spatial context each step, and increasing Δt to reduce these calls proved unviable at high RPM due to missed contacts and boundary escapes; further wall-clock gains are therefore most likely from SDF-derived calculation optimization.

The surrogate also exhibits strong generalization when driven by two previously unseen motions. It preserved stable dynamics and kept trajectory errors within the same bounds observed for the trained motion. Although the absolute scale of energy dissipation was underestimated (about 24% and 47%, respectively), the temporal evolution was captured with a near-perfect correlation ($r \approx 0.9998$), indicating that the model internalizes the underlying physics, but needs broader training data to calibrate energy magnitudes outside of its original set.

Geometric robustness was tested by inserting a cylindrical obstacle, absent from the training set, and combining it with an unseen driving motion. The surrogate remained numerically stable under this combined distribution and, once the barrier increased the ball's collision frequency, its energy-dissipation error (measured against the *original geometry* baseline used for hypothesis testing) fell to roughly 6%. While this comparison is not physically meaningful for the altered jar, it supports the idea that the earlier under-prediction stemmed from a distribution shift rather than a fundamental

limitation of the model.

Overall, these findings demonstrate that SDF-based graph surrogates can compress high-fidelity DEM physics into a lightweight neural simulator that is both fast, transferable, and can handle complex motions and geometries. Such capability opens avenues for iterative milling jar optimisation, large-scale parameter sweeps for mechanochemical scale-up, and closed-loop control strategies.

Several challenges remain:

- *Material diversity*: the current network is trained on a single material system; extending the feature set to particle radius, fill ratio, and restitution distributions is a logical next step.
- *Energy calibration*: the bias observed under out-of-distribution motions points to the need for data augmentation spanning a wider velocity and energy spectrum.
- *Uncertainty quantification*: ensemble or Bayesian message-passing variants would provide a direct performance indicator by relating the model's predictive variance to deviations from ground-truth DEM data, yielding confidence intervals around mean predictions. This is crucial because granular processes are inherently stochastic and sensitive to initial conditions, so quantifying predictive uncertainty helps detect out-of-distribution scenarios and supports risk-aware decision making in industrial deployment.

The present study marks an advance toward data-driven acceleration of granular process simulations. By releasing all code and preprocessing tools as open source, we hope to catalyze community adoption, foster reproducibility, and ultimately shorten the innovation cycle for sustainable mechanochemical technologies.

Appendix A. Sensitivity analysis

In this section, we analyze how the performance of the surrogate model is affected by a different selection of hyperparameters from those reported in Table 3.3, and an ablation of the global loss $\mathscr{L}_{\text{global}}$ (see Equation 6.28). Since the model has two objective predictions, we employ the following rationale: if the model maintains a stable error when predicting particle dynamics, then we test if the global energy dissipation is predicted accurately. We assess the influence of these hyperparameters based on the model's capacity to predict the standard milling process in the *Emax* machine.

History window size

The selection of the history window size is critical to allow the model to learn sufficient information about how a typical collision evolves in the system. Too long a window will effectively introduce noise that the model will not be able to resolve and lead to unstable dynamic predictions (see Figure A.2). Our results also indicate that selecting a window size that covers less than the median collision duration can lead to stable dynamic

prediction, but will result in less accurate energy dissipation predictions (see Figure A.1). Thus, we recommend selecting a window size that encompasses at least the median duration, and does not exceed this time by more than 40%. In the case of this work, the median collision duration is 0.0005 seconds, and each window frame contributes 0.0001 seconds.

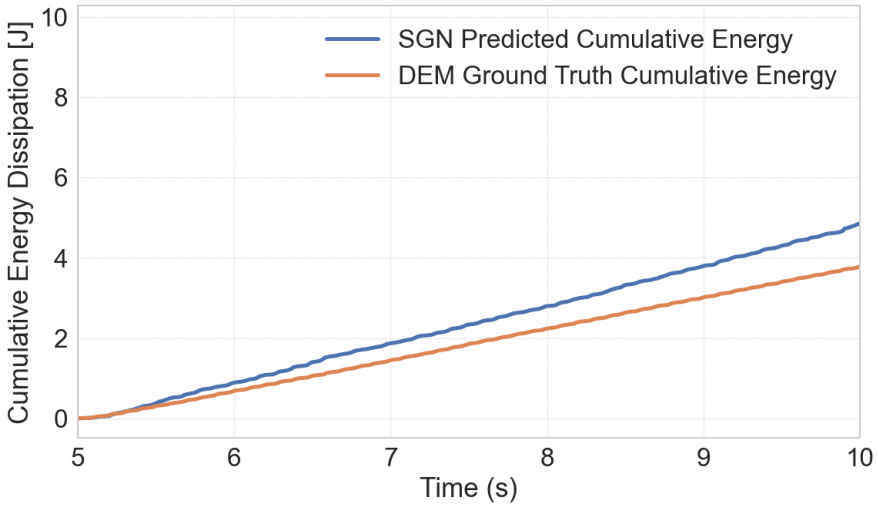


Figure A.1: Cumulative energy dissipation for window size = 3

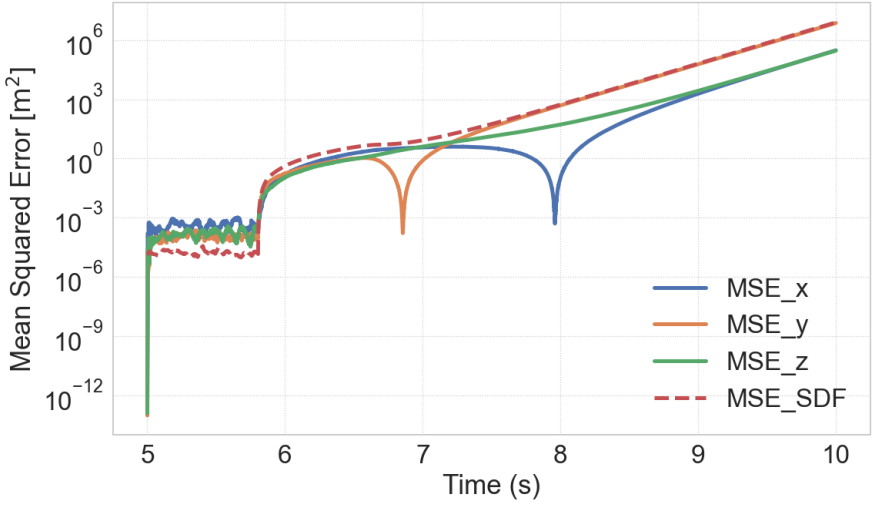


Figure A.2: MSE loss history for window size = 9

Interaction layers

The number of interaction layers defines how many times node messages are passed and aggregated before making a prediction. Intuitively, deeper interaction modules allow the model to capture more complex multi-body effects, but given the small number of particles in the system, the number of collisions with more than 2 elements involved is relatively scarce. As a result, using more than 1 layer leads to unstable predictions. When using 2 interaction layers, the dynamic predictions of the particle manage to stay within the jar bounds, but they become chaotic, leading the energy dissipation predictions to grow without bound (see Figure A.3). Using even more layers (i.e. 4) leads to unstable dynamic predictions (see Figure A.4). Nonetheless, we suspect that a system with significantly more multi-body interactions will necessitate more interaction layers. We recommend consulting the studies mentioned in the introduction since they cover systems where multi-body interactions are more prevalent.

Neuron number

The number of neurons in the model's hidden layers defines its representational capacity for capturing the nonlinear dynamics of particle collisions. Too few neurons constrain the model's ability to predict energy dissipation accurately, although the dynamics remain stable and precise (see Figure A.5).

Conversely, an excessively large hidden dimension increases the risk of overfitting to training noise, which can manifest as unstable long-term predictions or reduced generalization capability. However, determining the precise network width at which overfitting first appears would require a broader hyperparameter sweep, which was unnecessary for this work.

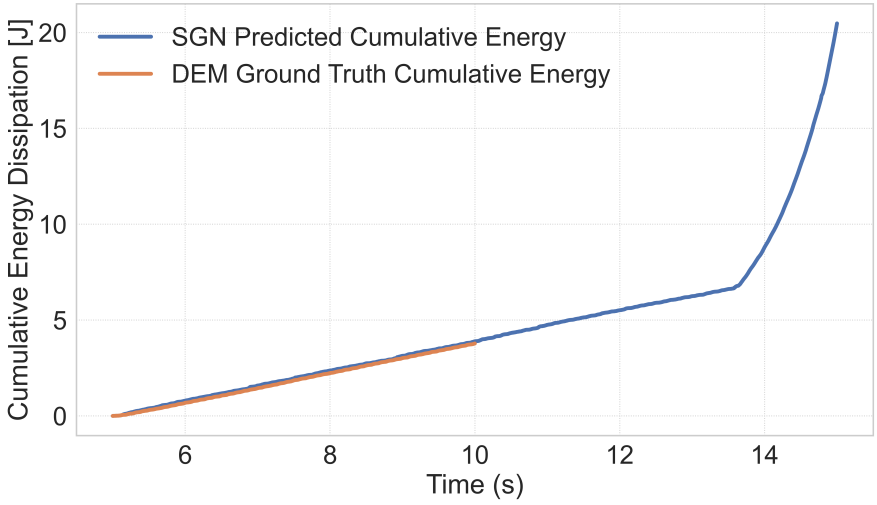


Figure A.3: Cumulative energy dissipation for interaction layer size = 2

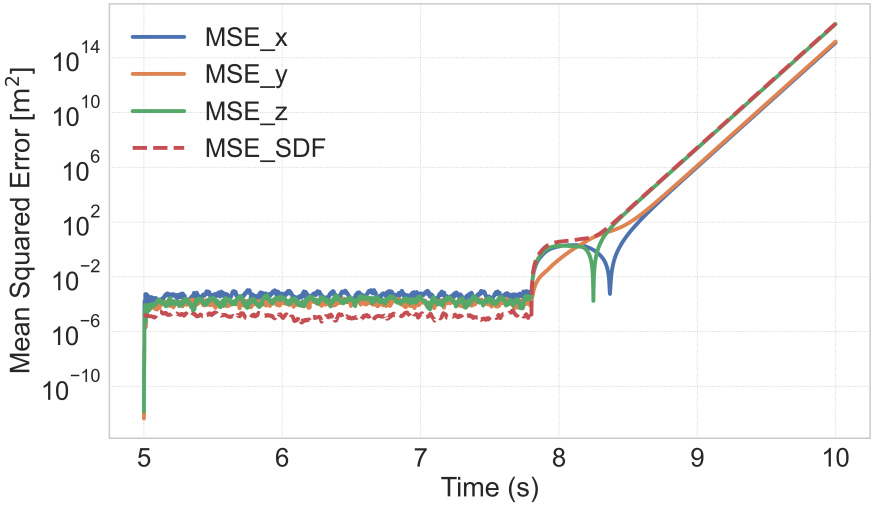


Figure A.4: MSE loss history for interaction layer size = 4

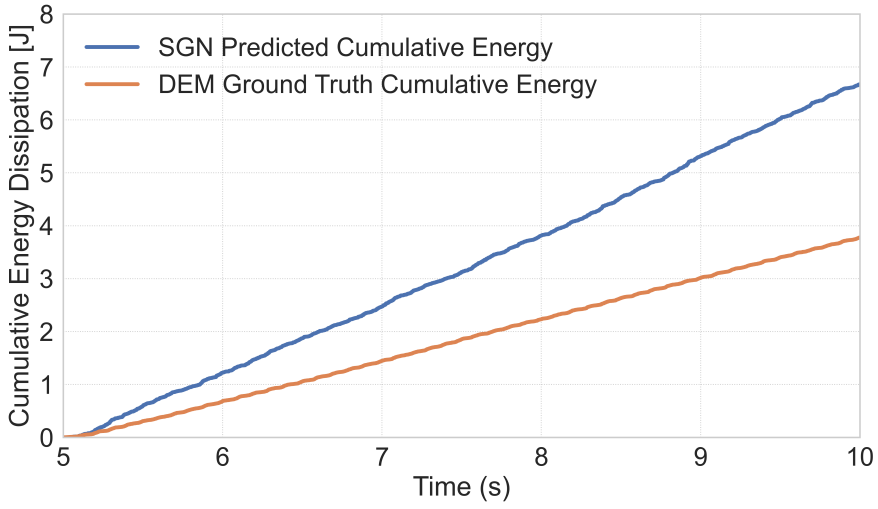


Figure A.5: Cumulative energy dissipation for neuron number = 64

MLP LAYER NUMBER

Lastly, the number of MLP layers defines the depth of successive nonlinear transformations applied to each node's aggregated features, thereby controlling the model's capacity to approximate complex mappings between the current particle states and their future dynamics. Using too few hidden layers (i.e., 2) leads to highly unstable dynamic predictions (Figure A.6).
Similarly, using too many hidden layers can introduce vanishing or exploding gradient issues during training, increase the model's susceptibility to overfitting, and substantially raise computational cost. In our experiments, since four MLP layers achieved stable convergence and accurate predictions, we did not investigate deeper architectures.

PW CONTACT THRESHOLD VALUE (ϕ_{pw})

The PW contact threshold (ϕ_{pw}) defines the near-wall region in which a particle-wall edge is created (edges added when $\phi_{SDF} \geq \phi_{pw}$, with $\phi_{SDF} < 0$ inside the jar). A shallower threshold (less negative, closer to zero) narrows this band and can miss near-wall interactions or lead to particle escape from the domain. A deeper threshold (more negative) widens the band, increasing PW edge density. We probe with two perturbations around the baseline -0.0052 m: a shallower -0.0049 m and a deeper -0.0055 m, adjusting the snap-back level as $\phi_{sb} = \phi_{pw} + 0.0003$ m to maintain $\phi_{pw} < \phi_{sb} < 0$.

In both cases, the kinematics of the system remain stable, but the energy dissipation prediction behavior differs. With a PW threshold closer to zero, grazing contacts are minimized, and the active PW-edge set shrinks to only the most wall-proximal, high-intensity events. This stronger set inflates the predicted per-step

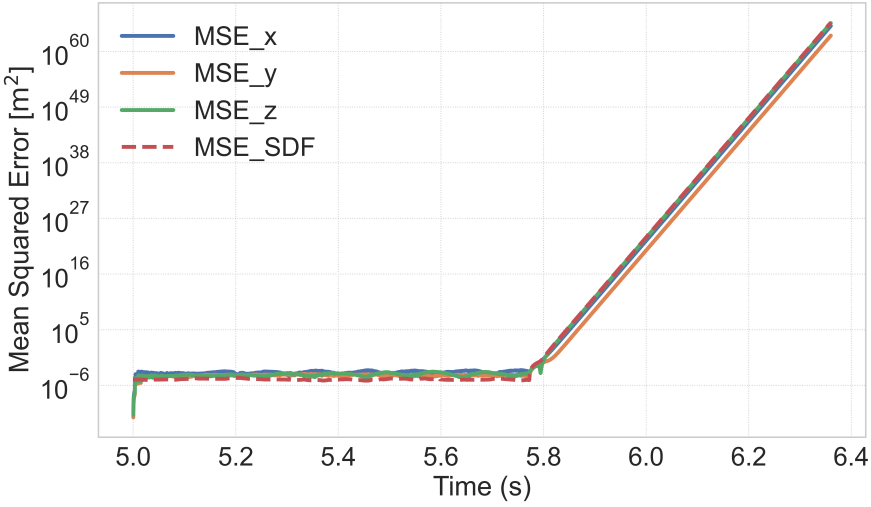


Figure A.6: MSE loss history for MLP layer size = 2

dissipation, so the cumulative energy overshoots the DEM ground truth; in this case, the relative error reaches 15.45% (see Figure A.7).

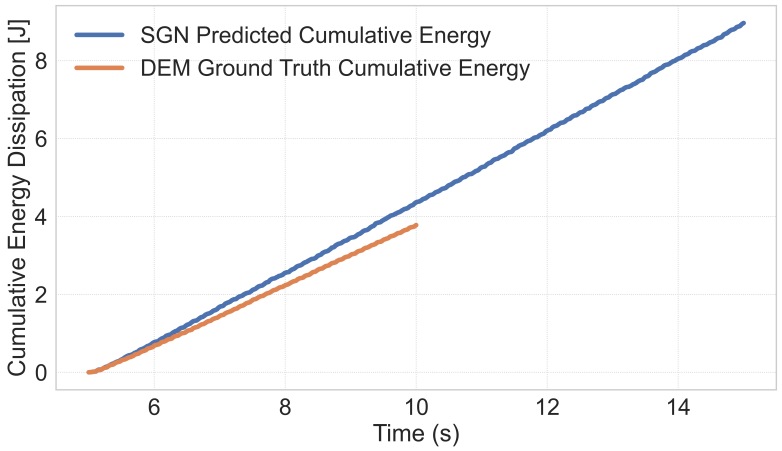


Figure A.7: Cumulative energy dissipation for shallower ϕ_{pw} = -0.0049 m

Conversely, when the PW threshold becomes more negative, more grazing interactions are detected, the active PW-edge set enlarges, and mean pooling over this larger, lower-intensity set reduces the predicted increment. It should be noted that deeper thresholds increase the risk of false positive detections. In this case, the prediction undershoots the DEM reference, with a relative error of 4.18% (Figure A.8). For a safe rollout, we recommend selecting a threshold that recovers ≥ 90% of DEM PW ground truth contacts.

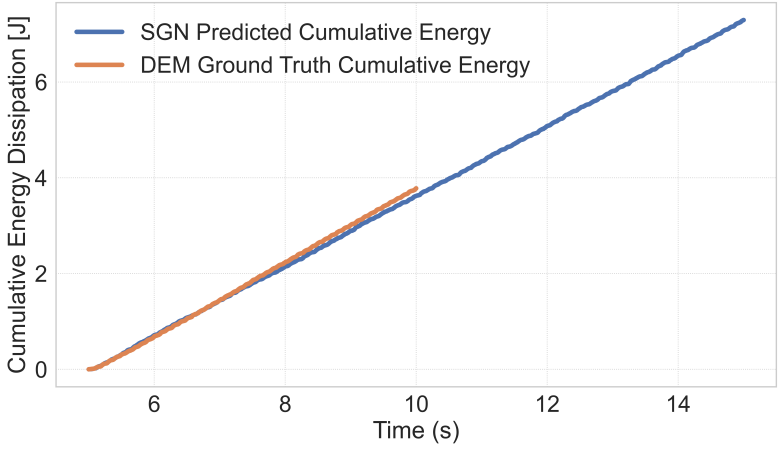


Figure A.8: Cumulative energy dissipation for shallower ϕ_{pw} = -0.0055 m

Global-loss ablation ($\mathscr{L}_{\text{global}}$)

Finally, we perform a ablation directly from Equation 6.28 by setting $\mathscr{L}_{\text{global}} = 0$, so the objective reduces to $\mathscr{L}_{\text{total}} = \alpha \mathscr{L}_{\text{node}}$ (all other settings in Table 3.3 unchanged). Without supervision on the global head, the predicted per-step dissipation becomes severely miscalibrated: the cumulative curve overshoots DEM with a relative error 514.35% at the end of the rollout, even though the kinematics remain stable (see Figure A.9). This confirms that $\mathscr{L}_{\text{global}}$ is necessary to calibrate the magnitude of dissipation; otherwise, its scale is unconstrained and drifts upward.

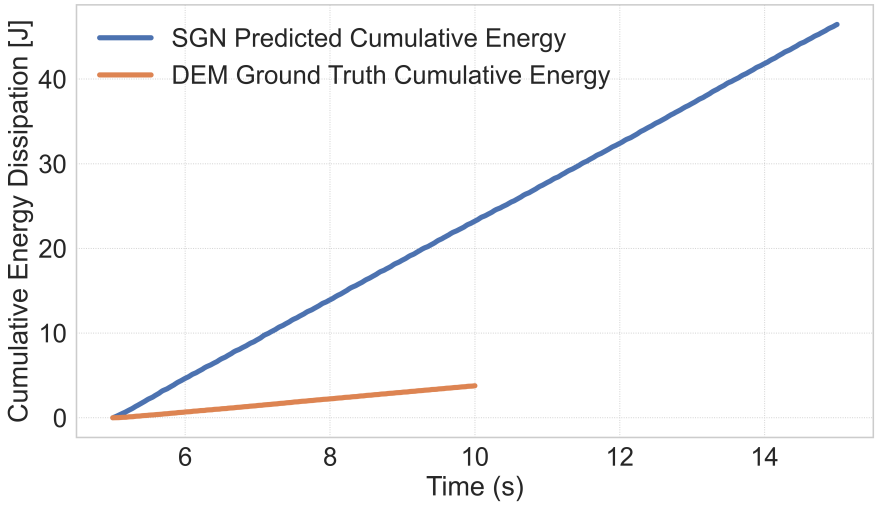


Figure A.9: Ablation of the global loss $\mathscr{L}_{\text{global}}$.

References

[1] A. Sanchez-Gonzalez, J. Godwin, T. Pfaff, R. Ying, J. Leskovec and P. W. Battaglia. *Learning to Simulate Complex Physics with Graph Networks*. 2020. arXiv: 2002.09405 [cs.LG]. URL: https://arxiv.org/abs/2002.09405.

[2] S. Li and M. Sakai. 'Advanced graph neural network-based surrogate model for granular flows in arbitrarily shaped domains'. In: *Chemical Engineering Journal* 500 (2024), p. 157349. ISSN: 1385-8947. DOI: https://doi.org/10.1016/j.cej.2024.157349. URL: https://www.sciencedirect.com/science/article/pii/S1385894724088405.

[3] Y. Zhao, H. Li, H. Zhou, H. R. Attar, T. Pfaff and N. Li. *A review of graph neural network applications in mechanics-related domains*. 2024. arXiv: 2407.11060 [cs.LG]. URL: https://arxiv.org/abs/2407.11060.

[4] S. Garrido Nuñez, D. L. Schott and J. T. Padding. 'Predictive models for energy dissipation in mechanochemical ball milling'. In: *Powder Technology* 457, 120919 (2025). ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2025.120919. URL: https://www.sciencedirect.com/science/article/pii/S0032591025003146.

[5] C. Burmeister, M. Hofer, P. Molaiyan, P. Michalowski and A. Kwade. 'Characterization of Stressing Conditions in a High Energy Ball Mill by Discrete Element Simulations'. In: *Processes* 10 (Apr. 2022), p. 692. DOI: 10.3390/pr10040692.

[6] W. Chen, M. Schoenitz, T. Ward, R. Dave and E. Dreizin. 'Numerical Simulation of Mechanical Alloying in a Shaker Mill by Discrete Element Method'. In: *KONA Powder and Particle Journal* 23 (Jan. 2005), pp. 152–162. DOI: 10.14356/kona.2005018.

[7] A. Oliveira, V. Rodriguez, R. de Carvalho, M. Powell and L. Tavares. 'Mechanistic modeling and simulation of a batch vertical stirred mill'. In: *Minerals Engineering* 156 (2020), p. 106487. ISSN: 0892-6875. DOI: https://doi.org/10.1016/j.mineng.2020.106487. URL: https://www.sciencedirect.com/science/article/pii/S0892687520303071.

[8] V. A. Rodriguez, L. Ribas, A. Kwade and L. M. Tavares. 'Mechanistic modeling and simulation of a wet planetary ball mill'. In: *Powder Technology* 429 (2023), p. 118901. ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2023.118901. URL: https://www.sciencedirect.com/science/article/pii/S003259102300685X.

[9] M. Sakai. 'How Should the Discrete Element Method Be Applied in Industrial Systems?: A Review'. In: *KONA Powder and Particle Journal* 33 (2016), pp. 169–178. DOI: 10.14356/kona.2016023.

[10] N. Govender, D. N. Wilke, C.-Y. Wu, R. Rajamani, J. Khinast and B. J. Glasser. 'Large-scale GPU based DEM modeling of mixing using irregularly shaped particles'. In: *Advanced Powder Technology* 29.10 (2018), pp. 2476–2490. ISSN: 0921-8831. DOI: https://doi.org/10.1016/j.apt.2018.06.028. URL: https://www.sciencedirect.com/science/article/pii/S0921883118303054.

[11] V. Martinez, T. Stolar, B. Karadeniz, I. Brekalo and K. Užarević. 'Advancing mechanochemical synthesis by combining milling with different energy sources'. In: *Nature Reviews Chemistry* 7.1 (Jan. 2023), pp. 51–65. ISSN: 2397-3358. DOI: 10.1038/s41570-022-00442-1. URL: https://doi.org/10.1038/s41570-022-00442-1.

[12] J.-L. Do and T. Friščić. 'Mechanochemistry: A Force of Synthesis'. In: *ACS Central Science* 3.1 (2017). PMID: 28149948, pp. 13–19. DOI: 10.1021/acscentsci.6b00277. eprint: https://doi.org/10.1021/acscentsci.6b00277. URL: https://doi.org/10.1021/acscentsci.6b00277.

[13] M. Fransen, A. Fürst, D. Tunuguntla, D. N. Wilke, B. Alkin, D. Barreto, J. Brandstetter, M. A. Cabrera, X. Fan, M. Guo, B. Kieskamp, K. Kumar, J. Morrissey, J. Nuttall, J. Ooi, L. Orozco, S.-A. Papanicolopulos, T. Qu, D. Schott, T. Shuku, W. Sun, T. Weinhart, D. Ye and H. Cheng. 'Towards Scientific Machine Learning for Granular Material Simulations: Challenges and Opportunities'. In: *Archives of Computational Methods in Engineering* (2025). ISSN: 1886-1784. DOI: 10.1007/s11831-025-10322-8.

[14] Y. Choi and K. Kumar. 'Graph Neural Network-based Surrogate Model for Granular Flows'. In: *arXiv* (2023). DOI: 10.48550/arXiv.2305.05218. eprint: 2305.05218. URL: https://arxiv.org/abs/2305.05218.

[15] Y. Jiang, X. Chen and coauthors. 'Integrating Graph Neural Network-Based Surrogate Modeling with Inverse Design for Granular Flows'. In: *Industrial & Engineering Chemistry Research* (2024). DOI: 10.1021/acs.iecr.4c00692. URL: https://pubs.acs.org/doi/10.1021/acs.iecr.4c00692.

[16] A. Mayr, S. Lehner, A. Mayrhofer, C. Kloss, S. Hochreiter and J. Brandstetter. *Boundary Graph Neural Networks for 3D Simulations*. 2023. arXiv: 2106.11299 [cs.LG]. URL: https://arxiv.org/abs/2106.11299.

[17] Z. Xie, X. Gu and Y. Shen. 'A Machine Learning Study of Predicting Mixing and Segregation Behaviors in a Bidisperse Solid–Liquid Fluidized Bed'. In: *Industrial & Engineering Chemistry Research* 61.24 (June 2022), pp. 8551–8565. ISSN: 0888-5885. DOI: 10.1021/acs.iecr.2c00071. URL: https://doi.org/10.1021/acs.iecr.2c00071.

[18] L. Lu, X. Gao, J.-F. Dietiker, M. Shahnam and W. A. Rogers. 'Machine learning accelerated discrete element modeling of granular flows'. In: *Chemical Engineering Science* 245 (2021), p. 116832. ISSN: 0009-2509. DOI: https://doi.org/10.1016/j.ces.2021.116832. URL: https://www.sciencedirect.com/science/article/pii/S0009250921003973.

[19] X. Guo, C. Hu, Y. Dai, H. Xu and L. Zeng. 'Learning dense gas-solids flows with physics-encoded neural network model'. In: *Chemical Engineering Journal* 485 (2024), p. 150072. ISSN: 1385-8947. DOI: https://doi.org/10.1016/j.cej.2024.150072. URL: https://www.sciencedirect.com/science/article/pii/S1385894724015584.

[20] N. Kishida, H. Nakamura, S. Ohsaki and S. Watano. 'Development of ultra-fast computing method for powder mixing process'. In: *Chemical Engineering Journal* 475 (2023), p. 146166. ISSN: 1385-8947. DOI: https://doi.org/10.1016/j.cej.2023.146166. URL: https://www.sciencedirect.com/science/article/pii/S1385894723048970.

[21] L. Benvenuti, C. Kloss and S. Pirker. 'Identification of DEM simulation parameters by Artificial Neural Networks and bulk experiments'. In: *Powder Technology* 291 (2016), pp. 456–465. ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2016.01.003. URL: https://www.sciencedirect.com/science/article/pii/S003259101630002X.

[22] A. Hajisharifi, F. Romanò, M. Girfoglio, A. Beccari, D. Bonanni and G. Rozza. 'A non-intrusive data-driven reduced order model for parametrized CFD-DEM numerical simulations'. In: *Journal of Computational Physics* 491 (2023), p. 112355. ISSN: 0021-9991. DOI: https://doi.org/10.1016/j.jcp.2023.112355. URL: https://www.sciencedirect.com/science/article/pii/S0021999123004503.

[23] P. Zhi and Y. Wu. 'Graph Neural Networks for Accelerating the Discrete Element Simulation of Granular Flow'. In: *Proceedings of the 16th World Congress on Computational Mechanics (WCCM 2024) and 4th Pan American Congress on Computational Mechanics (PANACM 2024), Track: Data Science, Machine Learning and Artificial Intelligence*. Ed. by A. Korobenko, M. Laforest, S. Prudhomme and R. Vaziri. Vancouver, Canada: Scipedia, July 2024, pp. 1–11. DOI: 10.23967/wccm.2024.123. URL: https://www.scipedia.com/public/Zhi_Wu_2024a.

[24] V. Sharma and O. Fink. 'Dynami-CAL GraphNet: A Physics-Informed Graph Neural Network Conserving Linear and Angular Momentum for Dynamical Systems'. In: *arXiv preprint arXiv:2501.07373* (2025).

[25] Y. Li, J. Bao, T. Chen, A. Yu and R. Yang. 'Prediction of ball milling performance by a convolutional neural network model and transfer learning'. In: *Powder Technology* 403 (2022), p. 117409. DOI: 10.1016/j.powtec.2022.117409.

[26] DCS Computing GmbH, JKU Linz and Sandia Corporation. *Gran model hertz model*. URL: https://www.cfdem.com/media/DEM/docu/gran_model_hertz.html. (accessed: 05.01.2023).

[27] Altair Engineering Inc. *The Hertz-Mindlin (no slip) model*. URL: https://help.altair.com/EDEM/Creator/Physics/Base_Models/Hertz-Mindlin_(no_slip.htm). (accessed: 08.08.2023).

[28] S. Garrido Nuñez, D. L. Schott and J. T. Padding. 'Optimization of operational parameters in the mechanochemical regeneration of sodium borohydride (NaBH4)'. In: *International Journal of Hydrogen Energy* 97 (2025), pp. 640–648. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2024.11.360. URL: https://www.sciencedirect.com/science/article/pii/S0360319924050511.

[29] Lucefin Group. *X46Cr13 technical card*. URL: https://www.lucefin.com/wp-content/files_mf/1.4034a420c25.pdf. (accessed: 29.08.2023).

[30] P. Santhanam and E. Dreizin. 'Predicting conditions for scaled-up manufacturing of materials prepared by ball milling'. In: *Powder Technology* 221 (May 2012), pp. 403–411. DOI: 10.1016/j.powtec.2012.01.037.

[31] S. Rosenkranz, S. Breitung-Faes and A. Kwade. 'Experimental investigations and modeling of the ball motion in planetary ball mills'. In: *Powder Technology* 212 (Sept. 2011), pp. 224–230. DOI: 10.1016/j.powtec.2011.05.021.

[32] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li and M. Sun. 'Graph neural networks: A review of methods and applications'. In: *AI Open* 1 (2020), pp. 57–81. ISSN: 2666-6510. DOI: https://doi.org/10.1016/j.aiopen.2021.01.001. URL: https://www.sciencedirect.com/science/article/pii/S2666651021000012.

[33] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals and G. E. Dahl. *Neural Message Passing for Quantum Chemistry*. 2017. arXiv: 1704.01212 [cs.LG]. URL: https://arxiv.org/abs/1704.01212.

[34] D. H. P. C. C. (DHPC). *DelftBlue Supercomputer (Phase 2)*. https://www.tudelft.nl/dhpc/ark:/44463/DelftBluePhase2. 2024.

7

Conclusions & recommendations

This thesis set out to propose a feasible pathway for closing the sodium borohydride ($NaBH_4$) cycle by combining systematic experimentation, physics-based modeling, and data-driven surrogates. In doing so, it also aimed to provide a reproducible blueprint for next-generation industrial-scale mechanochemistry. Below, the main conclusions and scientific contributions are described, followed by a set of concrete recommendations for researchers and technologists.

7.1. Conclusions & scientific contributions

- **Operational variables, interactions, and operating windows (chapter 2)**: A fractional design of experiments on the $NaBO_2 \cdot 4\,H_2O + MgH_2$ system revealed that milling time is the single most influential variable, followed by the combined effect of molar ratio and ball-to-powder ratio, whereas rotational speed plays only a secondary role within the explored range. By exploiting these insights, the highest yield reported in the literature was reproduced while running the mill 20% slower, cutting specific energy consumption without sacrificing conversion. The water-only quantification protocol further turned product work-up from a hazardous step into a benign, "ready-to-use" solution stage. Additionally, the wide range of operational settings explored enables a selection based on techno-economic interests, and not purely on yield performance as is typically done in state-of-the-art mechanochemical papers.

- **Linking internal dynamics to machine-level conditions (chapter 3)**: Discrete element method (DEM) simulations established mill-invariant mechanical descriptors, the mean normal and tangential energy dissipated per collision, $\bar{E}_n$ and $\bar{E}_t$, and the specific collision frequency per ball, $f_{\mathrm{col}}/n_{\mathrm{ball}}$, that relate directly to machine-level inputs (speed, fill ratio, ball size). Treating normal and tangential dissipation as two dimensionless groups collapses diverse operating conditions onto universal master curves for the Emax ball mill used in this work. These curves act as a mechanical fingerprint: once matched, any equipment, laboratory, or pilot operating at comparable mechanical work per collision has the potential to deliver similar chemical performance. This provides a common metric for reproducibility and transferability across mills and scales.

- **Normal and tangential partitioning (chapter 4)**: Leveraging the master-curve framework, experiments compared conditions at constant rotational speed and at constant total power to decouple how much energy is delivered from how it is partitioned between shear (tangential) and compression (normal). A low fill ratio of 6% that maximizes tangential dissipation emerged as the optimum for fractional conversion, attaining a record 94% regeneration yield while simultaneously trimming milling time, rpm, BPR, and ball count (−37.5% time, −40% BPR, −34% rpm). Although a tangential bias maximizes fractional yield, the energetic cost of achieving the highest conversion decreases mass-per-watt efficiency. Under constant-power conditions, a balanced regime (10% fill) yielded the highest converted mass per watt ($0.113\,\mathrm{g\,W^{-1}}$) compared to both tangential ($0.090\,\mathrm{g\,W^{-1}}$) and

normal ($0.108\,g\,W^{-1}$) biased cases. In short: keep the jar lightly loaded, tangential-rich, and do not blindly increase energy input; the optimal operating point depends on whether the objective prioritizes absolute conversion or energy-normalized productivity.

- **Yield prediction (chapter 5)**: Coupling the mechanical descriptors with reaction stoichiometry in machine learning models enabled accurate yield forecasts under data scarcity. A two-step Gaussian Process Regression (GPR) ensemble achieved R^2 = 0.83 and the lowest weighted MSE = 93.37 on out-of-sample predictions, consistently outperforming other models. Two-step XGBoost and Random Forest also delivered strong results. Beyond accuracy, model choice should reflect objectives and algorithmic properties (e.g., uncertainty quantification with GPR, interpretability, or speed). These predictors pre-screen unpromising combinations before a single gram of powder is milled, cutting time, materials, and energy costs; performance degrades when multiple variables move outside the training domain, reflecting regime-dependent, non-linear interactions. The dataset is fully transferable and extendable by independent groups and treats mechanical and chemical variables on equal footing.

- **Surrogate modeling (chapter 6)**: To reduce the computational cost of full DEM simulations, we introduced a physics-informed surrogate: a Signed-Distance-Function Graph Neural Network (SGN) trained on high-resolution data. The SGN reproduced bulk trajectories with an MSE of 2×10^{-4} m^2 and tracked cumulative energy dissipation within 2.8%, while operating with a time step 100× larger than DEM. It generalized to unseen jar motions, handled geometry modifications, and remained stable, making digital-twin optimization loops feasible on a laptop rather than an HPC cluster. Embedding boundary information via signed-distance fields lets the graph network respect wall constraints without hand-crafted collision rules, pointing to surrogate models for other particulate reactors. This is a proof-of-concept, and thus data-limited, yet it lays the groundwork for fast mechanics beyond full DEM and can incorporate experimentally measured signals in future iterations.

- **Standardization, reproducibility, and benchmarking**: By proving that chemical yield correlates monotonically with the tangential-to-normal dissipation ratio, and given that these metrics transfer across scales, the work provides missing variables linking ball-mill mechanics to solid-state kinetics. Full disclosure of experimental setup, coupled with master-curve normalization, allows laboratories to replicate not just nominal settings but the underlying impact statistics; this extends to larger-scale equipment. The use of matched constant-speed and constant-power experiments delivered the first systematic, unbiased benchmark of shear-dominated (tangential) versus compression-dominated (normal) loading in a bulk mechanochemical reaction, quantified linear penalties for losing tangential influence, and identified practical limits to energy productivity near full conversion.

- **Technical implications and TRL**: Laboratory batches now reach 94% $NaBH_4$ regeneration at room temperature, with reductions relative to previous state-of-

the-art results: -37.5% shorter time, -40% lower BPR (allowing more powder per batch), -34% lower rpm, and fewer balls, reducing operational costs and iron contamination. Master curves characterize ball mills onto scale-agnostic descriptors, offering practitioners a mechanical fingerprint to specify and compare reactors instead of relying on black-box heuristics; the framework is transferable to other solvent-free syntheses. Because the process is proven only in 125 mL laboratory jars with batch operation, it currently sits at TRL 3-4; nonetheless, the first quantitative scaling laws and data-based models provide key stepping stones toward TRL 5 pilot validation.

Taken together, these results form a closed feedback loop: experiments inform physics; physics-based simulations produce setup-independent, scale-agnostic descriptors and master curves; the descriptors, coupled with stoichiometry, seed machine learning predictors that reduce trial-and-error; and reduced-order surrogates accelerate mechanical data generation and lower computational barriers. In short, the thesis demonstrates a route from empirical screening to predictive design that is transferable across mills and scales, anchored to three mechanistic quantities: the mean normal energy dissipated per collision $\bar{E}_n$, the mean tangential energy dissipated per collision $\bar{E}_t$, and the specific collision frequency per ball $f_{\mathrm{col}}/n_{\mathrm{ball}}$.

7.2. Recommendations for future work

- **Metal wear & contamination**: Even under the most abrasive settings, Fe contamination stayed < 6 wt %, but long-term trends are unknown. Thus, carrying long-term endurance tests is advisable, along with other ball materials. The master curves enable a direct calculation of how the operational variables should change with different materials.

- **Master-curve generality**: The dissipated-energy master curves validated here for the Emax mill and the $NaBH_4$ system should now be tested in intrinsically different mill geometries (attritors, Simoloyer horizontal reactors, vibratory mills), and unrelated reactions such as solvent-free cocrystallisation or Li-ion-battery cathode recycling. Successfully collapsing these additional datasets onto similar dimensionless curves would confirm their universality and scale-up protocols derived from them.

- **Scale-up & continuous operation**: Currently, there are virtually no kilogram-per-hour demonstrations and batch loading incurs idle time. The first decision is whether continuous operation is essential for the commercial viability of $NaBH_4$. If the answer is yes, the DEM master curve logic could, in principle, be transferred to twin-screw extruders; however, the necessary long residence times seem incompatible. If a batch path is acceptable, a Simoloyer mill, which is essentially a high-energy, large-volume ball mill with cooling and inert gas purging, offers a pragmatic TRL 5-6 scale-up option.

- **Integrated system assessment**: Regeneration under MgH_2 produces 6 mol MgO per mol $NaBH_4$. Closing the loop, therefore, demands a viable MgO → MgH_2

route. This could be carbothermic or aluminothermic reduction of MgO followed by hydrogenation, or direct electro-magnesiation powered by renewable electricity. A full techno-economic scenario analysis must integrate this MgO reconversion energy, projected 2030 H_2 prices, and possible maritime fuel taxes to establish break-even costs for each logistics chain.

- **Alternative reducing agents**: Throughout the entirety of this thesis, MgH_2 was used as reducing agent in the mechanochemical regeneration. Other candidates of interest include Mg, Mg_2Si, and Al. The few published results exploring these agents should not be taken as hard limits of the potential of the reactions, as demonstrated by the first and third chapters. Instead, it seems more logical to define a desirable outcome based on economic targets and then optimize the mechanochemical process to arrive at those needs. This thesis lays out a straightforward methodology to study these alternative reactions.
- **Real-time monitoring**: Closed, opaque jars prevent in-situ spectroscopy, keeping reaction kinetics largely empirical. For this, retrofitting PXRD or Raman windows on pilot mills and then feeding spectra to the SGN surrogate for online control can deliver an unprecedented level of control, understanding, and eventually autonomous reactors.
- **Surrogate expansion**: To boost generalizability of the surrogate, it must be exposed to a far richer training dataset, including, but not limited to: multiple mills, fill ratios, and ball materials, plus experimental streams such as acoustic emission, or inline Raman. At the same time, exploring more expressive architectures (e.g. equivariant graph transformers) and embedding active-learning loops that trigger targeted experiments where the model is uncertain will widen its extrapolation capabilities and make it robust enough for truly autonomous, closed-loop control on industrial hardware.
- **Contact-resolved mechanochemistry**: Mechanical descriptors can be fed into FEA (Finite Element Analysis) to map contact events into time-resolved stresses, flash temperatures, and real contact area, summarized as duty cycle thresholds rather than peaks. Those pulse histories can then drive reactive MD (Molecular Dynamics) to return reaction probabilities, rate constants, and rank the roles of heating, plastic work, and pressure across materials.

Final remark

It must be emphasized that mechanochemistry is still a novel, fast-expanding field whose creative limits for reaction optimization are, for now, impossible to pin down. The design and tailoring of dedicated milling hardware, the integration of artificial intelligence routines, and novel ideas such as catalytically active milling balls can each push efficiencies to entirely new limits. Lastly, the methodological framework presented in this thesis offers a template that extends far beyond sodium borohydride, hinting at a future where statistics, mechanics, and chemistry intertwine to deliver sustainable processes at an industrial scale.

Acknowledgements

I've heard a lot of stories about less-than-optimal PhD experiences, and I'm quite happy to say that I don't have a single one to add to the bunch (except maybe my milling machine taking over 4 months to be repaired). I'd argue this overly positive experience is thanks to my promotors, Johan and Dingena. If everyone pursuing a PhD had the privilege of having people like you supporting them, I genuinely think the overall experience would be much better.

Johan, thank you for sharing all your knowledge, expertise, and confidence with me. I can't help but always be impressed by the sheer number of topics you have mastered and how you manage to leverage them! You have helped me learn so much and allowed me to pursue my own "research path", which is something I genuinely appreciate. The freedom you gave me to tackle the different questions and problems that came up over the last few years is something I'll always be grateful for. I think it has helped me substantially to become a better thinker and, in the meantime, also a better researcher. And hey, we're probably the only ones doing mechanochemistry in Delft, and I think we've done quite a decent job! The best proof of this is that we still have so many ideas to try, and just not enough time.

Dingena, thank you for always being so open and available for discussions. Your insights, experience, and guidance have helped me improve the quality of my work so much, and also how I present things (including this thesis!). A lot of the data in this thesis came from your ideas and rigor, and thanks to that, it has allowed me to explore unconventional ways to play with mechanochemistry. One thing that has always impressed me about you and your group is how diverse the research topics are. I find it really fascinating how these little discrete elements can be useful in so many scenarios. Over the last few years, I've been a witness to how you encourage your students, and I've tried to implement many of those aspects in my own way while helping others, which I think has worked quite positively!

Overall, I hope that both of you know how much I have enjoyed working and learning with you.

I would also like to thank my thesis committee, Wiebren de Jong, Fokko Mulder, Klaas Visser, Claudia Weidenthaler, and Ina Vollmer for taking the time to read my thesis and participate in my defence. Specifically, Klaas, I think that if I always had someone speaking so highly about my work, I could probably conquer the world. Ina, it has been such a pleasure and honor to meet you and get a glimpse of your work. I have no doubts that the goals and diligence you and your group have will make an impact in the field

(and I'm excited to read about it!).

This work is ultimately part of the SH2IPDRIVE project, and I would like to thank all my colleagues for their support, encouragement, and positive attitude. I would like to especially thank Julien, Erin, and Marcel. Julien, thank you especially for the first 2 years, for our weekly meetings, and for all the help in the lab. Erin, thank you for being the best front face that the project could have, and for always being so optimistic. And Marcel, thank you for being so friendly, and for always taking care of combining a lot of our work for the public. I don't think I've met anyone who makes cooler presentations than you.

In many ways, being part of TU Delft has always been a dream of mine. I'm not even sure how to thank all the people who shaped the life inflection points that allowed me to make it here. To my professors and friends back in Mexico, in some way, this feels like the conclusion of a long-lasting journey that has spanned over a decade. To the people at TU Eindhoven who decided to give me a scholarship, thank you for giving me the chance to completely change my life trajectory. And, in general, to the Dutch people, thank you for allowing me to join your society.

A mis papás, hermana y abuelos: gracias por todo su apoyo, cariño y paciencia conmigo. Sobre todo por siempre ayudarme a seguir mis sueños, por más ilógicos que sean. Nada de todo esto habría sido posible sin ustedes.

A mi querido perrito Colla, gracias por tanto sin esperar nada a cambio. Espero verte de nuevo.

Curriculum Vitæ

Santiago Garrido Nuñez

Education

2022–2026	PhD Computational Science Delft University of Technology, The Netherlands
2019–2021	MSc Energy Technology Eindhoven University of Technology, The Netherlands Thesis: CFD–DEM study on fluidisation of micron-sized Fe_2O_3
2014–2018	BSc Mechanical Engineering National Autonomous University of Mexico, Mexico City

Conferences

2025	AIChE Annual Meeting, Boston, MA, USA Oral presentation: ”Mechanochemical Regeneration of Sodium Borohydride $NaBH_4$: Optimization, Characterization and Prediction”
2025	INCOME, Berlin, Germany Oral presentation: “Effect of Tangential-to-Normal Stress Distribution in $NaBH_4$ Regeneration”
2025	Mech’Chem, New forces in mechanochemistry, Montpellier, France Oral presentation: “Enhancing mechanochemical sodium borohydride regeneration through process optimization and energy dissipation modeling”

List of Publications

5. S. Garrido Nuñez, D. L. Schott and J. T. Padding. ‘Influence of shear and compressive stress regimes on efficient NaBH4 mechanochemical regeneration’. In: *Chemical Engineering Journal* 525, 170168 (2025). ISSN: 1385-8947. DOI: https://doi.org/10.1016/j.cej.2025.170168. URL: https://www.sciencedirect.com/science/article/pii/S1385894725110127

4. S. Garrido Nuñez, D. L. Schott and J. T. Padding. ‘Accelerating granular dynamics simulations: A graph neural network surrogate for complex high-energy ball milling’. In: *Powder Technology* 468, 121653 (2026). ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2025.121653. URL: https://www.sciencedirect.com/science/article/pii/S0032591025010484

3. S. Garrido Nuñez, D. L. Schott and J. T. Padding. ‘Linking mechanics and chemistry: machine learning for yield prediction in NaBH4 mechanochemical regeneration’. In: *RSC Mechanochem.* 2 (2025), pp. 889–900. DOI: 10.1039/D5MR00076A. URL: http://dx.doi.org/10.1039/D5MR00076A

2. S. Garrido Nuñez, D. L. Schott and J. T. Padding. ‘Optimization of operational parameters in the mechanochemical regeneration of sodium borohydride (NaBH4)’. In: *International Journal of Hydrogen Energy* 97 (2025), pp. 640–648. ISSN: 0360-3199. DOI: https://doi.org/10.1016/j.ijhydene.2024.11.360. URL: https://www.sciencedirect.com/science/article/pii/S0360319924050511

1. S. Garrido Nuñez, D. L. Schott and J. T. Padding. ‘Predictive models for energy dissipation in mechanochemical ball milling’. In: *Powder Technology* 457, 120919 (2025). ISSN: 0032-5910. DOI: https://doi.org/10.1016/j.powtec.2025.120919. URL: https://www.sciencedirect.com/science/article/pii/S0032591025003146