OPEN ACCESS

A sequential two-step priming scheme reproduces diversity in synaptic strength and short-term plasticity

Kun-Han Lin[a,1] (ORCID), Holger Taschenberger[b,1,2] (ORCID), and Erwin Neher[a,c,2] (ORCID)

Glutamatergic synapses display variable strength and diverse short-term plasticity (STP), even for a given type of connection. Using nonnegative tensor factorization and conventional state modeling, we demonstrate that a kinetic scheme consisting of two sequential and reversible steps of release–machinery assembly and a final step of synaptic vesicle (SV) fusion reproduces STP and its diversity among synapses. Analyzing transmission at the calyx of Held synapses reveals that differences in synaptic strength and STP are not primarily caused by variable fusion probability (p_{fusion}) but are determined by the fraction of docked synaptic vesicles equipped with a mature release machinery. Our simulations show that traditional quantal analysis methods do not necessarily report p_{fusion} of SVs with a mature release machinery but reflect both p_{fusion} and the distribution between mature and immature priming states at rest. Thus, the approach holds promise for a better mechanistic dissection of the roles of presynaptic proteins in the sequence of SV docking, two-step priming, and fusion. It suggests a mechanism for activity-induced redistribution of synaptic efficacy.

synaptic transmission | short-term plasticity | synaptic vesicle priming | calyx of Held | numerical simulation

Chemical synapses change their strength during repetitive use in a synapse type-specific and activity-dependent manner. Such modifications occur on several timescales and define dynamic properties of synaptic networks (1, 2). Elucidating the biophysical mechanisms of synaptic plasticity is essential to understand information processing in circuits (3, 4). Kinetic schemes of synaptic transmission and plasticity provide a theoretical framework to mechanistically and quantitatively interpret functional deficits due to molecular perturbations of the release machinery, either experimentally induced or arising from genetically determined synaptopathies (5).

Short-term changes of synaptic strength such as paired-pulse facilitation (PPF) and short-term depression (STD) have been ascribed to changes in fusion probability (p_{fusion}) of synaptic vesicles (SVs) and/or changes in occupancy of presynaptic release sites (6). Some basic features of short-term plasticity (STP) are captured by a simple scheme postulating one kind of release site to which SVs are recruited, possibly in a Ca^{2+}-dependent manner, before being able to fuse upon action potential (AP) arrival (7, 8). However, numerous observations, including multiple kinetic components of STD (9) and its recovery (10, 11) and diverse STP even among synapses of a given type (12–15) are not easily accounted for by such a simple model.

To more faithfully reproduce the multifaceted features of STP, different multiple-state and/or multiple-site schemes of transmitter release have been proposed (16–24), which include parallel schemes in which more than one kind of SV can bind to one or more kinds of release sites, and sequential schemes in which SVs migrate between different kinds of release sites or states of maturation.

Motivated by converging evidence from molecular biology (25–29), electrophysiology (30–33), live-cell imaging (34, 35), and electron microscopy (EM) (36, 37) emphasizing the reversibility and multistep nature of the priming process, we explore here whether a recently proposed single-site multiple-state scheme of priming and fusion (38) can reproduce variable synaptic strength and diverse STP observed at the calyx of Held, a mammalian glutamatergic model synapse. The proposed kinetic scheme in its basic form (Fig. 1 *A* and *B*) assumes that SVs reversibly dock to a single type of release site and undergo two sequential priming steps to become fusion competent. Considering ultrastructural evidence for distinct docking states (36, 39–42), we refer to the two states and to the SVs residing in those states as loosely (LS) or tightly (TS) docked and SV_{LS} or SV_{TS}, respectively (see *SI Appendix*, Table S1 for a list of abbreviations).

By analyzing AP-evoked EPSC (eEPSC) trains elicited by a wide range of presynaptic firing frequencies with a combination of nonnegative tensor factorization (NTF)

Significance

Central nervous system synapses are diverse in strength and plasticity. Short-term plasticity has traditionally been evaluated with models postulating a single pool of functionally homogeneous fusion-competent synaptic vesicles. Many observations are not easily explainable by such simple models. We established and experimentally validated a scheme of synaptic vesicle priming consisting of two sequential and reversible steps of release–machinery assembly. This sequential two-step priming scheme faithfully reproduced plasticity at a glutamatergic model synapse. The proposed priming and fusion scheme was consistent with the measured mean responses and with the experimentally observed heterogeneity between synapses. Vesicle fusion probability was found to be relatively uniform among synapses, while the priming equilibrium at rest of mature versus immature vesicle priming states differed greatly.

[1]K.-H.L. and H.T. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: eneher@mpinat.mpg.de or taschenberger@mpinat.mpg.de.
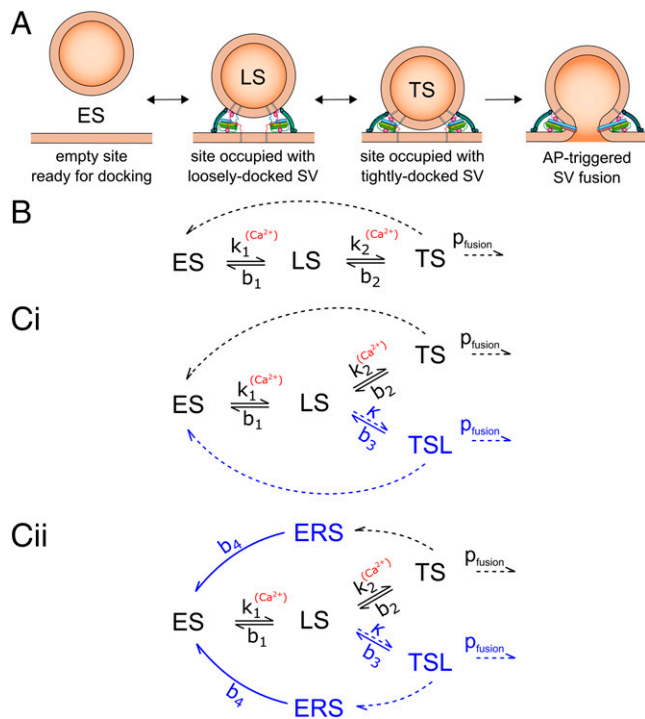
Fig. 1. Diagram of vesicle states and kinetic schemes for the numerical simulation of STP. (*A*) Basic sequential model for priming and fusion. SVs dock to an empty release site (ES) and undergo two priming steps to sequentially transition to the LS and TS states. Only SVs in state TS are fusion competent. (*B*) Kinetic scheme of state transitions for the basic model shown in (*A*). A simple three-state scheme is adequate for reproducing experimental data for f_{stim} = 1–20 Hz. (*C*) An extended reaction scheme with an additional TSL state is required for reproducing experimental data for $f_{stim} \geq$ 50 Hz. SV$_{TS}$s and SV$_{TSL}$s have the same p_{fusion}. The two states TS and TSL differ with respect to their stability. While TS has a lifetime in the range of 3–4 s, TSL relaxes back to LS within ~100 ms. Vacated release sites can either instantaneously return to ES and thereby be immediately available for SV docking (*C*, *i*) or else reside for some time in an ERS (*C*, *ii*). State transitions in (*B*) and (*C*) represented by dashed lines indicate instantaneous transitions, while those represented by solid lines occur with rate constants as shown (see also *SI Appendix*, Table S2). Elements shown in blue in (*C*) extend the kinetic scheme illustrated in (*B*).

and conventional state modeling, we reached a number of conclusions that provide views on the mechanisms of neurotransmitter release and STP. First, approximately 80% of available release sites are occupied at rest by primed SVs, which can be either in the LS or the TS. Second, different initial strength and diverse STP among synapses is primarily due to the variable relative abundance of SV$_{TS}$ over SV$_{LS}$, while p_{fusion} is quite uniform with a high value of ~0.4. Third, for frequencies of 5 to 20 Hz, steady-state release rates scale roughly linearly with presynaptic firing rates, thus maintaining largely frequency-invariant synaptic strength. Finally, at frequencies ≥ 50 Hz, additional kinetic features control release, such as an increase in p_{fusion} during trains, a speed-up of the priming process, and a decline of SV subpool occupancies reducing steady-state release.

Our numerical simulations, which mostly use experimentally determined or NTF-constrained model parameters, faithfully reproduce STP at calyx synapses over a wide range of activity levels and therefore provide a valuable framework for the mechanistic and quantitative interpretation of experimentally induced STP alterations. By emphasizing the multistep nature and reversibility of the priming process, the sequential two-step priming scheme suggests a mechanism for creating functional diversity among synapses and for an activity-induced redistribution of synaptic efficacy during AP trains.

Results

To validate the two-step priming scheme, we chose the following three-step approach. First, we acquired eEPSCs evoked by regular stimulus trains (0.5–200 Hz) from an ensemble of 35 rat calyx synapses. In addition, 100 and 200 Hz eEPSC trains preceded by two or four stimuli at 10 Hz were recorded. Second, all eEPSC train peaks were converted to quantal content, and such data from five to 200 Hz trains were subsequently subjected to NTF analysis. Third, model parameters were then initialized with values derived from NTF analysis or from analytical expressions regarding model predictions for, e.g., paired-pulse-ratio (PPR) and steady-state release at low-frequency stimulation (Eqs. **22–24** [equations referred to here and in the following are provided in the *SI Appendix*]). Subsequently, model fits were optimized by trial-and-error parameter variation to closely reproduce both the respective average eEPSC train response for each stimulation frequency (f_{stim}) and the NTF-derived base functions (BF). Including BFs in this optimization procedure ensures consistency of the model not only with mean train responses but also with the heterogeneity among synapses.

Estimating release from peak amplitudes neglects asynchronous release, which builds up during interstimulus intervals (ISIs) and decays after stimulation (43). However, this release component is small and decreases further during calyx maturation (44). We focused on presynaptic mechanisms regulating STP by 1) choosing P14–16 calyx synapses, which are little affected by AMPAR saturation and desensitization (45, 46), and 2) recording eEPSCs in the presence of 1 mM kynurenic acid (kyn) to alleviate remaining postsynaptic effects (45). Presynaptic inhibition of release via metabotropic GluRs is reportedly low at this age (47), which leaves the modulation of p_{fusion} and the dynamic regulation of priming as major determinants of STP.

Experimental eEPSC train data: Mean time courses and variability among calyx synapses. Fig. 2 *A* and *B* illustrate eEPSCs and the time courses of mean quantal content (m) in response to stimulus trains consisting of 15 APs (0.5, 1, and 2 Hz) or 40 APs (5–200 Hz). For a given synapse, m estimates for the initial eEPSCs (m_1) were similar across all f_{stim}. However, the mean m_1 across all f_{stim} varied nearly 10-fold between synapses (77–739 SVs, coefficient of variation (CV) = 0.42; Fig. 2*D*). The average over all 35 mean m_1 values was 377 ± 27 SV (Fig. 2*B*). During trains, the average m_j decreased monotonically toward a depressed steady state for all but the highest f_{stim}. During 200 Hz stimulation, net facilitation was observed for the average response—i.e., the PPR (m_2/m_1) was on average >1 (*SI Appendix*, Fig. S1A). PPR, when plotted against ISI, converts from facilitation into depression with a time constant of ~21 ms (*SI Appendix*, Fig. S1A) (48). Plotting PPR$_{200\ Hz}$ as a function of m_1 for individual synapses revealed large diversity and a negative correlation—i.e., synapses with larger m_1 predominantly showed paired-pulse depression while those with smaller m_1 often exhibited PPF (Fig. 2*D*). Such correlation is usually interpreted as an indication of heterogeneous initial p_{fusion} ($p_{fusion,1}$), because strong depression in synapses with high $p_{fusion,1}$ would occlude PPF (49). The two-step priming scheme (Fig. 1 *A* and *B*) allows an alternative view, as detailed in *SI Appendix*: It provides an approximate expression for calculating $p_{fusion,1}$ from PPR and STD during 10 Hz trains (Eq. **31**) for each synapse. Plotting such $p_{fusion,1}$ estimates versus the respective m_1 values revealed only weak correlation (Fig. 2*E*). The mean $p_{fusion,1}$ amounted to 0.43 with a CV of 0.20. The latter was substantially smaller than the CV of m_1 values, indicating that the variability of $p_{fusion,1}$ is not the principal cause for heterogeneous synaptic
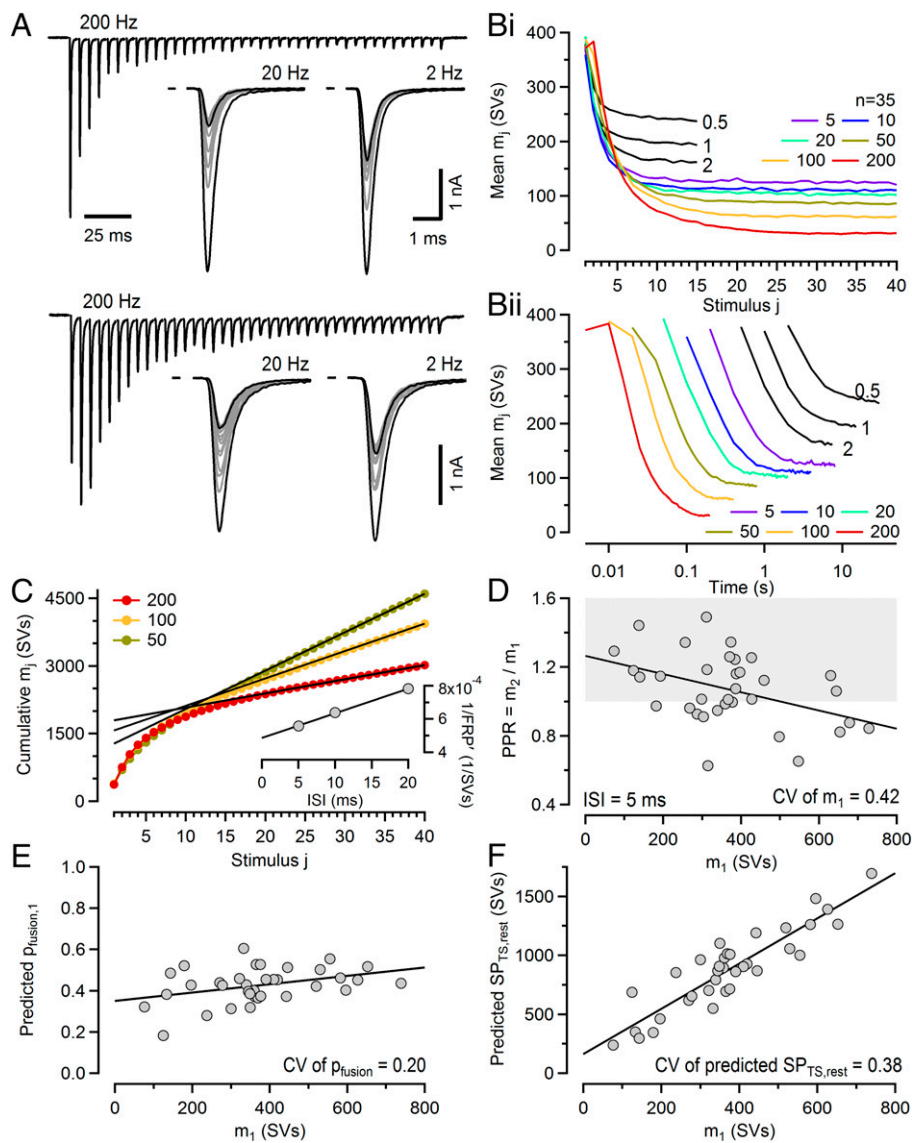
Fig. 2. STP in response to 0.5–200 Hz stimulus trains in post hearing-onset calyx of Held synapses. (*A*) Sample eEPSCs obtained from a strongly depressing (top) and a facilitating (bottom) synapse in response to 200 Hz (*Left*), 20 Hz (*Middle*), and 2 Hz (*Right*) stimulation. Only the initial 15 eEPSCs are superimposed for the 2 Hz and 20 Hz eEPSC trains. Each trace represents an average of three repetitions. (*B*) Mean quantal content (m_j) plotted against stimulus index j (*B, i*) or time (*B, ii*) for each eEPSC. Trains consisted of only 15 stimuli for the lowest three frequencies. The timing of eEPSC$_1$ was offset by one ISI in (*B, ii*) for clarity. Note logarithmic time axis in (*B, ii*). (*C*) Estimating the fast releasing pool (FRP) of SVs from eEPSC trains evoked by 50, 100, and 200 Hz stimulation which provided three FRP′ estimates that are uncorrected for incomplete pool depletion. The relationship between the three 1/FRP′ values and their respective ISIs was subsequently extrapolated to infinite f_{stim} (ISI = 0 ms) to obtain a mean FRP value that is corrected for incomplete pool depletion (*Inset*). (*D*) PPRs (m_2/m_1) for 200 Hz eEPSC trains negatively correlate with initial quantal content (m_1), which varies approximately 10-fold among calyx synapses (73–728 SVs). The gray shaded region indicates PPR >1. (*E* and *F*) Predictions for p_{fusion} (*E*) and $SP_{TS,rest}$ (*F*) for individual synapses obtained from their respective 10 Hz PPR and D_m values according to *SI Appendix*, Eq. 31.

strength. It is rather due to a variable size of the subpool of fusion-competent SVs at rest ($SP_{TS,rest}$), which is readily calculated for each synapse as the ratio $m_1/p_{fusion,1}$. Plotting $SP_{TS,rest}$ estimates as a function of the respective m_1 values revealed a strong linear correlation with an expected slope of ~1/0.43 (Fig. 2*F*). The estimated mean $SP_{TS,rest}$ (880 ± 57 SVs) had a CV of 0.38, which was close to the CV of m_1 values.

In sum, a large heterogeneity among synapses with respect to their SP_{TS} size at rest can explain strong variability in m_1, while p_{fusion} varies only little.

NTF analysis provides p_{fusion} estimates and yields constraints for SV subpool sizes and priming kinetics. We next derived suitable initial guesses for model parameters from NTF analysis. NTF decomposes complex data sets into a linear combination of components (50) and was recently adapted for the analysis of eEPSC trains (51). The algorithm considers time courses of quantal contents during trains as superpositions of contributions by two or more types of signal sources and assumes that their differential relative contributions account for STP diversity among synapses. In the context of the model (Fig. 1*A*), these sources correspond to the release contributions by SVs residing in certain states (such as TS and LS) prior to stimulation. NTF analysis returns the time courses of individual contributions as

BFs (51). For a given f_{stim}, BFs are the same for all synapses. They are normalized to a cumulative sum of 1, such that the product of a BF and the corresponding synapse-specific train quantal content (*M*) represents the time-resolved quantal release of that component for a given synapse (Fig. 3*B* and *SI Appendix*, Fig. S2 *A* and *B*). Two-component NTF fits provide BFs for release contributed by preexisting SV$_{TS}$s (BF$_{TS}$; Fig. 3*A*) and for the combined remaining release originating from SVs, which had been either loosely docked at stimulation onset or were newly recruited during the train (BF$_{LS,RS}$; Fig. 3*B*). The latter release component can be decomposed by subsequent 3-component NTF analysis (Fig. 3*C* and *SI Appendix*, Fig. S2 *A* and *B*). In short, NTF provides time courses of release contributions of SVs or sites, which had been in one of the states of the model (ES, TS, and LS; Fig. 1*A*) at stimulation onset.

NTF analysis cannot provide initial guesses for all model parameters, but it is very instrumental in constraining some of them. Two-component NTF, which is robust (51), separates release contributed by preexisting SV$_{TS}$s from other contributions. BF$_{TS}$ decay rapidly and approach zero after ~5 APs (Fig. 3 *A*). Provided that all preexisting SV$_{TS}$ are consumed, their initial value represents $p_{fusion,1}$ (0.39 ± 0.004 when averaged over all six f_{stim}), and the train quantal content (M_{TS}) associated with BF$_{TS}$ represents an estimate for the subpool SP_{TS} at
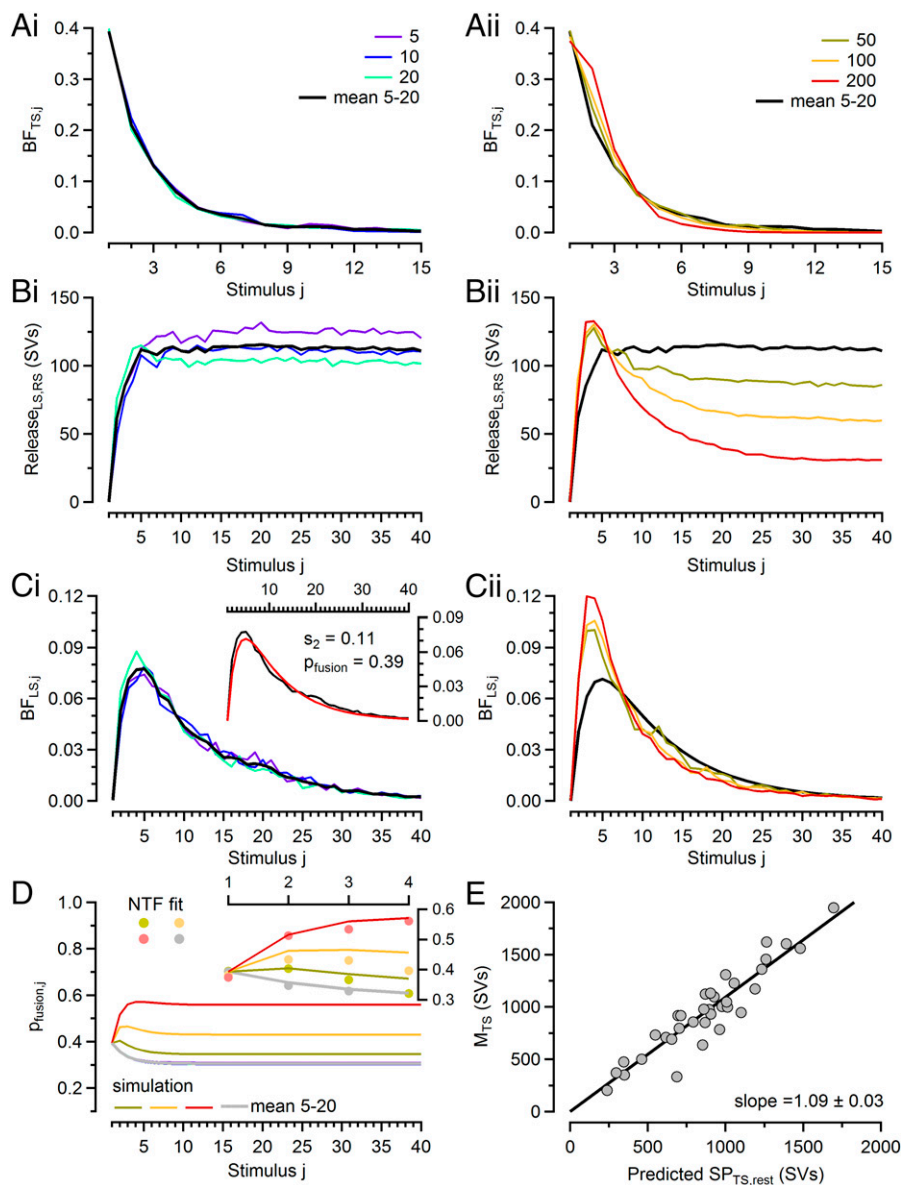
Fig. 3. NTF decomposition analysis of 5–200 Hz eEPSC trains. (*A*) Comparison of BF_{TS} reflecting the normalized release time course of preexisting SV_{TS}s for 5, 10, and 20 Hz (*A*, *i*) and 50, 100, and 200 Hz (*A*, *ii*) stimulation. The black traces in *A*, *i* and *A*, *ii* represent the mean BF_{TS} for 50, 10, and 20 Hz. Individual BFs in A1 are indistinguishable from their mean. (*B*) Comparison of mean $M_{LS,RS}$ · $BF_{LS,RS}$ reflecting the release contribution by SVs that were not tightly docked prior to stimulation, i.e., the sum of preexisting SV_{LS}s and newly recruited SVs, for 5, 10, and 20 Hz (*B*, *i*) and 50, 100, and 200 Hz (*B*, *ii*) stimulation. The black traces in *B*, *i* and *B*, *ii* represent the mean $BF_{LS,RS}$ for 50, 10, and 20 Hz. (*C*) Comparison of BF_{LS} reflecting the normalized release time course of preexisting SV_{LS}s for 5, 10, and 20 Hz (*C*, *i*) and 50, 100, and 200 Hz (*C*, *ii*) stimulation. The black traces in *C*, *i* and *C*, *ii* represent the mean BF_{LS} for 50, 10, and 20 Hz. The inset compares the average BF_{LS} of 5–20 Hz (black) with a fit (red) as described in Neher and Taschenberger (51), which can be used to obtain an estimate for s_2. (*D*) Simulated time course of p_{fusion} during stimulus trains calculated according to Eq. **37**. The inset compares NTF-derived p_{fusion} estimates for the four initial eEPSCs with simulated values. (*E*) Scatter plot of NTF-derived M_{TS} versus $SP_{TS,rest}$ as predicted for 10 Hz eEPSC trains from m_1 and p_{fusion} according to *SI Appendix*, Eq. **31** (see also Fig. 2*F*). The slope of the regression line (1.09) indicates that the NTF-derived M_{TS} is on average ~9% larger.

rest ($SP_{TS,rest}$; average 961 ± 68 SVs). BF_{TS}s for f_{stim} of 5 to 20 Hz are strikingly similar when plotted against stimulus number (Fig. 3 A, *i*). They decay exponentially, indicating nearly constant p_{fusion} throughout trains at these frequencies. BF_{TS}s for ≥50 Hz deviate from this pattern (Fig. 3 A, *ii*). Their second values are larger, followed by a steeper decline, which is especially prominent for 200 Hz. Quantitative analysis of the time courses of p_{fusion} during trains as derived from BF_{TS} (51) indicates a small decrease in p_{fusion} for f_{stim} of 5 to 20 Hz, while for ≥50 Hz p_{fusion} increases, consistent with the observed net facilitation (Fig. 3D). Fig. 3B shows the time courses for quantal release contributed by those SVs, which were not in the TS state at stimulus onset (Release$_{LS,RS}$). Again, for 5 to 20 Hz, these time courses are strikingly similar when plotted against the stimulus number, indicating similar contributions to release regardless of ISI duration (Fig. 3 B, *i*).

Three-component NTF decomposition allows the separation of release contributions from preexisting SV_{LS}s (Fig. 3 C) and newly recruited SVs (*SI Appendix*, Fig. S2A) (51). Provided that all preexisting SV_{LS}s are consumed, the train quantal content (M_{LS}) associated with the BF_{LS} represents an estimate for SP_{LS} at rest ($SP_{LS,rest}$; average 1,078 ± 75 SVs). BF_{LS}s start at a value

very close to zero since preexisting SV_{LS}s cannot fuse during the first AP but first need to undergo the LS → TS transition. For all frequencies, BF_{LS}s quickly increase to a maximum value after 2 to 3 APs before decaying exponentially to near zero. Again, the time courses of BF_{LS}s for 5 to 20 Hz are nearly indistinguishable (Fig. 3 C, *i*). As pointed out previously (51) and formally proven here (Eqs. **10–20**), this finding is consistent with a transition of a constant fraction (s_2) of SP_{LS} to SP_{TS} subsequent to an AP. The average BF_{LS} for 5 to 20 Hz trains can be used to obtain an estimate for s_2 (0.11; Fig. 3 C, *i*, Inset). The BF_{LS}s for 50 to 200 Hz reach higher peak values and thereafter approach zero with faster decay kinetics (Fig. 3 C, *ii*).

At steady state, the loss from SP_{TS} due to SV fusion is compensated by an equal number of SVs replenishing SP_{TS}. Because the time course of BF_{LS}s (Fig. 3 C, *i*) along with that of the release contributed by newly recruited SVs (Release$_{RS}$; *SI Appendix*, Fig. S2 A, *i*) are frequency invariant for f_{stim} = 5 to 20 Hz, subpool occupancies approach very similar steady-state values regardless of ISI durations. That way, SVs are supplied to SP_{TS} at a rate linearly increasing with f_{stim}, which yields a frequency invariant steady-state quantal content (m_{ss}) at 5 to 20 Hz (*SI Appendix*, Fig. S1 B–D) (52, 53). We refer to this feature as balanced priming.

In summary, NTF analysis provides initial estimates for key model parameters, such as initial p_{fusion}, the size of SP_{TS} and SP_{LS} at rest, and the fraction of SP_{LS} transferred to SP_{TS} subsequent to each AP. Several more parameter estimates can be derived from analytical expressions regarding model predictions for, e.g., PPR and steady-state release at low-frequency stimulation (Eqs. **22–25**). The time courses of the NTF components and their frequency invariance at 5 to 20 Hz suggests that APs act autonomously in this frequency range by releasing a certain fraction of SVs (p_{fusion}) and shifting the same amount between subpools, irrespective of the interval between successive stimuli.

Simulations using the basic model confirm predictions derived from NTF analysis. After initializing model parameters (Fig. 1*B*) as described, numerical simulations reproduce steady-state depression ($D_m = m_{ss}/m_1$) for $f_{stim} = 1$ to 20 Hz quite well (Fig. 4*D*,

dotted trace). We assumed for both priming rate constants k_1 and k_2 fixed resting values ($k_{1,rest}$ and $k_{2,rest}$) and linear slopes σ_1 and σ_2 to describe their Ca^{2+} dependence (Eqs. **4**, **5**). The resting sizes of subpools, $SP_{LS,rest}$ and $SP_{TS,rest}$, are thus determined by $k_{1,rest}$, b_1, $k_{2,rest}$, and b_2 (Eqs. **1–5**). During stimulation, SP_{TS} partially depletes, such that release decreases toward a steady-state value m_{ss} for which SV consumption and replenishment are at equilibrium (Fig. 4*D*). At $f_{stim} \leq 2$ Hz, this balance depends on all the parameters listed here, and increasing f_{stim} results in reduced replenishment and lower m_{ss} due to shorter ISIs (Fig. 4*D*). However, at f_{stim} of 5 to 20 Hz, the Ca^{2+}-dependent increments of k_1 and k_2 above their resting values dominate the priming rates. Unpriming, determined by b_1 and b_2, becomes negligible, and the net movement of SVs along the kinetic scheme occurs nearly exclusively in the priming direction. The resulting steady-state occupancy of subpools is dominated by p_{fusion} and the integrals
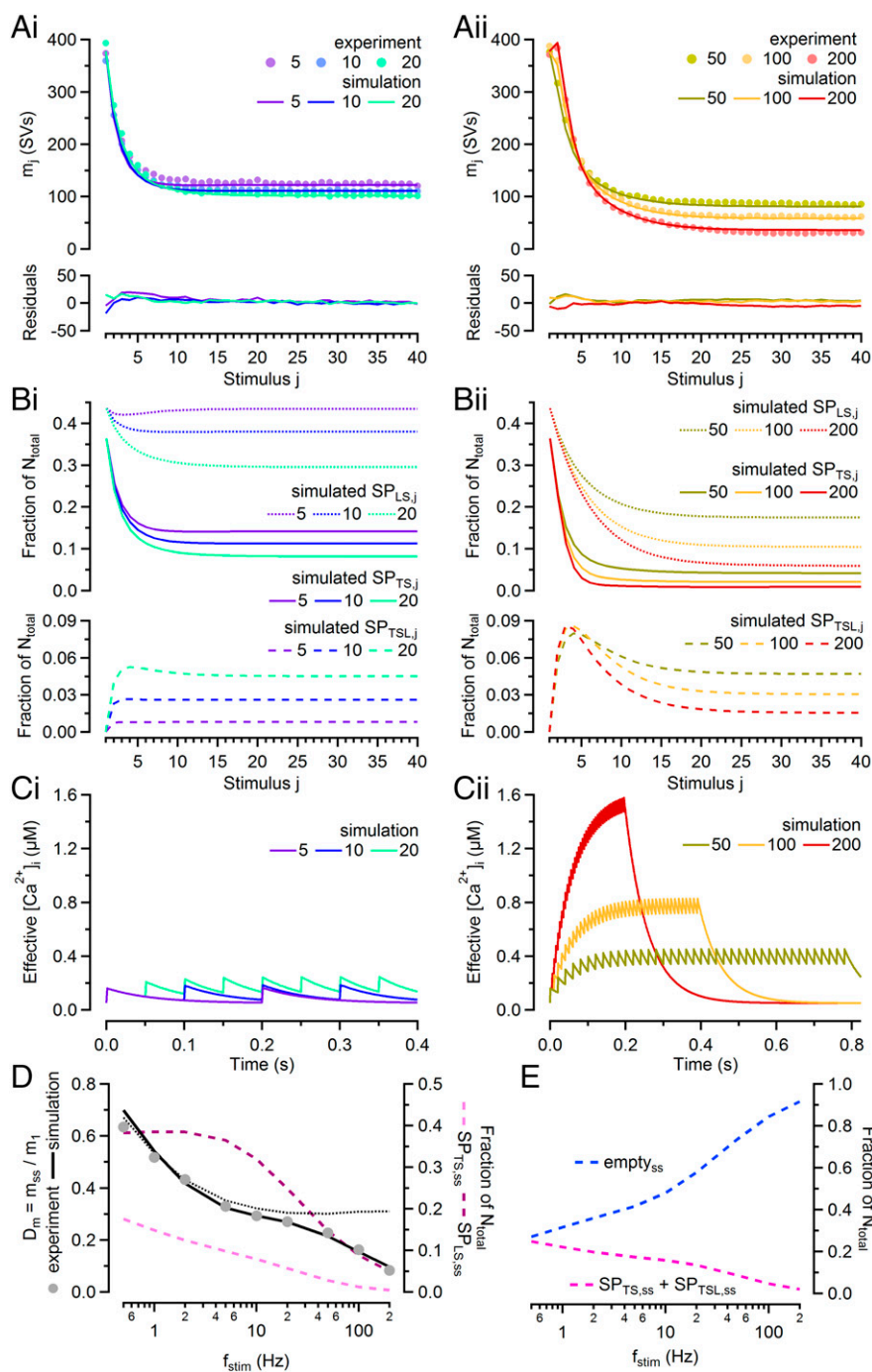

Fig. 4. Numerical simulations of STP in response to regular 5–200 Hz stimulus trains. (*A*) Experimental data (filled circles) and simulated m_j values (lines) plotted against stimulus j for 5, 10, and 20 Hz (*A, i*) and 50, 100, and 200 Hz (*A, ii*) trains. Residuals are shown in the small panels below. (*B*) Time course of subpool occupancies immediately before AP arrival for SP_{LS} and SP_{TS} (*Top*) and SP_{TSL} (*Bottom*) measured in fractions of total number of release sites (N_{tot}) for 5, 10, and 20 Hz (*B, i*) and 50, 100, and 200 Hz (*B, ii*) trains. Note different scaling of upper and lower panels. (*C*) Time course of the effective [Ca^{2+}] regulating priming (Eq. **6**) at high time resolution (1 ms) shown for comparison for 5, 10, and 20 Hz (*C, i*) and 50, 100, and 200 Hz (*C, ii*) trains. Only the initial portion of the [Ca^{2+}] trains is shown in *C, i*. (*D*) Normalized steady-state depression ($D_m = m_{ss}/m_1$, left axis) and steady-state occupancy of SP_{TS} and SP_{LS} measured in fractions of N_{tot} (right axis) are plotted logarithmically as a function of f_{stim}. While the steady-state occupancy of SP_{TS} gradually declines with increasing f_{stim}, SP_{LS} does not substantially deplete at steady state for $f_{stim} \leq 10$ Hz. The dotted trace is the prediction for D_m by the basic model (Fig. 1*B*). This assumes a strictly linear $k_1([Ca^{2+}])$, an empty SP_{TSL} at AP arrival, and no contribution of $y(t)$ to facilitation, but a decreasing $z(t)$ (Eq. **37**), as suggested by NTF for low frequencies. The basic model agrees quite well with the measured D_m (filled circles) for f_{stim} up to 20 Hz but clearly deviates for higher stimulation frequencies. The extended model (Fig. 1*C*) accurately describes data up to 200 Hz. (*E*) Simulated total numbers of empty release sites (blue dashed line) and release sites occupied with a fusion-competent SV (i.e., $SP_{TS} + SP_{TSL}$, red dashed line) calculated immediately before AP arrival at steady state in fractions of N_{tot} are plotted as a function of f_{stim}.

s_1 and s_2 over priming rates between consecutive APs (Eqs. **10**, **12**, and **24**), the latter representing fractions of upstream sub-pools converted per AP to the corresponding downstream sub-pools. s_1 and s_2 are related to the slopes σ_1 and σ_2 according to Eq. **12**. They are independent of f_{stim} as long as $[Ca^{2+}]$ transients are short relative to ISIs and $[Ca^{2+}]$-dependent priming dominates. Thus, m_{ss} tends toward a nearly frequency-independent plateau value (Eq. **24**) since both SV consumption and replenishment increase approximately linearly with frequency (Fig. 4D, dotted trace) (52, 53). Recovery from STD following conditioning 10 Hz stimulation is well approximated by single exponentials with time constants ~4 s (54), which can be reproduced in simulations by appropriate selection of $k_{1,rest}$ and $k_{2,rest}$.

In sum, STP during 1 to 20 Hz eEPSC trains is well described by a balanced priming scheme, as suggested by NTF, in which each AP transfers constant fractions (s_1 and s_2) of SVs from the respective upstream subpools to SP_{LS} and SP_{TS} and releases an almost constant fraction of SP_{TS} ($= p_{fusion}$). This leads to increasing STD (decrease of m_{ss}) for increasing f_{stim} up to 5 Hz, when Ca^{2+}-independent rates and Ca^{2+}-dependent rates are of similar magnitude. However, for $f_{stim} \geq 5$ Hz, when Ca^{2+}-dependent rates dominate, occupancy of both SP_{TS} and SP_{LS} is nearly frequency-independent, leading to relatively stable m_{ss} despite increasing f_{stim}. For $f_{stim} > 20$ Hz, however, this simple scheme will have to be extended, as detailed below.

Multiple mechanisms of release facilitation at stimulus frequencies ≥ 50 Hz. For $f_{stim} \geq 50$ Hz, large deviations are evident between experimental data and the basic model, which therefore needs to be extended to reproduce STP for f_{stim} up to 200 Hz. Inspection of the respective BFs reveals several changes at high f_{stim}: 1) p_{fusion} increases during trains, which is reflected in elevated second and third values of BF_{TS}, followed by a more rapid decay (Fig. 3 A, ii and D); 2) release contributed by those SVs not residing in TS prior to stimulation develops a peak around the third to fifth AP before decaying to steady-state levels lower than those at ≤ 20 Hz (Fig. 3 B, ii and C, ii); and 3) steady-state release contributed by newly recruited SVs decreases with increasing f_{stim}, causing stronger steady-state depression (Figs. 3 B, ii and 4D and SI Appendix, Fig. S2 A, ii). These changes in synapse behavior, which are likely caused by a summation of $[Ca^{2+}]$ transients or an incomplete relaxation of internal states of the release machinery, prompted us to consider options for extending the basic sequential model (Fig. 1). This also required slight adjustments of the parameters used so far due to some overlap of high- and low-frequency features.

The increase in p_{fusion} is likely a consequence of Ca^{2+} current facilitation (55–57), which we simulated with an empirical model (Eqs. **37–41**) to reproduce NTF analysis (Fig. 3D). However, the increase in p_{fusion} alone is insufficient to account for the measured PPF (48). In addition, the early peak at 200 Hz of the release component, reflecting the contribution by SVs that were not tightly docked prior to stimulation (Release$_{LS,RS}$; Fig. 3 B, ii), is not reproduced by the basic sequential model. We therefore explored additional mechanisms potentially accounting for release facilitation and enhanced priming at higher f_{stim}.

For $f_{stim} \geq 50$ Hz, AP-induced global $[Ca^{2+}]$ transients are expected to summate (Fig. 4C). Hence, a supralinear relationship between priming and $[Ca^{2+}]$ might generate release facilitation. However, simulations agreed only poorly with experimental data when we implemented this mechanism. Considering recent ultrastructural data of presynaptic active zones (AZs), which document rapidly reversible priming (37), we felt compelled to introduce a labile tightly docked SV state (TSL) and corresponding SV$_{TSL}$s.

We assume that each AP converts a certain fraction κ of SP_{LS} into such labile SV$_{TSL}$s, which constitute SP_{TSL} (Eqs. **43–46**). They have the same p_{fusion} as SV$_{TS}$s, but their backward transition LS ← TSL (b_3) is more rapid than that of LS ← TS (b_2). Thus, they contribute "extra" release to all responses of a high-frequency train except eEPSC$_1$, thereby enhancing PPR. They do not contribute to release, when ISIs are much longer than their mean decay time ($1/b_3$). A labile priming state can be implemented in series between LS and TS or as a parallel branch to the LS → TS transition. For both options, model parameters can be found, which reproduce STP accurately, even when the p_{fusion} of SV$_{TS}$s and that of SV$_{TSL}$s are constrained to have the same value. For brevity, we describe here only the version with the parallel branch (Fig. 1C). SP_{TSL} and SP_{LS} are incremented and decremented, respectively, following each AP by an amount of $\kappa \cdot SP_{LS}$. During ISIs, the time courses of subpool occupancies including SP_{TSL} are described by the differential equations Eqs. **43–46**. By introducing this extension of a TSL together with the facilitation of p_{fusion}, we were able to numerically simulate release during high-frequency trains with high fidelity (Fig. 5B) and identified $\kappa = 16\%$ and $b_3 = 11.1$ s^{-1} as best parameters (SI Appendix, Table S2). These numbers are compatible with recent EM data showing that the number of SVs tightly docked at the AZ is transiently increased 5 ms after an AP but has relaxed back to normal when flash-freezing is initiated 100 ms later (37). Similar shifts between loosely and tightly docked SVs were reported to be associated with the induction of long-term potentiation (LTP) in hippocampal synapses (39) as well as with the beta-adrenergic modulation of parallel fiber LTP in the cerebellum (40).

In sum, release facilitation during the onset of high-frequency trains is generated by an increase in p_{fusion}, by a more rapidly replenished SP_{TS} due to accelerated priming when global $[Ca^{2+}]$ summates, and further by a transient filling-up of SP_{TSL}.

Saturation of priming causes increased steady-state depression during high activity levels. At $f_{stim} \geq 50$ Hz, a progressive decline of release after the second or third stimulus toward a lower m_{ss} is observed (Figs. 4 A, ii and 5B). It results in a saturation of the relationship between release rate and f_{stim} (SI Appendix, Fig. S1C). We explored two options for modeling this: 1) a saturation of the priming rate with increasing f_{stim} and 2) a delayed availability of recently used sites for SV docking by introduction of a refractory release site state (58, 59).

As the simplest case, we describe here the replacement of the linear relationship between k_1 and $[Ca^{2+}]$ (Eq. **4**) by a Michaelis—Menten (MM) type saturating relationship (Eq. **42**). This introduces a parameter $K_{0.5}$, which is the $[Ca^{2+}]$ at half-maximum k_1. A $K_{0.5}$ of 280 nM and a slight adjustment of other parameters adequately reproduces experimental data for all f_{stim} (Fig. 4A and SI Appendix, Table S2). Note that the simulated mean steady-state depression $D_m = m_{ss}/m_1$ exactly superimposes on experimental data over two orders of magnitude of f_{stim} (Fig. 4D).

Having established an adequate fit to the experimental data, we then examined time courses of model quantities such as effective $[Ca^{2+}]$, p_{fusion}, and subpool occupancies (Figs. 4 B and C and 3D). At 10 Hz stimulation, the occupancy of SP_{LS} remains relatively constant during the train, while SP_{TS} rapidly depletes to a level of 0.310 times its resting value. This decay is the major cause of observed STD at 10 Hz ($m_{ss}/m_1 = 0.312 \pm 0.013$). At 200 Hz stimulation, both SP_{LS} and SP_{TS} strongly deplete, and the depletion of the former is a consequence of the MM type saturation of k_1. Depletion of subpools is partially compensated by an increase in p_{fusion} (Fig. 3D). Model predictions for steady-state subpool sizes and depression are plotted against f_{stim} in Fig. 4 D
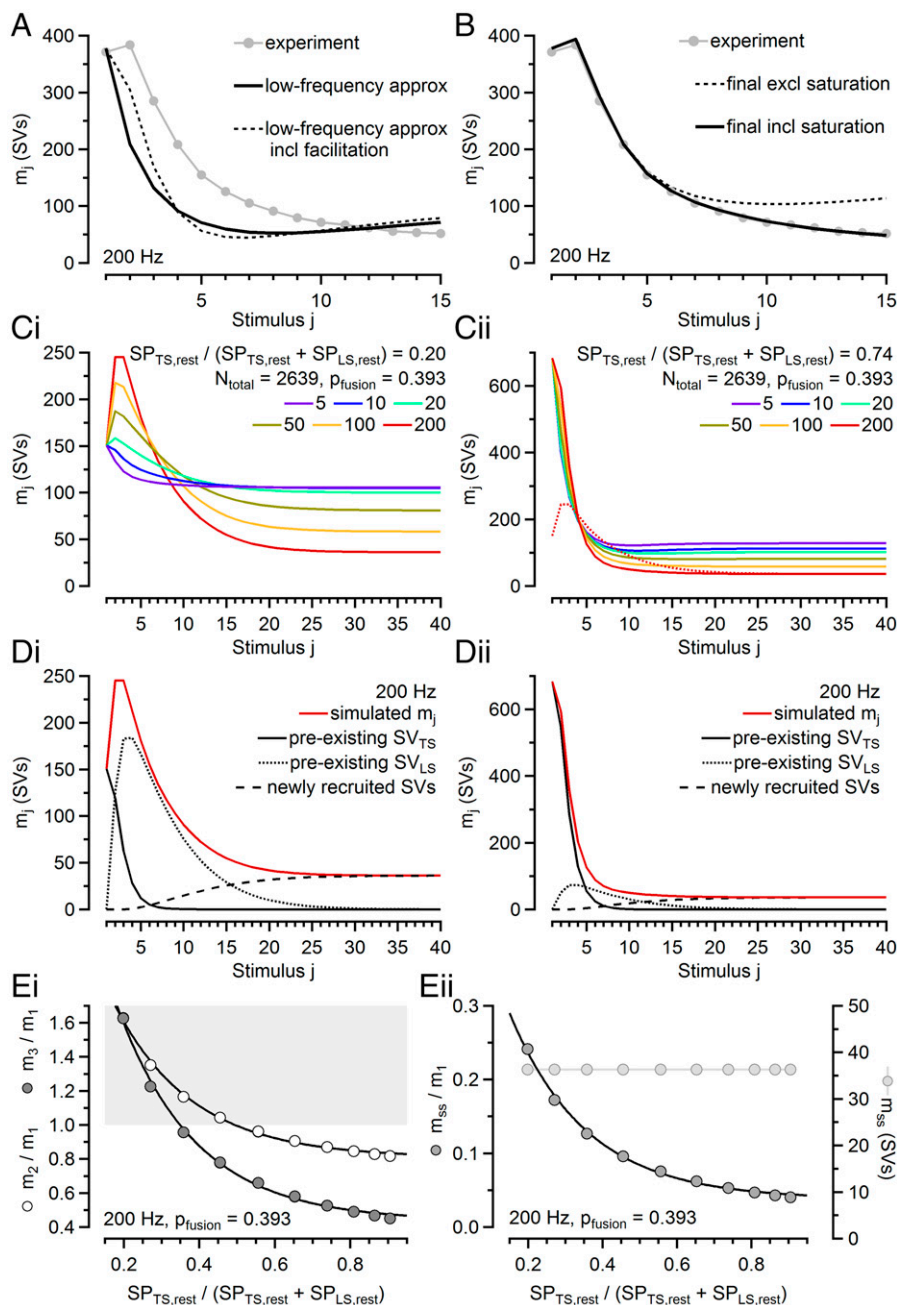
Fig. 5. Model predictions for facilitation, depression, and STP heterogeneity. (*A* and *B*) Experimental data (gray symbols and lines) and several model predictions (solid and dotted black lines) for 200 Hz eEPSC trains plotted superimposed versus stimulus index *j*. (*A*) The solid trace simulates release using the basic model with constant p_{fusion}, which is sufficient to account for experimental data up to f_{stim} = 20 Hz. The dashed trace includes the facilitation of p_{fusion}, as reported by NTF. Neither simulation can predict facilitation as experimentally observed for f_{stim} = 200 Hz. (*B*) Simulated 200 Hz trains after extending the model to include an additional release component mediated by SV_{TSL}s (dotted trace) and then further refined by assuming a saturating MM–type relationship between $[Ca^{2+}]$ and k_1 (solid trace), which accurately reproduces both PPF and STD. (*C*) Time courses of m_j in response to 5–200 Hz stimulus trains were simulated using standard values for all model parameters except for b_2, which was either increased (*C*, *i* and *D*, *i*) or decreased (*C*, *ii* and *D*, *ii*) such that the fraction $SP_{TS,rest}/(SP_{LS,rest} + SP_{TS,rest})$ was reduced to ~20% or enhanced to ~74%, respectively. The red dotted trace in *C*, *ii* represents the m_j time course for the simulated 200 Hz train shown in (*C*, *i*). (*D*) Simulated contributions to release during 200 Hz stimulation by preexisting SV_{TS}s (solid black) and by preexisting SV_{LS}s (dotted black) or newly recruited SVs (dashed black). Simulated total release (m_j, solid red) is shown for comparison. (*E*) Ratios m_2/m_1 (PPR) and m_3/m_1 (*E*, *i*) and relative steady-state depression (m_{ss}/m_1, *E*, *ii*) plotted versus the relative fraction of $SP_{TS,rest}$ for 14 simulations similar to those shown in (*C*) and (*D*). Either unpriming rate constant b_2 or priming rate constant k_2 for the LS ↔ TS transition were increased or decreased to generate relative SP_{TS} fractions in the range from 0.2 to 0.9. Note that $p_{fusion,1}$ was kept at 0.39 and standard values were used for all other model parameters. The gray shaded region in (*E*, *i*) indicates ratios >1.

and *E*, illustrating that the occupancy of SP_{LS} at steady state remains nearly constant and close to its initial value ($SP_{LS,rest}/N_{tot}$ = 0.44; *SI Appendix*, Table S2) for all f_{stim} ≤10 Hz (Fig. 4*D*). At 200 Hz stimulation, ~92% of the release sites are vacant at steady state (Fig. 4*E*). Simulated BFs are shown in *SI Appendix*, Fig. S3. Considering that NTF-derived BFs only provided initial guesses for various model parameters and that the latter were subsequently adjusted to optimize model predictions, the agreement between simulated (*SI Appendix*, Fig. S3) and NTF-derived (Fig. 3) BFs is reassuring.

Contributions of simulated NTF components to total release at 200 Hz are shown in Fig. 5*D* for two cases of differing initial subpool sizes. In comparison to the NTF decomposition of the measured mean 200 Hz eEPSC train (*SI Appendix*, Fig. S2 *B, ii*), Fig. 5*D* illustrates the differential contribution to release by SV_{LS}s following conversion into SV_{TS}s or SV_{TSL}s. In the two modeled synapses, this generates either pronounced net facilitation (Fig. 5 *D, i*) or faster and deeper depression (Fig. 5 *D, ii*) during 200 Hz trains.

As an alternative to a saturation of k_1, we introduced an empty but refractory release site state (ERS) (59–61). In this case, release sites vacated following SV fusion are not instantaneously converted into empty sites available for SV docking but shift into ERS, from which they become available for SV docking with a rate constant b_4 (Fig. 1 *C, ii*). A b_4 of 3.6 s^{-1} and a slight optimization of other parameters (*SI Appendix*, Table S2) resulted in model predictions, which for our standard parameter set were hardly discernible from those of the alternative approach of using an MM–type saturation of k_1 (Fig. 1 *C, i*). Significant differences between the two types of models were only observed when the occupancy of release sites at rest was reduced below 70%.

In sum, decreasing steady-state release at f_{stim} ≥50 Hz indicates a saturation of the pool replenishment process, which can be simulated equally well by a saturation of the Ca^{2+}-dependent acceleration of k_1 or else by introducing a refractory state of release sites, which after their use become available for new SV docking with first-order kinetics.
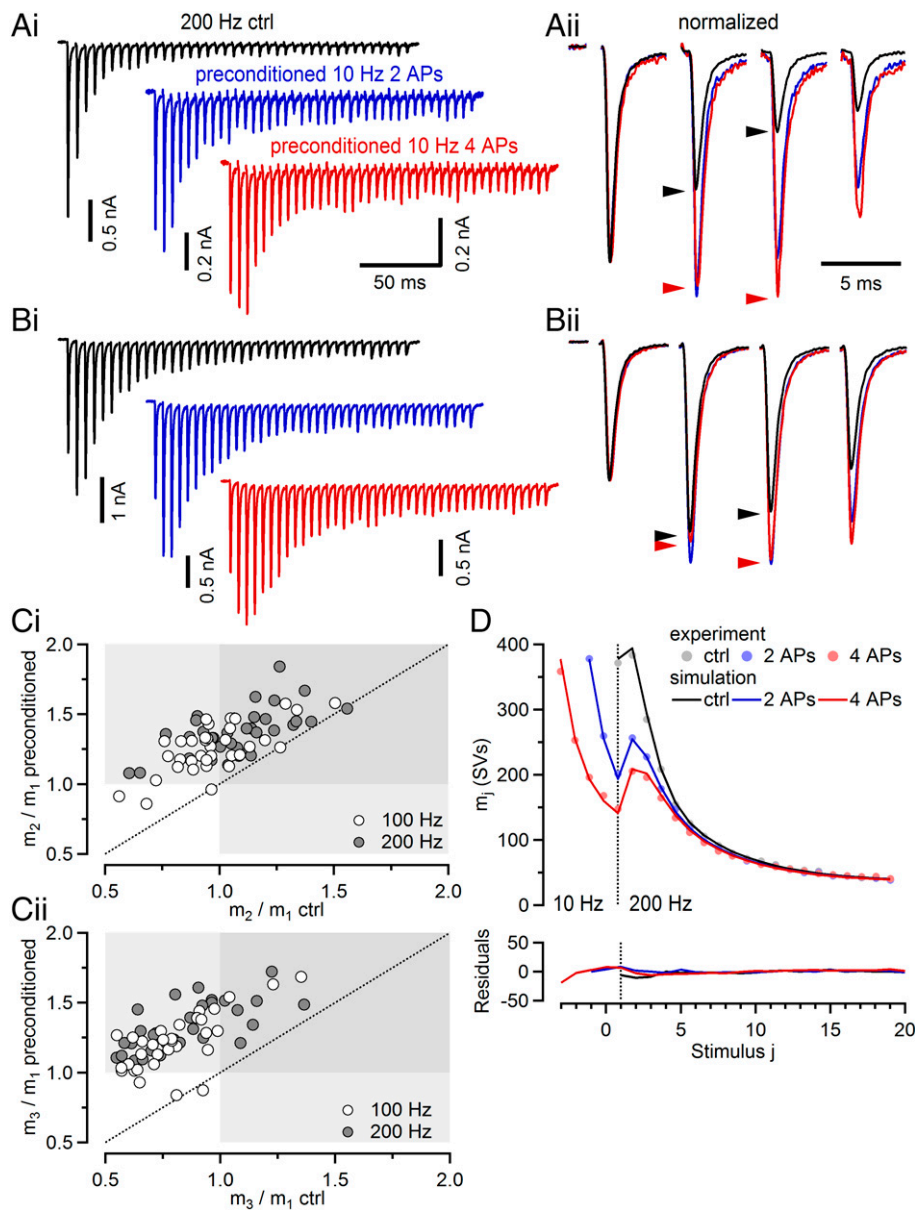
Fig. 6. Numerical simulations of STP in response to complex stimulus train patterns. (*A* and *B*) Sample 200 Hz eEPSCs recorded without (*A, i* and *B, i, Top*) or with 10 Hz preconditioning using 2 APs (*A, i* and *B, i, Middle*) or 4 APs (*A, i* and *B, i, Bottom*) in a strongly depressing (*A*) and a facilitating (*B*) synapse. We found that 10 Hz preconditioning converted depression into facilitation (*A*) and augmented existing facilitation (*B*). The initial four eEPSCs are shown after normalization to the peak of $eEPSC_1$ at an expanded timescale in (*A, ii*) and (*B, ii*). Time calibration bars in (*A, i*) and (*A, ii*) also apply to (*B, i*) and (*B, ii*). (*C*) Ratios m_2/m_1 (*C, i*) and m_3/m_1 (*C, ii*) measured after 10 Hz preconditioning with four APs plotted against the respective values obtained without preconditioning for all 35 synapses. Nearly all values lie above the unity line (dotted line). Values from both 200 Hz (filled circles) and 100 Hz (open circles) eEPSC trains are plotted. The gray shaded regions indicate ratios >1. (*D*) Simulated mean m_j values and experimental data including the conditioning responses plotted superimposed against stimulus index. Note the excellent agreement between experimental data (circles) and model predictions (solid lines). Residuals are plotted in the small panel (*Bottom*). The dotted vertical line marks the transition from 10 Hz conditioning to 200 Hz stimulation.

STP induced by conditioning stimulation and accelerated eEPSC recovery after depleting high-frequency trains. We next studied STP during 100 and 200 Hz eEPSC trains preconditioned by two or four APs delivered at 10 Hz, which causes pronounced release facilitation (62). Two examples for eEPSCs elicited by such a stimulus pattern, which are part of our standard NTF protocol, are shown in Fig. 6 *A* and *B*. Nearly all synapses showed more or less pronounced facilitation following 10 Hz preconditioning (Fig. 6*C*). The conditioning low-frequency stimulation depletes SP_{TS}, thereby exposing release facilitation at the onset of the high-frequency train due to the increase in p_{fusion} and because of the rapid conversion of SV_{LS}s to fusion-competent SV_{TS}s and SV_{TSL}s. The agreement between model predictions and average data was remarkable (Fig. 6*D*), considering that model parameters were largely determined on the basis of regular trains without preconditioning.

Finally, we tested the adequacy of the model against general features of eEPSC recovery from STD. As described above, the slow recovery after 10 Hz trains (54) can be reproduced by appropriate selection of $k_{1,rest}$ and $k_{2,rest}$. Following high-frequency stimulation, the recovery time course gains a rapid component (10, 63). After cessation of stimulation, three

processes overlap: a fast drop of p_{fusion} from a facilitated value back to $p_{fusion,1}$, a rapid decrease of SP_{TSL}, and an increase in SP_{TS} due to accelerated priming while $[Ca^{2+}]$ is elevated. The net result may either be an accelerated recovery or a transient decrease of synaptic strength, depending on the relative magnitudes of these processes and their kinetics (63). Ca^{2+}-accelerated priming is sensitive to the $[Ca^{2+}]$ buildup during trains, which is determined by the decay of individual $[Ca^{2+}]$ transients. At ≤ 20 Hz, when individual $[Ca^{2+}]$ transients do not overlap, the Ca^{2+}-dependence of k_1 and k_2 is relevant only in terms of the integral over transients—i.e., large transients with short duration are as effective in promoting priming as small transients with correspondingly longer duration. At ≥ 50 Hz, however, $[Ca^{2+}]$ transients summate. During stimulus trains, the increased rate of priming is balanced by release. After stimulation, refilling of partially depleted subpools continues at an increased rate until $[Ca^{2+}]$ has decayed back to $[Ca^{2+}]_{rest}$. Lengthening the decay of the $[Ca^{2+}]$ transient at the expense of its amplitude shifts a larger proportion of its priming-promoting effect into the early recovery time course. In simulations, the amount of accelerated eEPSC recovery can be adjusted to match experimental data by varying the value of the model parameter τ_{Ca} (*SI Appendix*, Table S2). An example for a

simulation of a stimulation pattern probing recovery is given in *SI Appendix*, Fig. S4. It should be noted, though, that experimentally observed decay time courses of global $[Ca^{2+}]$ are often biphasic (63–65). More detailed simulation of the Ca^{2+} signal relevant for controlling Ca^{2+}-dependent priming (i.e., effective $[Ca^{2+}]$) may improve the accuracy of model predictions.

Discussion

We analyzed eEPSC trains elicited by a wide range of presynaptic firing frequencies using a combined electrophysiological and modeling approach that was aided by nonnegative tensor factorization (51). Our analysis leads to four principal conclusions: 1) The experimental data are well compatible with a reversible priming process leaving ~20% of the release sites empty at rest while the remaining 80% are occupied by SVs, which are in one of two states (LS and TS, constituting subpools SP_{LS} and SP_{TS}, respectively). Both subpools equilibrate dynamically during presynaptic activity. 2) Different initial strength and diverse STP among calyx synapses is primarily due to variable SP_{LS} and SP_{TS} sizes at rest. Functional diversity across all synapses is consistent with relatively uniform p_{fusion} despite the large variability in initial strength. 3) Fusion-competent docked and primed $SV_{TS}s$ have a high p_{fusion} of ~0.4, consistent with the experimentally observed rapidly progressing STD during high-frequency stimulation. 4) Depending on f_{stim}, release occurs with different characteristics: *i*) At very low f_{stim}, the number of newly primed SVs per ISI is primarily determined by the resting values $k_{1,rest}$ and $k_{2,rest}$ and the activity-independent unpriming rate constants b_1 and b_2, such that increasing f_{stim} leads to decreasing steady-state release. *ii*) For intermediate f_{stim}, the Ca^{2+}-dependent increases of k_1 and k_2 above their resting values dominate. Because individual AP-evoked $[Ca^{2+}]$ transients do not overlap, each AP causes a forward transition of a constant fraction of SVs to the respective downstream subpool, irrespective of the interval between consecutive APs. Since each AP also triggers the fusion of an almost constant fraction of $SV_{TS}s$ (p_{fusion}), release tends toward a steady state, which is independent of frequency. *iii*) At high f_{stim}, additional kinetic features control release, such as an increase in p_{fusion} during trains, a speed-up of the priming process, a transiently increased occupancy of SP_{TSL} contributing to release facilitation, and a frequency-dependent decline of SP_{LS} and SP_{TS} occupancies and steady-state release rate. The first three features contribute to PPF, while the fourth causes STD. For modeling these high-frequency features, we had to extend the kinetic scheme (Fig. 1*C*).

Before discussing these aspects individually, we summarize the assumptions underlying the relationship between NTF analysis and model fitting.

Assumptions underlying NTF decomposition of quantal release during eEPSC trains. NTF-based decomposition of eEPSC trains into release components rests on the following four assumptions: 1) Contributions to release are nonnegative. 2) Release can be decomposed into distinct components, each of which represents the contribution by SVs that had been in a certain state prior to stimulation (LS, TS, or undocked). 3) Forward transition rates of SV priming and fusion strongly dominate over backward transition rates in the f_{stim} range used for NTF analysis (5–200 Hz). 4) Heterogeneity among synapses with respect to initial strength and STP characteristics is primarily due to differences in the relative abundance of $SV_{LS}s$ and $SV_{TS}s$ at rest—i.e., individual synapses are endowed with variable fractions of docked SVs equipped with a mature release machinery.

NTF decomposes release into components constrained only by nonnegativity. Therefore, a given NTF fit result should not be regarded as a unique solution but rather as one of many options consistent with both the mean eEPSC trains of all synapses examined and their variability. Provided that the above-mentioned assumptions hold, NTF analysis delivers good initial guesses for model parameters, which together with trial-and-error parameter optimization leads to a number of important conclusions discussed in the following sections.

A simple equation for estimating p_{fusion} from low-frequency stimulation-induced STD. When analyzing 5 to 20 Hz eEPSC trains, 1) we postulated a linear relationship between $[Ca^{2+}]$ and the priming rate constants k_1 and k_2 and 2) we assumed that the integral over the AP-induced $[Ca^{2+}]$ transient is constant over that f_{stim} range. This results in balanced priming, with the average SV recruitment rate and the average release rate both being proportional to f_{stim} at steady-state conditions. However, the same conclusions can be reached with much-less-specific assumptions. Any process by which an AP causes the release of a certain fraction of fusion-competent SVs and also the transition of a certain fraction of immature SVs to the fusion-competent state will reach a frequency-independent steady state. This would hold even if recruitment were nonlinearly dependent on $[Ca^{2+}]$ as long as $[Ca^{2+}]$ transients summated only little, such that each AP exerted its Ca^{2+}-dependent effect independently. This is largely the case for $f_{stim} \leq 20$ Hz. The choice of the exact time course of $[Ca^{2+}]$ transients has little influence on the outcome of simulations in this f_{stim} range, because different $[Ca^{2+}]$ transients will similarly influence release as long as they have the same time integral and decay within the ISI back to $[Ca^{2+}]_{rest}$ (Eqs. **12–14**).

The simplicity of conditions in the 5 to 20 Hz range (almost constant p_{fusion}, balanced loss and gain for SP_{LS}, fixed fraction of source pools transferred to downstream pools per AP) allowed us to derive a simple equation for calculating $p_{fusion,1}$ from the two easily measurable quantities, PPR $= m_2/m_1$ and $D_m = m_{ss}/m_1$. In this expression ($p_{fusion,1} = (1 - PPR)/(1 - D_m)$; Eq. **31**), the numerator recapitulates the standard use of PPR as an indicator for p_{fusion} (14, 66). The equation shows that $1 - PPR$ is actually a good estimate for p_{fusion} in case of complete steady-state depression ($D_m = 0$), provided that all simplification made in the derivation of Eq. **31** can be applied. However, SV pool depletion is usually far from complete for 5 to 20 Hz stimulation for which Eq. **31** holds. Thus, the correction by the denominator is substantial. The $p_{fusion,1}$ value obtained that way (0.43) is quite close to that derived from NTF analysis (0.39), the latter being used in all numerical simulations. Estimates for $p_{fusion,1}$ obtained according to Eq. **31** for individual synapses show less variability than corresponding m_1 values, which argues against $p_{fusion,1}$ being the main source of heterogeneity in synaptic strength among synapses (Fig. 2 *D–F*).

Comparison to p_{fusion} estimates derived by "traditional" quantal analysis methods. For a parallel priming and fusion scheme, the mean p_{fusion} is simply the weighted average of the p_{fusion} values pertaining to individual subpools of fusion-competent SVs. The mean p_{fusion} may be low, if "reluctantly releasing" SVs (9, 18, 67, 68) contribute to release. In the context of the sequential priming scheme discussed here, our analysis suggests a relatively high p_{fusion} for $SV_{TS}s$, which are the only fusion-competent SVs in resting synapses (Fig. 1*B*). In contrast, "traditional" estimates often calculate p_{fusion} as the ratio of the quantal content of a single synaptic response over the size of the pool of readily releasable SVs as determined by high-frequency stimulus trains

(69). Under the conditions described here, such stimulus trains deplete not only SP_{TS} but rather the sum $SP_{TS} + SP_{LS}$ due to the rapid LS → TS transition during high-frequency stimulation. Release probability as measured by traditional methods, therefore, is the product $p_{fusion} \cdot SP_{TS}/(SP_{LS} + SP_{TS})$, i.e., p_{fusion} times the probability of a docked SV being in the tightly docked state equipped with a mature release machinery (51). Considering a sequential priming scheme with only one fusion-competent state, p_{fusion} as determined by NTF or model fitting is a quantity strictly reflecting the SV fusion process, whereas p_{fusion} as determined by traditional methods is a quantity depending on the fusion process and on the priming equilibria at rest (70, 71) (*SI Appendix*, Fig. S5). This insight, if applicable to a given type of synapse, has strong implications for the interpretation of perturbations of synapse function by mutagenesis or pharmacological tools: An observed change in p_{fusion} is generally interpreted as a change in the fusion machinery, involving SNARE proteins, synaptotagmins, and complexins, or else reflecting changes in the microdomain $[Ca^{2+}]$ signal, which depends on Ca^{2+} currents, Ca^{2+} buffers, and coupling distances between voltage-gated Ca^{2+} channels (VGCCs) and Ca^{2+} sensors for SV fusion. According to our interpretation, an observed change in a traditional p_{fusion} estimate may well reflect changes in the LS ↔ TS equilibrium at rest, possibly involving presynaptic proteins such as Munc13, DOC2, CAPS, and synaptotagmin7 (27, 30). A modulatory influence of second messengers ($[Ca^{2+}]$, diacylglycerol (DAG), phosphatidylinositol 4,5-bisphosphate (PIP_2)) may show up in traditional quantal analysis as a change in p_{fusion} while NTF analysis reports it as a shift in the state of priming at rest. More precisely, the ratio of traditional p_{fusion} estimates over NTF-derived p_{fusion} estimates equals $SP_{TS}/(SP_{LS} + SP_{TS})$. Provided that conditions can be found for a given type of synapse for which Eq. **31** applies, $p_{fusion,1}$ can be estimated from PPR and D_m measured at a suitable frequency, and $SP_{TS,rest}$ is readily calculated as the ratio $m_1/p_{fusion,1}$. Assuming further that pool estimates obtained by analyzing depleting high-frequency eEPSC trains represent the sum $SP_{LS,rest} + SP_{TS,rest}$, then both $SP_{TS,rest}$ and $SP_{LS,rest}$ can be determined approximately from a small number of measurements without kinetic modeling or NTF analysis.

Synaptic facilitation and saturation of priming shape the release time course during high-frequency trains. For the f_{stim} range of 5 to 20 Hz, NTF provides quite stringent constraints on model parameters and eEPSC trains can be modeled as independently occurring release events largely unaffected by the ISI length. In contrast, our findings regarding STP mechanisms at $f_{stim} \geq 50$ Hz can be interpreted in alternative ways. Consecutive release events during 50 to 200 Hz stimulation may interact in various ways: through the summation of global $[Ca^{2+}]$ transients, through kinetic limitations such as the saturation of priming rates, and via synergies, e.g., a supralinear dependence of p_{fusion} on increasing local $[Ca^{2+}]$ due to VGCC facilitation. Together, these interactions shape synaptic facilitation and depression, which can be prominently observed during 200 Hz stimulation during which initial and transient net facilitation of eEPSCs is followed by strong depression toward a small m_{ss} (Fig. 4 *A, ii*). To illustrate how a given interaction influences the release time course, we compare in Fig. 5 *A* and *B* experimental data with model predictions. The basic model, which is sufficient for describing low frequency–induced STP, fails to reproduce the 200 Hz data (Fig. 5*A*, solid trace). Allowing p_{fusion} to vary as predicted by NTF analysis and reproduced by the empirical *y*–*z* formalism (Fig. 3*D* and Eq. **37**) is insufficient to explain the extent of PPF (Fig. 5*A*, dotted trace): Not only is m_2 far below the

experimental value, but m_3 and m_4 are even more so. Adding the contribution of the TSL state to release leads to an adequate fit up to m_8 (Fig. 5*B*, dotted trace). However, it predicts subsequent release to rebound due to enhanced priming by the gradual buildup of $[Ca^{2+}]$ during 200 Hz trains. Only the additional implementation of k_1 saturation reproduces the experimentally observed release time course over the entire stimulation period (Fig. 5*B*, solid trace). Alternatively, the experimental data can be simulated satisfactorily by the introduction of an ERS into the model (Fig. 1 *C, ii* and *SI Appendix*, Table S2).

Comparison with parallel release models and past work on the calyx of Held synapse. Previous studies of calyx synapses using long step depolarizations or presynaptic Ca^{2+} uncaging for triggering SV fusion identified two kinetically distinct release components mediated by two SV pools differentially coupled to presynaptic VGCCs (72, 73): a slowly releasing pool (SRP) and a fast-releasing pool (FRP) (68, 74, 75). The FRP accounts for the majority of SVs fusing during AP-evoked release, while the contribution of the SRP is only minor (76). The FRP may thus be regarded as the calyx–synapse equivalent of the readily releasable pool as estimated with AP trains in other synapses. Notably, SP_{LS} and SP_{TS} are not congruent with SRP and FRP but rather represent a subdivision of the FRP.

In a sequential priming scheme as used here, kinetic components may not be readily linked to specific state transitions. In contrast, it is intuitively easy to consider two kinetic components as independent contributions of two SV populations. Their intuitive tangibility explains why parallel kinetic schemes are particularly attractive. Previously, we thus described some of the STP features, based on a smaller data set, in terms of a parallel model consisting of a rapidly releasing SV subpool (called superprimed SVs) and a slowly releasing one (called normally primed SVs) (16) in analogy to studies at cultured hippocampal synapses (77). Superprimed SVs share many properties with SV_{TS}s. Normally primed SVs in the context of a parallel kinetic scheme represent fusion-competent SVs, albeit with a low fusogenicity (16). However, in the framework of the sequential scheme, proposed here, only SV_{TS}s are considered fusion competent. Therefore, and in view of the morphological evidence for "loose" and "tight" coupling, we consciously avoided the previous terminology.

Release mediated by SV_{TS}s has many features in common with the release contributed by so-called preprimed SVs, postulated for glutamatergic synapses in the cortex (78). Likewise, the sequential scheme proposed for cerebellar parallel fiber-molecular–layer interneuron synapses (22, 23) is related to the sequential scheme favored here: Both assume reversible priming steps in sequence, postulate certain fractions of upstream SV pools being transferred to a downstream pool upon AP arrival, and separate contributions to release in terms of the state of SVs prior to stimulation.

The identity of the $[Ca^{2+}]$ signal regulating the priming rate. The identity of the $[Ca^{2+}]$ signal that regulates priming (effective $[Ca^{2+}]$; *SI Appendix*, Table S2) is not unequivocally established. For f_{stim} of 5 to 20 Hz, each AP shifts certain constant fractions of SVs from one pool to the next downstream pool, which suggests a local, short-lived action of the effective $[Ca^{2+}]$ transients. If Ca^{2+}-dependent priming depends on Munc13, then a local $[Ca^{2+}]$ transient similar to that triggering SV fusion may be most relevant since Munc13 is an integral part of the AZ (79–81). On the other hand, the priming rate may saturate far below the peak $[Ca^{2+}]$ within such local domains if the respective Ca^{2+} sensors have high affinity. A complete

description of Ca^{2+}-dependent priming would have to consider endogenous Ca^{2+} buffers, which shape the time course of both the local domain $[Ca^{2+}]$ (82, 83) and the global volume–averaged $[Ca^{2+}]$ (64, 65, 84, 85). Here we chose to explore the influence of the $[Ca^{2+}]$ time course with the simplest possible model, characterized by an AP-induced increment in $[Ca^{2+}]$ proportional to the Ca^{2+} influx, followed by an exponential decay toward $[Ca^{2+}]_{rest}$ (Fig. 4C and SI Appendix, Table S2; Eq. **6**).

Sources and consequences of functional diversity among individual synapses. Fig. 5 C and D illustrates results of two simulations during which all model parameters were identical except for b_2, which was either increased or decreased to obtain a relative size of $SP_{TS,rest}$ of 20% (Fig. 5 C, i and D, i) or 74% (Fig. 5 C, ii and D, ii), respectively. We then systematically varied b_2 or $k_{2,rest}$ (Fig. 5E) and plotted the ratios m_2/m_1 and m_3/m_1 for simulated 200 Hz eEPSC trains as a function of the relative size of $SP_{TS,rest}$ (Fig. 5 E, i). We observed negative correlations, reminiscent of the relationship between PPR and the measured m_1 (Fig. 2D). A similar negative correlation was found between the simulated steady-state depression m_{ss}/m_1 and simulated $SP_{TS,rest}$ (Fig. 5 E, ii), illustrating that the magnitude of depression strongly depends on the relative abundance of $SV_{TS}s$ at rest. Fig. 5, ii also plots the absolute m_{ss} as a function of the relative size of $SP_{TS,rest}$ for simulated 200 Hz trains. As expected, m_{ss} is independent of the distribution of subpool sizes at rest since the Ca^{2+}-dependent parameters of the priming scheme dominate at high frequencies over the small rate constants, which set the distribution of states at rest.

Recent studies emphasize variations in coupling distances between presynaptic VGCCs and docked SVs as a predominant mechanism generating differences in p_{fusion} (for review, see reference 86). Our implementation of NTF analysis, on the other hand, postulates differences in the maturation state of primed SVs as the main source of intersynapse variability. It is well conceivable that a different NTF implementation, allowing several states or types of release sites to contribute to release, would come up with BFs for such states with different p_{fusion}, as one might expect for different coupling distances. However, our assignment of variability to differences in the maturation state of primed SVs is self-consistent in the sense that SI Appendix, Eq. **31**, when applied to individual synapses, confirms that variable p_{fusion} contributes little to their heterogeneity in synaptic strength (Fig. 2E). Of course, this does not preclude that spatial coupling between Ca^{2+} entry and the Ca^{2+} sensor for SV fusion may be heterogeneous.

The identification of molecular mechanisms causing functional synaptic diversity has recently attracted great attention due to the recognition of its importance for maximizing the capacity of information processing of neuronal networks (87) and of findings regarding the modulation of heterogeneity by astrocytes (88). Activity-induced changes in synaptic weight such as the redistribution of synaptic efficacy as a consequence of the induction of LTP (89) have emerged as important elements in the

description of the dynamics of neuronal networks (90, 91). However, the understanding of molecular mechanisms underlying such phenomena has lagged behind. Such redistribution is a basic property of our model, controlled by the distribution of SVs between primed states at rest and not involving a change in p_{fusion}. In Fig. 5C, we compare simulated time courses of the release of two synapses, one with a low ratio of SP_{TS}/SP_{LS} (Fig. 5 C, i) and the other with a high ratio of SP_{TS}/SP_{LS} (Fig. 5 C, ii) at rest. Despite the 4.5-fold difference in initial synaptic strength ($m_1 = 150$ vs. 684 SVs), the cumulative release during the first 20 stimuli at 200 Hz is quite similar (2,198 vs. 2,695 SVs; SI Appendix, Fig. S5B). This suggests changes in the LS ↔ TS equilibrium at rest and priming proteins, such as Munc13s, as the molecular basis of a redistribution of synaptic efficacy.

Materials and Methods

Preparation. Juvenile, posthearing onset (P14–16) Wistar rats of either sex were used. All experiments complied with the German Protection of Animals Act and with the guidelines for the welfare of experimental animals issued by the European Communities Council Directive. Acute brainstem slices were prepared similarly as previously described (16). See SI Appendix for details.

Electrophysiology. Whole-cell patch-clamp recordings were made from principal neurons of the medial nucleus of trapezoid body (MNTB) at room temperature as previously described (16). See SI Appendix for details.

Decomposition of quantal release into distinct components using NTF. NTF of eEPSC trains was performed similarly as previously described (51). See SI Appendix for details.

Simulation of AP-evoked SV release and short-term plasticity. We used a 2-step priming and fusion scheme (Fig. 1) to numerically simulate SV fusion and short-term plasticity. Details of the model and approximations for steady-state conditions that were used to constrain model parameters are given in SI Appendix.

Data, Materials, and Software Availability. Data sets shown in Fig. 4A and the program code required to re-create Fig. 4 B and C and all other numerical simulations are available at an open repository under https://doi.org/10.5281/zenodo.6818173 (92). All other data are included in the article and/or SI Appendix.

Author affiliations: [a]Emeritus Laboratory of Membrane Biophysics, Max Planck Institute for Multidisciplinary Sciences, 37077 Göttingen, Germany; [b]Department of Molecular Neurobiology, Max Planck Institute for Multidisciplinary Sciences, 37075 Göttingen, Germany; and [c]Cluster of Excellence "Multiscale Bioimaging", Georg August University, 37075 Göttingen, Germany

1. L. F. Abbott, W. G. Regehr, Synaptic computation. *Nature* **431**, 796–803 (2004).
2. W. G. Regehr, Short-term presynaptic plasticity. *Cold Spring Harb. Perspect. Biol.* **4**, a005702 (2012).
3. M. Bartos, I. Vida, P. Jonas, Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nat. Rev. Neurosci.* **8**, 45–56 (2007).
4. E. S. Fortune, G. J. Rose, Short-term synaptic plasticity as a temporal filter. *Trends Neurosci.* **24**, 381–385 (2001).
5. M. Verhage, J. B. Sørensen, SNAREopathies: Diversity in mechanisms and symptoms. *Neuron* **107**, 22–37 (2020).
6. R. S. Zucker, W. G. Regehr, Short-term synaptic plasticity. *Annu. Rev. Physiol.* **64**, 355–405 (2002).
7. M. Silva, V. Tran, A. Marty, Calcium-dependent docking of synaptic vesicles. *Trends Neurosci.* **44**, 579–592 (2021).
8. E. Neher, T. Sakaba, Multiple roles of calcium ions in the regulation of neurotransmitter release. *Neuron* **59**, 861–872 (2008).
9. K. L. Moulder, S. Mennerick, Reluctant vesicles contribute to the total readily releasable pool in glutamatergic hippocampal neurons. *J. Neurosci.* **25**, 3842–3850 (2005).
10. L. Y. Wang, L. K. Kaczmarek, High-frequency firing helps replenish the readily releasable pool of synaptic vesicles. *Nature* **394**, 384–388 (1998).
11. J. S. Lee, W. K. Ho, E. Neher, S. H. Lee, Superpriming of synaptic vesicles after their recruitment to the readily releasable pool. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 15079–15084 (2013).
12. D. Debanne, N. C. Guérineau, B. H. Gähwiler, S. M. Thompson, Paired-pulse facilitation and depression at unitary synapses in rat hippocampus: Quantal fluctuation affects subsequent release. *J. Physiol.* **491**, 163–176 (1996).
13. G. Grande, L. Y. Wang, Morphological and functional continuum underlying heterogeneity in the spiking fidelity at the calyx of Held synapse in vitro. *J. Neurosci.* **31**, 13386–13399 (2011).
14. L. E. Dobrunz, C. F. Stevens, Heterogeneity of release probability, facilitation, and depletion at central synapses. *Neuron* **18**, 995–1008 (1997).

15. A. Losonczy, L. Zhang, R. Shigemoto, P. Somogyi, Z. Nusser, Cell type dependence and variability in the short-term plasticity of EPSCs in identified mouse hippocampal interneurones. *J. Physiol.* **542**, 193–210 (2002).

16. H. Taschenberger, A. Woehler, E. Neher, Superpriming of synaptic vesicles as a common basis for intersynapse variability and modulation of synaptic strength. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E4548–E4557 (2016).

17. J. Trommershäuser, R. Schneggenburger, A. Zippelius, E. Neher, Heterogeneous presynaptic release probabilities: Functional relevance for short-term plasticity. *Biophys. J.* **84**, 1563–1579 (2003).

18. F. Doussau *et al.*, Frequency-dependent mobilization of heterogeneous pools of synaptic vesicles shapes presynaptic plasticity. *eLife* **6**, e28935 (2017).

19. S. Hallermann *et al.*, Bassoon speeds vesicle reloading at a central excitatory synapse. *Neuron* **68**, 710–723 (2010).

20. B. Pan, R. S. Zucker, A general model of synaptic transmission and short-term plasticity. *Neuron* **62**, 539–554 (2009).

21. T. Gabriel *et al.*, A new kinetic framework for synaptic vesicle trafficking tested in synapsin knock-outs. *J. Neurosci.* **31**, 11563–11577 (2011).

22. T. Miki, Y. Nakamura, G. Malagon, E. Neher, A. Marty, Two-component latency distributions indicate two-step vesicular release at simple glutamatergic synapses. *Nat. Commun.* **9**, 3943 (2018).

23. V. Tran, T. Miki, A. Marty, Three small vesicular pools in sequence govern synaptic response dynamics during action potential trains. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2114469119 (2022).

24. A. Martínez-Valencia, G. Ramírez-Santiago, F. F. De-Miguel, Dynamics of neuromuscular transmission reproduced by calcium-dependent and reversible serial transitions in the vesicle fusion complex. *Front. Synaptic Neurosci.* **13**, 785361 (2022).

25. J. S. Dittman, Unc13: A multifunctional synaptic marvel. *Curr. Opin. Neurobiol.* **57**, 17–25 (2019).

26. D. Fasshauer, W. K. Eliason, A. T. Brünger, R. Jahn, Identification of a minimal core of the synaptic SNARE complex sufficient for reversible assembly and disassembly. *Biochemistry* **37**, 10354–10362 (1998).

27. E. A. Prinslow, K. P. Stepien, Y. Z. Pan, J. Xu, J. Rizo, Multiple factors maintain assembled trans-SNARE complexes in the presence of NSF and αSNAP. *eLife* **8**, e38880 (2019).

28. Y. Lai *et al.*, Molecular mechanisms of synaptic vesicle priming by Munc13 and Munc18. *Neuron* **95**, 591–607.e10 (2017).

29. T. Xu *et al.*, Inhibition of SNARE complex assembly differentially affects kinetic components of exocytosis. *Cell* **99**, 713–722 (1999).

30. E. He *et al.*, Munc13-1 and Munc18-1 together prevent NSF-dependent de-priming of synaptic vesicles. *Nat. Commun.* **8**, 15915 (2017).

31. T. Miki *et al.*, Actin- and myosin-dependent vesicle loading of presynaptic docking sites prior to exocytosis. *Neuron* **91**, 808–823 (2016).

32. T. Sakaba, Two Ca^{2+}-dependent steps controlling synaptic vesicle fusion and replenishment at the cerebellar basket cell terminal. *Neuron* **57**, 406–419 (2008).

33. J. R. Kobbersmed *et al.*, Rapid regulation of vesicle priming explains synaptic facilitation despite heterogeneous vesicle: Ca^{2+} channel distances. *eLife* **9**, e51032 (2020).

34. D. Zenisek, J. A. Steyer, W. Almers, Transport, capture and exocytosis of single synaptic vesicles at active zones. *Nature* **406**, 849–854 (2000).

35. V. N. Murthy, C. F. Stevens, Reversal of synaptic vesicle docking at central synapses. *Nat. Neurosci.* **2**, 503–507 (1999).

36. S. Chang, T. Trimbuch, C. Rosenmund, Synaptotagmin-1 drives synchronous Ca^{2+}-triggered fusion by C_2B-domain-mediated synaptic-vesicle-membrane attachment. *Nat. Neurosci.* **21**, 33–40 (2018).

37. G. F. Kusick *et al.*, Synaptic vesicles transiently dock to refill release sites. *Nat. Neurosci.* **23**, 1329–1338 (2020).

38. E. Neher, N. Brose, Dynamically primed synaptic vesicle states: Key to understand synaptic short-term plasticity. *Neuron* **100**, 1283–1291 (2018).

39. J. H. Jung, L. M. Kirk, J. N. Bourne, K. M. Harris, Shortened tethering filaments stabilize presynaptic vesicles in support of elevated release probability during LTP in rat hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2018653118 (2021).

40. R. Martín *et al.*, β-adrenergic receptors/epac signaling increases the size of the readily releasable pool of synaptic vesicles required for parallel fiber LTP. *J. Neurosci.* **40**, 8604–8617 (2020).

41. R. Fernández-Busnadiego *et al.*, Cryo-electron tomography reveals a critical role of RIM1α in synaptic vesicle tethering. *J. Cell Biol.* **201**, 725–740 (2013).

42. C. Imig *et al.*, The morphological and molecular nature of synaptic vesicle priming at presynaptic active zones. *Neuron* **84**, 416–431 (2014).

43. V. Scheuss, H. Taschenberger, E. Neher, Kinetics of both synchronous and asynchronous quantal release during trains of action potential-evoked EPSCs at the rat calyx of Held. *J. Physiol.* **585**, 361–381 (2007).

44. C. H. Yang, W. K. Ho, S. H. Lee, Postnatal maturation of glutamate clearance and release kinetics at the rat and mouse calyx of Held synapses. *Synapse* **75**, e22215 (2021).

45. H. Taschenberger, R. M. Leão, K. C. Rowland, G. A. Spirou, H. von Gersdorff, Optimizing synaptic architecture and efficiency for high-frequency transmission. *Neuron* **36**, 1127–1143 (2002).

46. H. Taschenberger, V. Scheuss, E. Neher, Release kinetics, quantal parameters and their modulation during short-term depression at a developing synapse in the rat CNS. *J. Physiol.* **568**, 513–537 (2005).

47. S. Iwasaki, T. Takahashi, Developmental regulation of transmitter release at the calyx of Held in rat auditory brainstem. *J. Physiol.* **534**, 861–871 (2001).

48. M. Müller, F. Felmy, R. Schneggenburger, A limited contribution of Ca^{2+} current facilitation to paired-pulse facilitation of transmitter release at the rat calyx of Held. *J. Physiol.* **586**, 5503–5520 (2008).

49. R. S. Zucker, Changes in the statistics of transmitter release during facilitation. *J. Physiol.* **229**, 787–810 (1973).

50. A. Cichocki, R. Zdunek, A. H. Phan, S. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation* (Wiley, Chichester, UK, 2009).

51. E. Neher, H. Taschenberger, Non-negative matrix factorization as a tool to distinguish between synaptic vesicles in different functional states. *Neuroscience* **458**, 182–202 (2021).

52. J. Turecek, S. L. Jackman, W. G. Regehr, Synaptic specializations support frequency-independent Purkinje cell output from the cerebellar cortex. *Cell Rep.* **17**, 3256–3268 (2016).

53. J. Turecek, S. L. Jackman, W. G. Regehr, Synaptotagmin 7 confers frequency invariance onto specialized depressing synapses. *Nature* **551**, 503–506 (2017).

54. H. von Gersdorff, R. Schneggenburger, S. Weis, E. Neher, Presynaptic depression at a calyx synapse: The small contribution of metabotropic glutamate receptors. *J. Neurosci.* **17**, 8137–8146 (1997).

55. J. G. Borst, B. Sakmann, Facilitation of presynaptic calcium currents in the rat brainstem. *J. Physiol.* **513**, 149–155 (1998).

56. M. F. Cuttle, T. Tsujimoto, I. D. Forsythe, T. Takahashi, Facilitation of the presynaptic calcium current at an auditory synapse in rat brainstem. *J. Physiol.* **512**, 723–729 (1998).

57. K. H. Lin, S. Oleskevich, H. Taschenberger, Presynaptic Ca^{2+} influx and vesicle exocytosis at the mouse endbulb of Held: A comparison of two auditory nerve terminals. *J. Physiol.* **589**, 4301–4320 (2011).

58. V. Haucke, E. Neher, S. J. Sigrist, Protein scaffolds in the coupling of synaptic exocytosis and endocytosis. *Nat. Rev. Neurosci.* **12**, 127–138 (2011).

59. N. Hosoi, M. Holt, T. Sakaba, Calcium dependence of exo- and endocytotic coupling at a glutamatergic synapse. *Neuron* **63**, 216–229 (2009).

60. Y. Hua *et al.*, Blocking endocytosis enhances short-term synaptic depression under conditions of normal availability of vesicles. *Neuron* **80**, 343–349 (2013).

61. T. Sakaba *et al.*, Fast neurotransmitter release regulated by the endocytic scaffold intersectin. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 8266–8271 (2013).

62. M. Müller, J. D. Goutman, O. Kochubey, R. Schneggenburger, Interaction between facilitation and depression at a large CNS synapse reveals mechanisms of short-term plasticity. *J. Neurosci.* **30**, 2007–2016 (2010).

63. N. Lipstein *et al.*, Munc13-1 is a Ca^{2+}-phospholipid-dependent vesicle priming hub that shapes synaptic short-term plasticity and enables sustained neurotransmission. *Neuron* **109**, 3980–4000.e7 (2021).

64. K. H. Lin, H. Taschenberger, E. Neher, Dynamics of volume-averaged intracellular Ca^{2+} in a rat CNS nerve terminal during single and repetitive voltage-clamp depolarizations. *J. Physiol.* **595**, 3219–3236 (2017).

65. M. Müller, F. Felmy, B. Schwaller, R. Schneggenburger, Parvalbumin is a mobile presynaptic Ca^{2+} buffer in the calyx of Held that accelerates the decay of Ca^{2+} and short-term facilitation. *J. Neurosci.* **27**, 2261–2271 (2007).

66. W. J. Betz, Depression of transmitter release at the neuromuscular junction of the frog. *J. Physiol.* **206**, 629–644 (1970).

67. L. G. Wu, J. G. Borst, The reduced release probability of releasable vesicles during recovery from short-term synaptic depression. *Neuron* **23**, 821–832 (1999).

68. T. Sakaba, E. Neher, Quantitative relationship between transmitter release and calcium current at the calyx of Held synapse. *J. Neurosci.* **21**, 462–476 (2001).

69. E. Neher, Merits and limitations of vesicle pool models in view of heterogeneous populations of synaptic vesicles. *Neuron* **87**, 1131–1142 (2015).

70. M. Tanaka, T. Sakaba, T. Miki, Quantal analysis estimates docking site occupancy determining short-term depression at hippocampal glutamatergic synapses. *J. Physiol.* **599**, 5301–5327 (2021).

71. M. S. Thanawala, W. G. Regehr, Presynaptic calcium influx controls neurotransmitter release in part by regulating the effective size of the readily releasable pool. *J. Neurosci.* **33**, 4625–4633 (2013).

72. Z. Chen, B. Das, Y. Nakamura, D. A. DiGregorio, S. M. Young Jr., Ca^{2+} channel to synaptic vesicle distance accounts for the readily releasable pool kinetics at a functionally mature auditory synapse. *J. Neurosci.* **35**, 2083–2100 (2015).

73. K. Wadel, E. Neher, T. Sakaba, The coupling between synaptic vesicles and Ca^{2+} channels determines fast neurotransmitter release. *Neuron* **53**, 563–575 (2007).

74. T. Sakaba, E. Neher, Calmodulin mediates rapid recruitment of fast-releasing synaptic vesicles at a calyx-type synapse. *Neuron* **32**, 1119–1131 (2001).

75. J. S. Lee, W. K. Ho, S. H. Lee, Actin-dependent rapid recruitment of reluctant synaptic vesicles into a fast-releasing vesicle pool. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E765–E774 (2012).

76. T. Sakaba, Roles of the fast-releasing and the slowly releasing vesicles in synaptic transmission at the calyx of Held. *J. Neurosci.* **26**, 5863–5871 (2006).

77. O. M. Schlüter, J. Basu, T. C. Südhof, C. Rosenmund, Rab3 superprimes synaptic vesicles for release: Implications for short-term synaptic plasticity. *J. Neurosci.* **26**, 1239–1246 (2006).

78. B. Gustafsson, R. Ma, E. Hanse, The small and dynamic pre-primed pool at the release site; A useful concept to understand release probability and short-term synaptic plasticity? *Front. Synaptic Neurosci.* **11**, 7 (2019).

79. H. Sakamoto *et al.*, Synaptic weight set by Munc13-1 supramolecular assemblies. *Nat. Neurosci.* **21**, 41–49 (2018).

80. S. Reddy-Alla *et al.*, Stable positioning of Unc13 restricts synaptic vesicle fusion to defined release sites to promote synchronous neurotransmission. *Neuron* **95**, 1350–1364.e12 (2017).

81. M. R. Karlocai *et al.*, Variability in the Munc13-1 content of excitatory release sites. *eLife* **10**, e67468 (2021).

82. E. Eggermann, P. Jonas, How the "slow" Ca^{2+} buffer parvalbumin affects transmitter release in nanodomain-coupling regimes. *Nat. Neurosci.* **15**, 20–22 (2011).

83. N. P. Vyleta, P. Jonas, Loose coupling between Ca^{2+} channels and release sensors at a plastic hippocampal synapse. *Science* **343**, 665–670 (2014).

84. T. Collin *et al.*, Developmental changes in parvalbumin regulate presynaptic Ca^{2+} signaling. *J. Neurosci.* **25**, 96–107 (2005).

85. O. Caillard *et al.*, Role of the calcium-binding protein parvalbumin in short-term synaptic plasticity. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13372–13377 (2000).

86. G. Bornschein, H. Schmidt, Synaptotagmin Ca^{2+} sensors and their spatial coupling to presynaptic Ca_v channels in central cortical synapses. *Front. Mol. Neurosci.* **11**, 494 (2019).

87. B. Barbour, N. Brunel, V. Hakim, J. P. Nadal, What can we learn from synaptic weight distributions? *Trends Neurosci.* **30**, 622–629 (2007).

88. M. Letellier *et al.*, Astrocytes regulate heterogeneity of presynaptic strengths in hippocampal networks. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E2685–E2694 (2016).

89. H. Markram, M. Tsodyks, Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* **382**, 807–810 (1996).

90. A. J. Watt, N. S. Desai, Homeostatic plasticity and STDP: Keeping a neuron's cool in a fluctuating world. *Front. Synaptic Neurosci.* **2**, 5 (2010).

91. G. A. Carpenter, B. L. Milenova, Redistribution of synaptic efficacy supports stable pattern learning in neural networks. *Neural Comput.* **14**, 873–888 (2002).

92. K.-H. Lin, H. Taschenberger, E. Neher, A sequential two-step priming scheme reproduces diversity in synaptic strength and short-term plasticity. Zenodo. https://zenodo.org/record/6818173. Deposited 11 July 2022.