Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a t-rex roaring up into the air (B) a corgi wearing a top hat (C) an orc forging a hammer on an anvil (D) a chimpanzee stirring a bubbling purple potion in a cauldron (E) a wedge of cheese on a silver platter (F) an iguana holding a balloon (G) a frog wearing a sweater (H) a group of squirrels rowing crew (I) an astronaut chopping vegetables in a sunlit kitchen (J) the titanic, aerial view (K) a barbecue grill cooking sausages and burger patties (L) a wooden desk and chair from an elementary school

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a tiger dressed as a military general (B) a kitten standing on top of a giant tortoise (C) a cake in the shape of a train (D) a squirrel DJing (E) a beautifully carved wooden knight chess piece (F) the Statue of Liberty, aerial view (G) a steam engine train, high resolution (H) a bulldozer clearing away a pile of snow (I) a cocker spaniel wearing a crown (J) a plush triceratops toy, studio lighting, high resolution (K) a stack of pancakes covered in maple syrup (L) a humanoid robot playing the cello

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) an expensive office chair (B) a train engine made out of clay (C) a robot stegosaurus (D) a squirrel-lizard hybrid (E) a red convertible car with the top down (F) a cauldron full of gold coins (G) a ficus planted in a pot (H) a teal moped (I) a shiny beetle (J) a pirate collie dog, high resolution (K) a piglet sitting in a teacup (L) a silver humanoid robot flipping a coin

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a terracotta bunny (B) an orangutan using chopsticks to eat ramen (C) a beautiful violin sitting flat on a table (D) a lemur drinking boba (E) a knight holding a lance and sitting on an armored horse (F) a humanoid robot using a laptop (G) a gummy bear driving a convertible (H) A car made out of sushi (I) An iridescent steampunk patterned millipede with bison horns (J) an origami crane (K) a palm tree, low poly 3d model (L) a highland cow

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a chihuahua wearing a tutu (B) an unstable rock cairn in the middle of a stream (C) an exercise bike (D) a small saguaro cactus planted in a clay pot (E) a yorkie dog dressed as a maid (F) a ripe strawberry (G) a dachsund wearing a boater hat (H) an orchid flower planted in a clay pot (I) a fluffy cat lying on its back in a patch of sunlight (J) a ceramic lion (K) a car made out pizza (L) a corgi taking a selfie

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a robot and dinosaur playing chess, high resolution (B) a fox and a hare tangoing together (C) a bumblebee sitting on a pink flower (D) a panda wearing a chef's hat and kneading bread dough on a countertop (E) a metal sculpture of a lion's head, highly detailed (F) a frazer nash super sport car (G) a cat lying on its side batting at a ball of yarn (H) a green dragon breathing fire (I) a squirrel dressed like Henry VIII king of England (J) a ghost eating a hamburger (K) a shiba inu playing golf wearing tartan golf clothes and hat (L) A DSLR photo of pyramid shaped burrito with a slice cut out of it

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a red-eyed tree frog (B) an adorable piglet in a field (C) a teddy bear pushing a shopping cart full of fruits and vegetables (D) a plate piled high with chocolate chip cookies (E) a toad catching a fly with its tongue (F) a Christmas tree with donuts as decorations (G) a monkey-rabbit hybrid (H) Mont Saint-Michel, France, aerial view (I) Coffee cup with many holes (J) a pug wearing a bee costume (K) a fox playing the cello (L) a lion reading the newspaper

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a beagle eating a donut (B) a typewriter (C) a cute steampunk elephant (D) a toy robot (E) a sheepdog running (F) a corn cob and a banana playing poker (G) a squirrel made out of fruit (H) a roulette wheel (I) a bald eagle (J) a plush dragon toy (K) a kitten looking at a goldfish in a bowl (L) a hippo wearing a sweater

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a goose made out of gold (B) a chimpanzee holding a peeled banana (C) a bulldozer (D) a marble bust of a mouse (E) a ceramic upside down yellow octopus holding a blue green ceramic cup (F) a bear dressed in medieval armor (G) a chihuahua lying in a pool ring (H) a confused beagle sitting at a desk working on homework (I) an iridescent metal scorpion (J) a rotary telephone carved out of wood (K) a kangaroo sitting on a bench playing the accordion (L) a pair of tan cowboy boots, studio lighting, product photography

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a few pool balls sitting on a pool table (B) a wizard raccoon casting a spell (C) a lemur taking notes in a journal (D) a banana peeling itself (E) a goat drinking beer (F) a silverback gorilla holding a golden trophy (G) a humanoid robot holding a human brain (H) a lionfish (I) a llama wearing a suit (J) an ice cream sundae (K) a tiger wearing a tuxedo (L) a robot couple fine dining

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a ladybug (B) a squirrel flying a biplane (C) a corgi taking a selfie (D) a pig playing the saxophone (E) a fox holding a videogame controller (F) a beautiful rainbow fish (G) baby elephant jumping on a trampoline (H) a baby dragon drinking boba (I) an amigurumi motorcycle (J) a corgi sneezing (K) a swan and its cygnets swimming in a pond (L) a pig playing a drum set

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) two macaw parrots sharing a milkshake with two straws (B) a pug made out of modeling clay (C) a wide angle DSLR photo of a mythical troll stirring a cauldron (D) a rabbit cutting grass with a lawnmower (E) a yorkie dog wearing extremely cool sneakers (F) a baby dragon hatching out of a stone egg (G) two gummy bears playing dominoes (H) a corgi puppy (I) a classic Packard car (J) an orangutan making a clay bowl on a throwing wheel (K) an adorable kitten lying next to a flower (L) a tiger dressed as a military general

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a pig wearing a backpack (B) an origami bulldozer sitting on the ground (C) a corgi standing up drinking boba (D) an otter lying on its back in the water holding a flower (E) a colorful camping tent in a patch of grass (F) two foxes tango dancing (G) an origami motorcycle (H) a tiger waiter at a fancy restaurant (I) a panda wearing a necktie and sitting in an office chair (J) a chimpanzee wearing headphones (K) a monkey-rabbit hybrid (L) a candelabra with many candles on a red velvet tablecloth

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a chimpanzee dressed as a football player (B) a pigeon standing on a manhole cover (C) cats wearing eyeglasses (D) A DMC Delorean car (E) a plush toy of a corgi nurse (F) a blue tulip (G) a crocodile playing a drum set (H) a straw basket with a cobra coming out of it (I) an origami hippo in a river (J) a woolly mammoth standing on ice (K) a cat wearing a bee costume (L) a squirrel giving a lecture writing on a chalkboard

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).
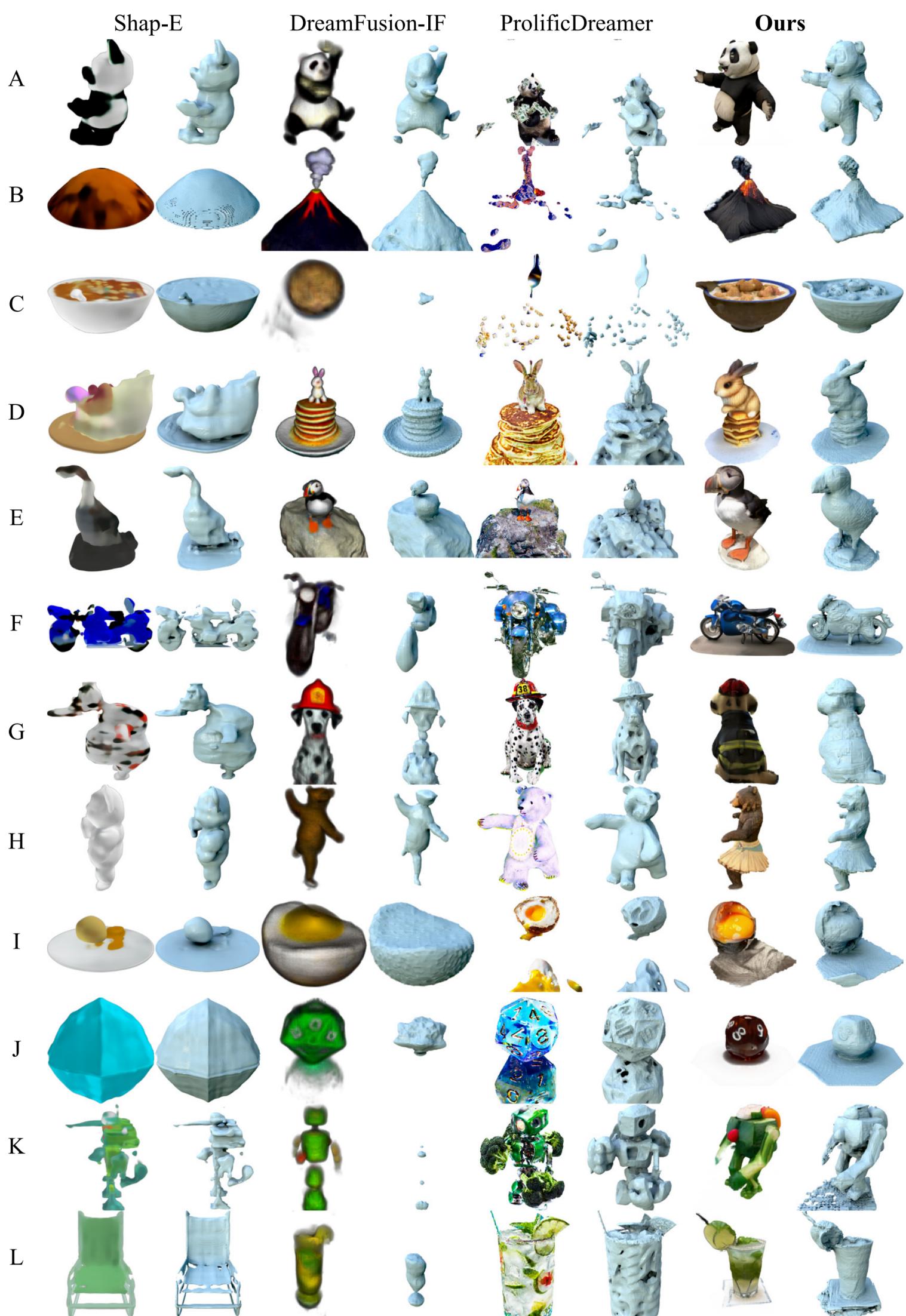
Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a peacock on a surfboard (B) a humanoid robot lying on a couch using a laptop (C) a squirrel holding a bowling ball (D) an eggshell broken in two with an adorable chick standing next to it (E) a hotdog in a tutu skirt (F) edible typewriter made out of vegetables (G) a baby monkey riding on a pig (H) a knight chopping wood (I) a yorkie dog eating a donut (J) a hippo made out of chocolate (K) a plate of fried chicken and waffles with maple syrup on them (L) a drum set made of cheese

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a panda throwing wads of cash into the air (B) an erupting volcano, aerial view (C) a bowl of cereal and milk with a spoon in it (D) a baby bunny sitting on top of a stack of pancakes (E) a puffin standing on a rock (F) a blue motorcycle (G) a dalmation wearing a fireman's hat (H) a bear dancing ballet (I) a cracked egg with the yolk spilling out on a wooden table (J) a 20-sided die made out of glass (K) a robot made out of vegetables (L) a mojito in a beach chair

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a blue lobster (B) an old vintage car (C) a lion's mane jellyfish (D) a recliner chair (E) a squirrel-octopus hybrid (F) a red eyed tree frog, low poly (G) a robot tiger (H) A zoomed out DSLR photo of phoenix made of splashing water (I) a mandarin duck swimming in a pond (J) a green monster truck (K) Neuschwanstein Castle, aerial view (L) a squirrel dressed up like a Victorian woman

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a Space Shuttle (B) a red cardinal bird singing (C) the Imperial State Crown of England
(D) a beagle in a detective's outfit (E) a marble bust of a fox head (F) a fox working on a
jigsaw puzzle (G) a blue poison-dart frog sitting on a water lily (H) a tree stump with an axe
buried in it (I) a vintage record player (J) a robot cat knocking over a chess piece on a board
(K) a majestic sailboat (L) a walrus smoking a pipe

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and
Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) Chichen Itza, aerial view (B) a mouse playing the tuba (C) a corgi lying on its back with its tongue lolling out (D) a cat wearing a lion costume (E) a panda rowing a boat in a pond (F) a rabbit, animated movie character, high detail 3d model (G) a greyhound dog racing down the track (H) an exercise bike in a well lit room (I) a golden goblet, low poly (J) Wedding dress made of tentacles (K) an astronaut riding a kangaroo (L) a chimpanzee looking through a telescope

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) Two locomotives playing tug of war (B) a fox holding a videogame controller (C) a frazer nash super sport car (D) a tiger made out of yarn (E) a model of a house in Tudor style (F) a completely destroyed car (G) a rabbit digging a hole with a shovel (H) a praying mantis wearing roller skates (I) a pile of dice on a green tabletop (J) a shiny silver robot cat (K) a team of butterflies playing soccer on a field (L) a mouse holding a candlestick

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) aerial view of a ruined castle (B) a raccoon astronaut holding his helmet (C) an orangutan playing accordion with its hands spread wide (D) a group of dogs eating pizza (E) a dachsund riding a unicycle (F) a humanoid robot sitting looking at a Go board with some pieces on it (G) a plush t-rex dinosaur toy, studio lighting, high resolution (H) a model of the eiffel tower made out of toothpicks (I) an airplane made out of wood (J) a red pickup truck driving across a stream (K) a tarantula, highly detailed (L) a rainforest bird mating ritual dance

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a steampunk space ship designed in the 18th century (B) a baby grand piano viewed from far away (C) a tiger karate master (D) Michelangelo style statue of an astronaut (E) a red rotary telephone (F) a lemur taking notes in a journal (G) a silver platter piled high with fruits (H) a teddy bear taking a selfie (I) a tiger playing the violin (J) a tiger eating an ice cream cone (K) a kingfisher bird (L) a spanish galleon sailing on the open sea

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a porcelain dragon (B) Sydney opera house, aerial view (C) a drying rack covered in clothes (D) a roast turkey on a platter (E) an astronaut standing on the surface of mars (F) a human skeleton relaxing in a lounge chair (G) a brightly colored mushroom growing on a log (H) a bear playing electric bass (I) a pair of headphones sitting on a desk (J) An anthropomorphic tomato eating another tomato (K) a shiny red stand mixer (L) a dog made out of salad

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).
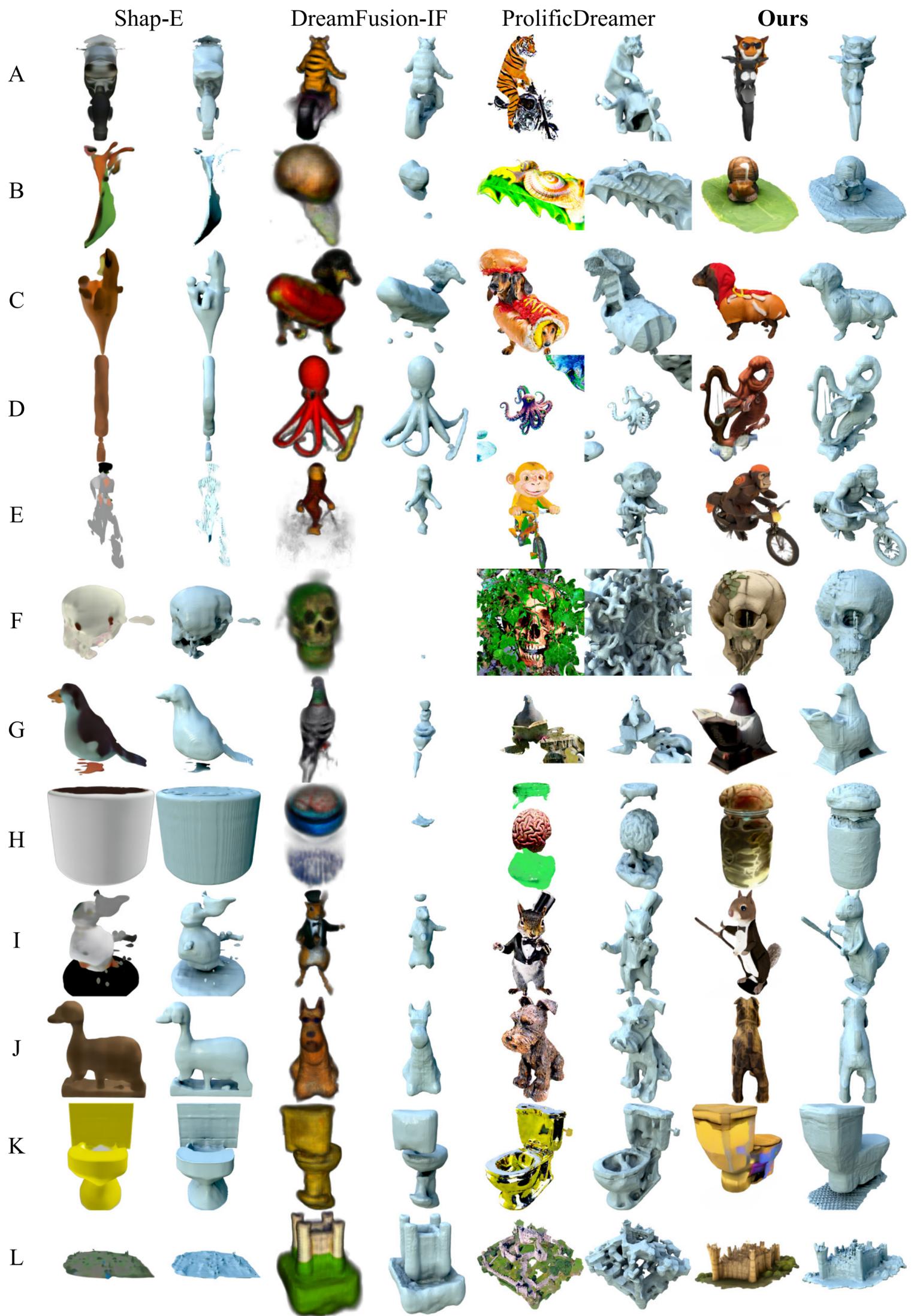
Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a cat magician making a white dove appear (B) a pug made out of metal (C) a bear dressed as a lumberjack (D) a marble bust of a cat, a real mouse is sitting on its head (E) a bichon frise wearing academic regalia (F) a cat with a mullet (G) an amigurumi bulldozer (H) a wine bottle and full wine glass on a chessboard (I) a capybara wearing a top hat, low poly (J) a robot dinosaur (K) a lobster playing the saxophone (L) a gummy bear playing the saxophone

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and Ours (20 seconds).

Uncurated comparison between Shap-E, DreamFusion, ProlificDreamer and our method.

The prompts used in this figure are:
(A) a tiger wearing sunglasses and a leather jacket, riding a motorcycle (B) a snail on a leaf
(C) a dachsund dressed up in a hotdog costume (D) an octopus playing the harp (E) a monkey
riding a bike (F) a human skull with a vine growing through one of the eye sockets (G) a
pigeon reading a book (H) a brain in a jar (I) a squirrel wearing a tuxedo and holding a
conductor's baton (J) miniature schnauzer wooden sculpture, high quality studio photo (K) a
toilet made out of gold (L) an English castle, aerial view

Inference time: Shap-E (6 seconds), DreamFusion (1.5 hours), ProlificDreamer (10 hours), and
Ours (20 seconds).