Model Predictive Control of an Integrated Buck Converter for Digital SoC Domains in 65nm CMOS
Xun Sun, Akshat Boora, Rajesh Pamula, Chi-Hsiang Huang, Diego Peña-Colaiocco, and Visvesh S. Sathe
University of Washington, Seattle, WA (E-mail: xuns@uw.edu)

Abstract
This paper describes a digital control architecture for integrated voltage regulators (IVRs) that achieves time-optimal transient supply-voltage ($V_{dd}$) response under random load-current ($I_{load}$) fluctuation. Implementing low-complexity low-latency Model Predictive Control (MPC) is key to achieving a measured 2.49X settling-time ($t_{sett}$) improvement over optimally tuned Proportional Integral Differential (PID) control.

Introduction
Minimizing peak $V_{dd}$ dropout ($V_{drop}$) margins in response to $I_{load}$ transients is increasingly critical to SoC efficiency. Towards this goal, the increased adoption of adaptive clocking [1]-[2] motivates the additional objective of $t_{sett}$ minimization: the time taken to restore $V_{dd}$ to a reference ($V_{ref}$). Several techniques for aggressive $t_{sett}$ reduction including Time-Optimal Control (TOC) [3] have been proposed [4]. However, these techniques either pose stability analysis challenges, or work only for load-step current waveforms—not random loading. For this reason, PID control continues to dominate. Model Predictive Control (MPC) is a modern, widely adopted control technique capable of providing stable and optimal response under random disturbances. However, its computational complexity—requiring a constrained optimization problem to be solved at each time step—has thus far restricted its use to applications that can tolerate tens of microseconds of computational latency [6]: achieving sub-10ns loop delay latencies, frequently required of IVR control, remains elusive. This paper describes how control and datapath optimizations, combined with the computational speed in modern CMOS, realize an IVR buck regulator with optimal transient response.

MPC Buck Architecture
Fig. 1 shows the proposed MPC buck regulator test-chip architecture. A Solver module incorporates the sampled $V_{dd}$ at cycle $n$ ($V_{dd}[n]$) into an accurate (non-linear) predictive model that tracks and updates the impact of the applied duty cycle ($d[n]$) on inter-dependent buck state variables—cycle-averaged inductor current ($I_L$) and capacitor voltage ($V_{dd}$). Traditional MPC uses this predictive model to plan an optimal buck duty-cycle trajectory for the next $k$-cycles ($d^*[n+i]$), $i=0,1...k$ which minimizes any desired objective function subject to constraints. However, to realize IVR-feasible low-latency control of an (approximately) 2nd order buck, the proposed formulation sets $k=2$, and avoids constrained optimization instead solving a much simpler problem of constraint satisfaction: the Solver must plan a 2-cycle trajectory ($d^*[n]$, $d^*[n+1]$) so that $V_{dd}[n+2]=V_{ref}$ and $I_L[n+1]=I_{load}$, to satisfy the minimum $t_{sett}$ criterion under non-duty-cycle saturated conditions ($d[n]=0$ or 1) to achieve $V_{dd}$ recovery (Fig. 2).

MPC also achieves optimal response under practical worst-case droop conditions involving duty-cycle saturation. For a large $I_{load}$ step-up, such a response involves (1) maximally slewing $I_L$ ($d^*[n]=1$) beyond $I_{load}$ to stop and reverse $V_{dd}$ droop recovery before (2) slewing $I_L$ back toward $I_{load}$ ($d^*[n]=0$) at time $n_{max}$. Eventually, (3) $I_L=I_{load}$ at $n_{match}$, after which $I_L$ tracks $I_{load}$ (Fig. 3). Timing $n_{max}$ precisely is critical for a stable, minimal $t_{sett}$ response and is achieved by tracking the surplus charge $Q_N[n]$, delivered to the load if $I_L$ were maximally slewed back to $I_{load}$ at cycle $n$. Avoiding an under-damped response requires that $Q_N[n_{max}]-C_1(V_{ref}-V_{dd}[n])=0$, so that the surplus current provides just the right quantity of charge to restore $V_{dd}$ without overshoot. Continuously evaluating the stability constraint (Dataflow graph shown in Fig. 4) to identify $n_{max}$ allows stable, rapid MPC response even with random $I_{load}$ disturbances. In contrast, traditional TOC techniques sense the $I_{load}$ step magnitude and commit to a determined $n_{max}$, unable to adjust to further $I_{load}$ changes until $V_{dd}$ is achieved.

Although the proposed MPC architecture avoids computationally intensive constrained optimization, datapath optimization remains necessary to lower computational latency and power dissipation. First, only $d^*[n]$ is computed in cycle $n$, because the computation of $d^*[n+1]$ is speculative under random loading conditions. The dataflow graph of the $d^*[n]$ computation (Fig. 5) is optimized by migrating computations away from the critical path between $V_{dd}[n]$ and $d^*[n]$ to reduce both latency and power dissipation (Fig. 3).

Test Chip and Measurement Results
Fig. 6 details the MPC-buck architecture implemented in 65nm CMOS (Die photograph in Fig. 11). An off-chip 1µH inductor and 1µF capacitor were used by the converter which was clocked at 10MHz. The error between $V_{dd}$ and $V_{ref}$ is first quantized by a 5mV resolution comparator bank before being forwarded to the MPC controller. A Digital Pulse Width Modulator (DPWM) converts the controller output code ($d[8:0]$) to pulses $p$ and $n$. Delayed-clocked registers enable glitch-free same-cycle feedback while providing sufficient time for the comparator banks (2ns) and the MPC controller (4.5ns) to evaluate. The Solver consumes 0.7mW at 10MHz.

Fig. 7 shows measured MPC-buck step-up and -down responses to a 305mA/0.12ns $I_{load}$ step at $V_{dd}=1V$ and a 191mA/0.12ns step at 0.7V. The test chip includes a tunable PID controller for a baseline comparison. $V_{drop}$ and $t_{sett}$ parameters corresponding to an optimally tuned PID controller are also annotated. Despite non-idealities resulting from non-zero $d_{min}$ and non-uniform quantization, the system achieves near-time-optimal transient response. Compared to the baseline, MPC achieves marginal $V_{drop}$ improvement ($d[n]=0$ or 1 during droop/surge) in both designs but a 2.49X $t_{sett}$ reduction. MPC can be applied to other VR modalities (LDOS, SC-regulators) and further leverage area and speed benefits offered by advanced CMOS nodes. These improvements, obtained through enhanced control, are compatible and expected to be additive with other circuit and package technology advances [7].

MPC stability, whose analysis is well documented in the literature [8], is also experimentally validated in this work. The synthetic load was configured with a pseudo-random $I_{load}$ sequence in the 0-200mA range and the system transient response was measured (Fig. 8). Over 50ns (corresponding to over 500,000 $I_{load}$ transitions), stable $V_{dd}$ regulation was observed from the MPC-buck. Also shown is the stable transient $V_{dd}$ response to an $I_{load}$ pulse, where the step-down transition occurs during transient recovery of the preceding $I_{load}$ step-up.
Fig. 9. Fig.10 presents a comparison with related work.

References


Acknowledgments

Alvin Loke, Carlos Tokunaga for helpful discussions. This work was funded by SRC (TXACE, Task 2810.035) and NSF (CAREER).

Fig. 6: MPC Test chip architecture. Delayed clocks provide time for Comparator and Solver evaluation with glitch-free operation.

Fig. 7: Measured Vref response to a load-step for MPC and optimally tuned PID controllers. Both incur |n| saturation resulting in similar Vref waveforms.

Fig. 8: Measured MPC Vref waveforms under random loading conditions.

Fig. 9: Measured MPCI Vref response to Iload pulse with step-down occurring during τcrit.

Fig. 10: Comparison with related work.